Tackling Climate Change with Machine Learning

David Rolnick[1]*, Priya L. Donti[2], Lynn H. Kaack[3], Kelly Kochanski[4], Alexandre Lacoste[5], Kris Sankaran[6,7], Andrew Slavin Ross[9], Nikola Milojevic-Dupont[10,11], Natasha Jaques[12], Anna Waldman-Brown[12], Alexandra Luccioni[6,7], Tegan Maharaj[6,8], Evan D. Sherwin[2], S. Karthik Mukkavilli[6,7], Konrad P. Körding[1], Carla Gomes[13], Andrew Y. Ng[14], Demis Hassabis[15], John C. Platt[16], Felix Creutzig[10,11], Jennifer Chayes[17], Yoshua Bengio[6,7]

[1]University of Pennsylvania, [2]Carnegie Mellon University, [3]ETH Zürich, [4]University of Colorado Boulder, [5]Element AI, [6]Mila, [7]Université de Montréal, [8]École Polytechnique de Montréal, [9]Harvard University, [10]Mercator Research Institute on Global Commons and Climate Change, [11]Technische Universität Berlin, [12]Massachusetts Institute of Technology, [13]Cornell University, [14]Stanford University, [15]DeepMind, [16]Google AI, [17]Microsoft 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.[1] Storms, droughts, fires, and flooding have become stronger and more frequent [3]. 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 [4]. 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 planning for resilience and disaster management, given an understanding of climate and extreme events. Such a diversity of problems can be seen as an opportunity: there are many ways to have an impact.

*D.R. conceived and edited this work, with P.L.D., L.H.K., and K.K. Authors P.L.D., L.H.K., K.K., A.L., K.S., A.S.R., N.M-D., N.J., A.W-B., A.L., T.M., and E.D.S. researched and wrote individual sections. S.K.M., K.P.K., C.G., A.Y.N., D.H., J.C.P., F.C., J.C., and Y.B. contributed expert advice. Correspondence to drolnick@seas.upenn.edu.

[1]For a layman's introduction to the topic of climate change, see [1, 2].

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 good,[2] there remains the need for a concerted effort to identify how these tools may best be applied to tackle 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 strategies we highlight include climate mitigation and adaptation, as well as meta-level tools that enable other strategies. In order to maximize the relevance of our recommendations, we have consulted experts across many fields (see Acknowledgments) in the preparation of this paper.

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 climate models afford an exciting domain for interpretable ML (see §7). 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), tools for helping reduce personal energy consumption (see §10.2), 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 §2.4), techniques for automatically assessing the energy consumption of buildings in cities (see §3.1), 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.

[2] See the AI for social good movement (e.g. [5, 6]), ML for the developing world [7], the computational sustainability movement (e.g. [8–12], the American Meteorological Society's Committee on AI Applications to Environmental Science, and the field of Climate Informatics (`www.climateinformatics.org`) [13], as well as the relevant survey papers [14–16].

	Causal inference	Computer vision	Interpretable models	NLP	RL & Control	Time-series analysis	Transfer learning	Uncertainty quantification	Unsupervised learning
1 Electricity systems									
Enabling low-carbon electricity		•	•		•	•		•	•
Reducing current-system impacts		•				•		•	•
Ensuring global impact		•					•		•
2 Transportation									
Reducing transport activity		•				•		•	•
Improving vehicle efficiency		•			•				
Alternative fuels & electrification					•				•
Modal shift	•	•				•		•	
3 Buildings and cities									
Optimizing buildings	•				•	•	•		
Urban planning		•				•	•		•
The future of cities				•			•	•	•
4 Industry									
Optimizing supply chains		•			•	•			
Improving materials									•
Production & energy		•	•		•				
5 Farms & forests									
Remote sensing of emissions		•							
Precision agriculture		•			•	•			
Monitoring peatlands		•							
Managing forests		•			•	•			
6 Carbon dioxide removal									
Direct air capture									•
Sequestering CO_2		•						•	•
7 Climate prediction									
Uniting data, ML & climate science		•	•			•		•	
Forecasting extreme events		•	•			•		•	
8 Societal impacts									
Ecology		•					•		
Infrastructure					•	•		•	
Social systems		•				•			•
Crisis		•		•					
9 Solar geoengineering									
Understanding & improving aerosols						•		•	
Engineering a planetary control system					•			•	
Modeling impacts						•		•	
10 Individual action									
Understanding personal footprint	•			•	•	•			
Facilitating behavior change				•					•
11 Collective decisions									
Modeling social interactions			•		•				
Informing policy	•	•		•				•	•
Designing markets					•	•			•
12 Education				•	•				
13 Finance				•		•		•	

Table 1: Climate change solution domains, corresponding to sections of this paper, matched with selected areas of ML that are relevant to each.

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 strategies.

- *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 areas 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 applications that will have their primary impact after 2040. While extremely important, these may in some cases be less pressing than those which can help act on climate change in the near term.

- *Uncertain Impact* denotes applications where the impact on GHG emissions is uncertain (for example, the *Jevons paradox* may apply[3]) or where there is potential for undesirable side effects (*negative externalities*).

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 applications 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 strategies. 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 address 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.

[3]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.

Mitigation

1 Electricity Systems *by Priya L. Donti*

AI has been called the new electricity, given its potential to transform entire industries [17]. 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 [18–20].

Electricity systems[4] are responsible for about a quarter of human-caused greenhouse gas emissions each year [26]. Moreover, as buildings, transportation, and other sectors seek to replace GHG-emitting fuels (§2-3), demand for low-carbon electricity will grow. To reduce emissions from electricity systems, society must

- Rapidly transition to low-carbon[5] 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 CO_2-emitting power plants, since the transition to low-carbon power 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 (Fig. 1). Such contributions include accelerating the development of clean energy technologies, improving forecasts of demand and clean energy, improving electricity system optimization and management, and enhancing system monitoring. These contributions require a variety of ML paradigms and techniques, as well as close collaborations with the electricity industry and other 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 tackling 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 not instantaneously[6]). These two types of sources affect electricity systems differently, and so present distinct opportunities for ML techniques.

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

[4]Throughout 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 [21–25].

[5]We 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 or easily replenished energy sources such as the sun, wind, or water, but need not necessarily be carbon-free (as in the case of some biomass [27]). Similarly, not all low-carbon energy is renewable (as in the case of nuclear energy).

[6]Nuclear power plants are often viewed as inflexible since they can take hours or days to turn on or off, and are often left on (at full capacity) to operate as *baseload*. That said, nuclear power plants may have some flexibility to change their power generation for load-following and other electric grid services, as in the case of France [28].

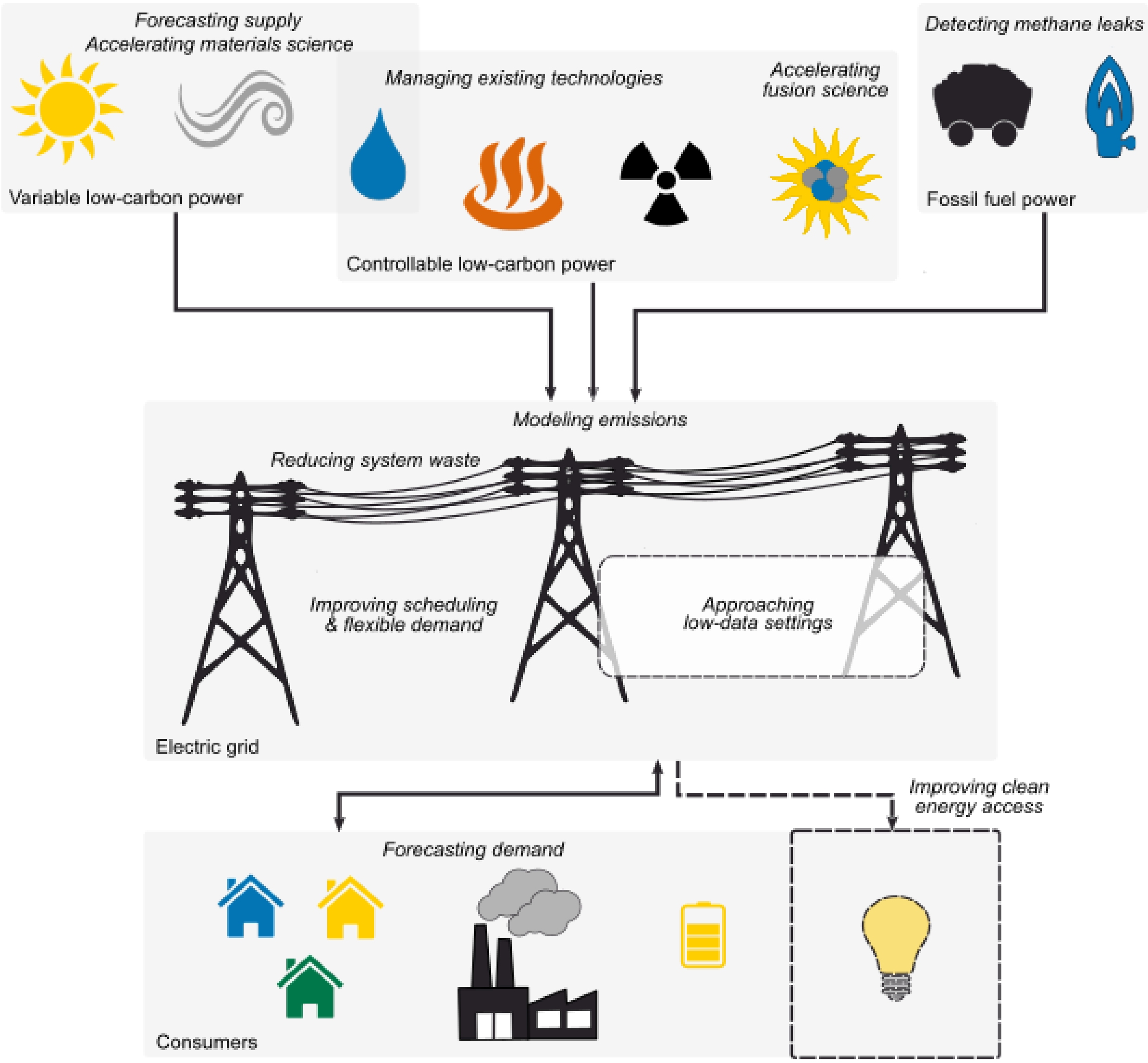

Figure 1: Selected opportunities to reduce GHG emissions from electricity systems using machine learning.

turbine, or other variable electricity generator, there is some mix of natural gas plants, storage, or other controllable sources ready to buffer changes in its output (e.g. when unexpected clouds block the sun or the wind blows less strongly than predicted). Today, this buffer is often provided by coal and natural gas plants run in a CO_2-emitting standby mode called *spinning reserve*. In the future, this role is expected to be played by energy storage technologies such as batteries (§2.3), pumped hydro, or power-to-gas [29].[7] 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.

[7] It is worth noting that in systems with many fossil fuel plants, storage may increase emissions depending on how it is operated [30, 31].

Forecasting supply and demand *High Leverage*

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 allow system operators to reduce their reliance on polluting standby plants and to proactively manage increasing amounts of variable sources. Better long-term forecasts can help system operators (and investors) determine where and when 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. ML can help on all these fronts.

To date, many ML methods have been used to forecast electricity supply and demand. These methods have employed historical data, physical model outputs, images, and even video data to create short- to medium-term forecasts of solar power [32–40], wind power [41–45], "run-of-the-river" hydro power [19], demand [46–49], or more than one of these [50, 51] 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 work has attempted to understand specific categories of demand, for instance by clustering households [52, 53] or by disaggregating electricity signals using game theory, optimization, regression, and/or online learning [54–56].

While much of this previous work has used domain-agnostic techniques, ML algorithms of the future will need to 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 (§7) and in hybrid physics-plus-ML modeling techniques [33–35]. 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 [57]. In addition to incorporating system physics, ML models should also directly optimize for system goals [58–60]. For instance, the authors of [58] use a deep neural network to produce demand forecasts that optimize for electricity scheduling costs rather than forecast accuracy; 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 generator should produce, interpretable 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 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 ML work in 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 every controllable generator should produce. This process is slow and complex, as it is governed by NP-hard optimization problems such as *unit commitment* and *optimal power flow* that must be coordinated across multiple time scales (from sub-second to days ahead). Further, scheduling will become even more complex as electricity systems include more storage, variable generators, 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 and improving the quality of optimization solutions. 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 approximate or simplify existing optimization problems [63–65], to find good starting points for optimization [66], or to learn from the actions of power system control engineers [67]. Dynamic scheduling [68, 69] and safe reinforcement learning could also be used to balance the electric grid in real time; 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 using energy storage, flexible demand, low-carbon generators, and other resources connected to the electric grid. One strategy is to explicitly design local control algorithms; for instance, recent work has controlled energy storage and *solar inverters* using supervised learning techniques trained on historical optimization data [70–73]. Another strategy is to let storage, demand, and generation respond to real-time prices[8] that reflect (for example) how emissions-intensive electricity currently is. In this case, ML can help both to design real-time prices and to respond to these prices. Previous work has used dynamic programming to set real-time electricity prices [78] and reinforcement learning to set real-time prices in more general settings [79]; similar techniques could be applied to create prices that instead optimize for GHG emissions. Techniques such as agent-based models [80–83], online optimization [84], and dynamic programming [85] can then help maximize profits for decentralized storage, demand, and generation, given real-time prices. In general, much more work is needed to test and scale existing decentralized solutions; barring deployment on real systems, platforms such as PowerTAC [86] can provide large-scale simulated electricity markets on which to perform these tests.

Accelerating materials science *High Leverage* *Long-term*

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 [87, 88]. ML 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 [88–90]. 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 [91]. (Additional applications of ML to batteries are discussed in §2.3.)

More generally in materials science, ML techniques including supervised learning, active learning, and generative models have been used to help synthesize, characterize, model, and design materials, as described in reviews [87, 92] and more recent work [93]. As discussed in [87], novel challenges for ML in materials science include coping with moderately sized datasets and inferring physical principles from trained models [94]. In addition to advancing technology, ML can inform policy for accelerated materials science; for instance, previous work has applied natural language processing to patent data to understand the solar panel innovation process [95]. 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 (§4.2) or create better CO_2 sorbents (§6.1).

Additional applications

There are many additional opportunities for ML to advance variable power generation. For instance, it is important to ensure that low-carbon variable generators produce energy as efficiently and profitably as

[8]For discussions and examples of different types of advanced electricity markets, see [74–77].

possible. Prior work has attempted to maximize electricity production by controlling movable solar panels [96, 97] or wind turbine blades [98] using reinforcement learning or Bayesian optimization. Other work has used graphical models to detect faults in rooftop solar panels [99] and genetic algorithms to optimally place wind turbines within a wind farm [100]. ML can also help control batteries located at solar and wind farms to increase these farms' profits, for instance by storing their electricity when prices are low and then selling it when prices are high; prior work has used ML to forecast electricity prices [101, 102] or reinforcement learning to control batteries based on current and historical prices [103].

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 often unknown to electricity system operators, previous work has used computer vision techniques on satellite imagery to generate size and location data for rooftop solar panels [104, 105]. Further, to ensure that the distribution system runs smoothly, recent work has employed techniques such as matrix completion and deep neural networks to estimate the state of the system when there are few sensors [106–108].

1.1.2 Controllable sources

Controllable low-carbon electricity sources can help achieve climate change goals while requiring very few changes to how the electric grid is run (since today's fossil fuel power plants are also controllable). ML can support existing controllable technologies while accelerating the development of new technologies such as nuclear fusion power plants.

Managing existing technologies

Many controllable low-carbon technologies are already commercially available; these technologies include geothermal, nuclear fission, and (in some cases[9]) dam-based hydropower. ML can provide valuable input in planning where these technologies should be deployed and can also help maintain already-operating power plants. For instance, recent work has proposed to use ML to identify and manage sites for geothermal energy, using satellite imagery and seismic data [110]. Previous work has also used multi-objective optimization to place hydropower dams in a way that satisfies both energy and ecological objectives [111]. Finally, ML can help maintain nuclear fission reactors (i.e., nuclear power plants) by detecting cracks and anomalies from image and video data [112] or by preemptively detecting faults from high-dimensional sensor and simulation data [113]. (The authors of [114] speculate that ML and high performance computing could also be used to help simulate nuclear waste disposal options or even design next-generation nuclear reactors.)

Accelerating fusion science ***High Leverage*** ***Long-term***

Nuclear fusion reactors [115] 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 [116]. 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 [117].

[9]Dam-based hydropower may produce methane, primarily due to biomass that decomposes when a hydro reservoir floods, but the amount produced varies between power plants [109].

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 [118–123] 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 [123]. 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 [124]; 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[10] 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 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-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 ***High Leverage*** ***Uncertain Impact***

Reducing emissions from fossil fuels is a necessary stopgap 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 [131, 132] or detect existing leaks [133–135], and there is a great deal of opportunity in this space to improve and scale existing strategies. In addition to leak detection, ML can help reduce emissions from freight transportation of solid fuels (§2), identify and manage storage sites for CO_2 sequestered from power plant flue gas (§6.2), and optimize power plant parameters to reduce CO_2 emissions. In all these cases, projects should be pursued with great care so as not to impede or prolong the transition to a low-carbon electricity system; ideally, projects should be preceded by system impact analyses to ensure that they will indeed decrease GHG emissions.

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 [136], bipartite ranking [137], and neural network-plus-clustering techniques [138] on electric grid data, and future work will need to improve and/or localize these approaches 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*

[10] Plasma simulation frameworks for tokamak reactors include RAPTOR [125, 126], ASTRA [127], CRONOS [128], PTRANSP [129], and IPS [130].

capture the emissions effects of small changes in demand at any given time. To inform consumers about marginal emissions factors, WattTime [139] estimates these factors in real time for the United States using regression-based techniques, and the electricityMap project [140] provides multi-day forecasts for Europe using ensemble models on electricity and weather data. 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 [141]. 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 methods produce 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.

1.3 Ensuring global impact

Much of the discussion around electricity systems often focuses on settings such as the United States with near universal electricity access and relatively abundant data. However, many places that do not share these attributes are still integral to tackling climate change [26] and warrant serious consideration. To ensure global impact, ML can help improve electricity access and translate electricity system insights from high-data to low-data contexts.

Improving clean energy access

Improving access to clean electricity can address climate change while simultaneously improving social and economic development [142, 143]. Specifically, clean electricity provided via electric grids, *microgrids*, or off-grid methods can displace diesel generators, wood-burning stoves, and other carbon-emitting energy sources. Figuring out what clean 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 to inform electrification initiatives [144]. ML and statistics can also help operate rural microgrids through accurate forecasts of demand and power production [145, 146], since small microgrids are even harder to balance than country-scale electric grids. Generating data to aid energy access policy and better managing energy access strategies are therefore two areas in which ML may have promising applications.

Approaching low-data settings ***High Leverage***

While ML methods have often been applied to grids with widespread sensors, 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 ML techniques such as transfer learning to translate insights from high-data to low-data settings (especially since all electric grids share the same underlying system physics). Developing data-efficient ML techniques will likely also be useful in low-data settings; for instance, in [147], 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 can also help generate information within low-data settings. For instance, recent work has estimated the layout of electricity grids in regions where they are not explicitly mapped, using computer vision on satellite imagery along with graph search techniques [148]. Companies have also proposed to use satellite imagery to measure power plant CO_2 emissions [149] (also see §5.1). Other recent work has modeled electricity consumption using regression-based techniques on cellular network data [150], which may prove useful in settings with many cellular towers but few electric grid sensors. Although low-data settings are generally underexplored by the ML community, electricity systems research in these settings presents opportunities for both innovative ML and climate change mitigation.

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 [20]. 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. Interpretable ML may enable stakeholders outside ML to better understand and apply models in real-world settings. Similarly, it will be important to develop hybrid ML models that explicitly account for system physics (see e.g. [147, 151–153]), 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 *by Lynn H. Kaack*

Transportation systems form a complex web that is fundamental to an active and prosperous society. Globally, the transportation sector accounts for about a quarter of energy-related CO_2 emissions [4]. In contrast to the electricity sector, however, transportation has not made significant progress to lower its CO_2 emissions [154] and much of the sector is regarded as hard to decarbonize [155]. 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 [156]. 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.[11] At present, more than two-thirds of transportation emissions are from road travel [156], but air travel has the highest emission intensity and is responsible for an increasingly large share. Strategies to reduce GHG emissions[12] from transportation include [156]:

- Reducing transport activity.
- Improving vehicle efficiency.
- Alternative fuels and electrification.
- Modal shift (shifting to lower-carbon options, like rail).

Each of these mitigation strategies offers opportunities for ML (Fig. 2). While many of us probably think of autonomous vehicles and ride-sharing when we think of transport and ML, these technologies have uncertain impacts on GHG emissions [160], potentially even increasing them. We discuss these disruptive technologies in §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 can help in implementing such interventions, for example by providing better demand forecasts. Typically, ML strategies 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 that 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 occurs 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 [161].

Understanding transportation data

Many areas of transportation lack data, and decision-makers often design infrastructure and policy with uncertain information. In recent years, new types of sensors have become available, and ML can turn this raw data into useful information. Traditionally, traffic is monitored with ground-based counters that are installed on selected roads. A variety of technologies are used, such as inductive loop detectors or pneumatic tubes.

[11] Carbon intensity is measured in grams of CO_2-equivalent per person-km or per ton-km, respectively.

[12] For general resources on how to decarbonize the transportation sector, see the AR5 chapter on transportation [156], and [157–159].

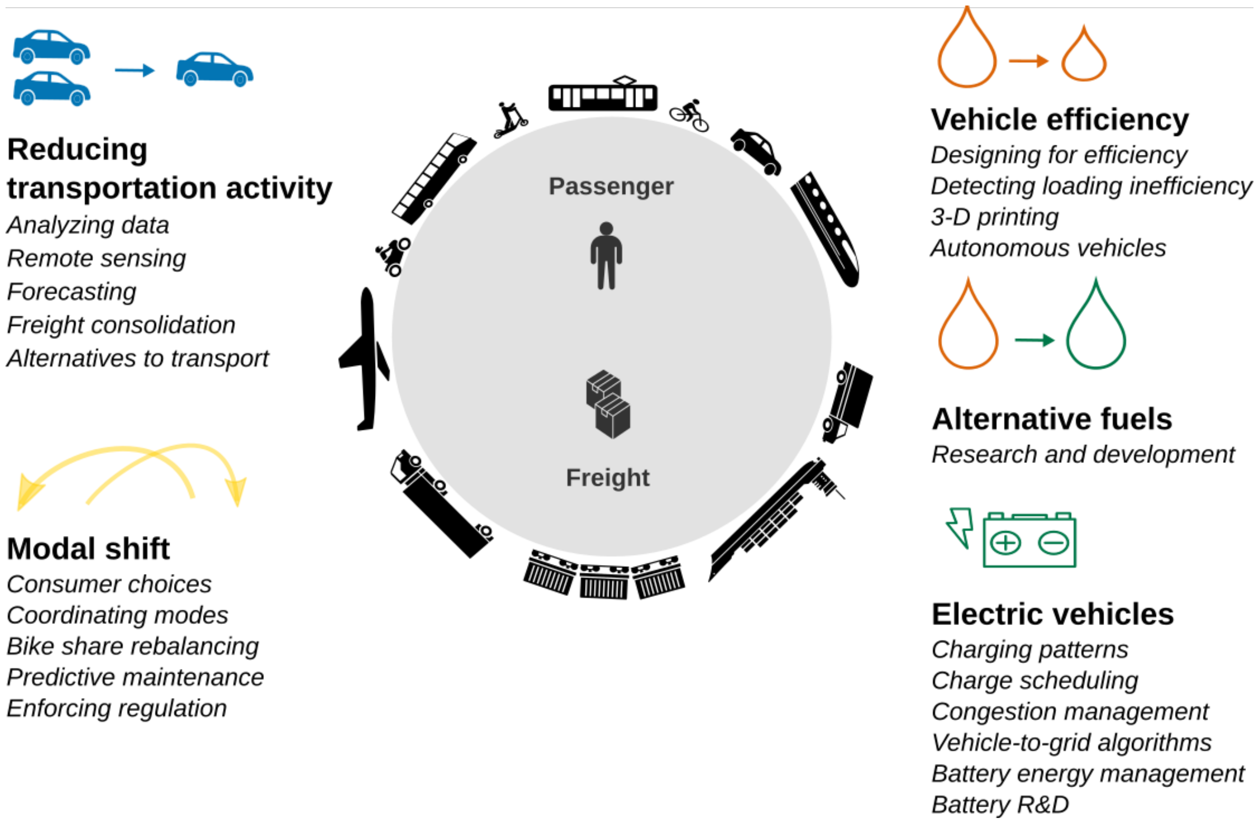

Figure 2: Selected strategies to mitigate GHG emissions from transportation using machine learning.

Traffic is sometimes monitored with video systems, in particular when counting pedestrians and cyclists, which can be automated with computer vision [162]. 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 [163–165]. 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 [166–169], and image counts can serve to estimate average vehicle traffic [170]. Similarly, ML methods can help with imputing missing data for precise bottom-up estimation of GHG emissions [171] and they are also applied in simulation models of vehicle emissions [172].

Modeling demand *High Leverage*

Modeling demand and planning new infrastructure can significantly shape how long trips are and which transport modes are chosen by passengers and shippers – for example, discouraging sprawl and creating new transportation links can both reduce GHG emissions. ML can provide information about mobility patterns, which is directly necessary for agent-based travel demand models, one of the main transport planning tools [173]. For example, ML makes it possible to estimate origin-destination demand from traffic counts [174], and it offers new methods for spatio-temporal road traffic forecasting – which do not always outperform other statistical methods [175] but may transfer well between areas [176]. Also, short-term forecasting of public transit ridership can improve with ML; see for example [177, 178]. ML is particularly relevant for deducing information from novel data – for example, learning about the behavior of public transit users from smart card data [179, 180]. Also, mobile phone sensors provide new means to understand personal travel demand and the urban topology, such as walking route choices [181]. Similarly, ML-based modeling of demand can help mitigate climate change by improving operational efficiency of modes that emit significant CO_2, such as aviation. ML can help predict runway demand and aircraft taxi time in order to reduce the excess fuel burned in the air and on the ground due to congestion in airports [182, 183].

Shared mobility *Uncertain Impact*

In the passenger sector, shared mobility (such as on-demand ride services or vehicle-sharing[13]), is undoubtedly disrupting the way people travel and think about vehicle ownership, and ML plays an integral part in optimizing these services (e.g. [184]). 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 [185]. On the other hand, shared mobility can lead to higher utilization of each vehicle, which means a more efficient use of materials [186]. 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, they can be used to predict if a customer decides to share a ride with other passengers from an on-demand ride service [187]. For decision-makers it is important 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. Some local governments are beginning to require data-sharing from these providers (see §3.3).

Car-sharing services using autonomous vehicles could yield GHG emission savings when they encourage people to use public transit for part of the journey [188] or with autonomous electric vehicles [189]. However, using autonomous shared vehicles alone could increase the total vehicle-miles traveled and therefore do not necessarily lead to lower emissions as long as the vehicles have internal combustion engines (or electrical engines on a "dirty" electrical grid) [190, 191]. 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, it is important that industry and public policy 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 [192, 193].

Freight routing and consolidation *High Leverage*

Bundling shipments together, which is referred to as freight consolidation, dramatically reduces the number of trips (and therefore the 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 [159]. 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. Many problem settings are addressed with methods from the field of operations research. There is evidence that ML can improve upon these methods, in particular mixed-integer linear programming [194]. Other proposed and deployed applications of ML include predicting arrival times or demand, identifying and planning around transportation disruptions [195], and clustering suppliers by their geographical location and common shipping destinations. Proposed planning approaches include designing allocation algorithms and

[13] In this section, we discuss shared cars; see §2.4 for bike shares and electric scooters.

freight auctions, and ML has for example been shown to help pick good algorithms and parameters to solve auction markets [196].

Alternatives to transport *Uncertain Impact*

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 [159]. ML can be a valuable tool for improving AM processes [197]. ML can also help to improve virtual communication [198]. If passenger trips are replaced by telepresence, travel demand can be reduced, as has been shown for example in public agencies [199] and for scientific teams [200]. 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 [201].

2.2 Improving vehicle efficiency

Most vehicles are not very efficient compared to what is technically possible: for example, aircraft carbon intensity is expected to decline by more than a third with respect to 2012, simply by virtue of newer models replacing aging jets [202]. Both the design of the vehicle and the way it is operated can increase the fuel economy. Here, we discuss how ML can help design more efficient vehicles and the impacts that autonomous driving may 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 [203].

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 [204]. Hybrid electric vehicles, which are more efficient than combustion engines alone, rely on power management methods that can be improved with ML [205]. Aerodynamic efficiency improvements need turbulence modeling that is often computationally intensive and relies heavily on ML-based surrogate models [206]. Aerodynamic improvements can not only be made by vehicle design but also by rearranging load. Lai et al. [207] 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 [159, 186]. ML is applied to improve those processes, for example through failure detection [208, 209] or material design [210].

Autonomous vehicles *Uncertain Impact*

Machine learning is essential in the development of autonomous vehicles (AVs), including in such basic tasks as following the road and detecting obstacles [211].[14] While AVs could reduce energy consumption – for example, by reducing traffic congestion and inducing efficiency through 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 [160, 212] and for broader impacts on mobility see [213].) 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 [214], though they come with regulatory challenges [215]. 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

[14] Providing details on the general role of ML for AVs is beyond the scope of this paper.

electrifying long-distance road freight [216]. 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. [217, 218] 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. [219] use data from smart phone sensors to identify driving behavior that leads to higher energy consumption in electric vehicles.

2.3 Alternative fuels and electrification

Electric vehicles *High Leverage*

Electric vehicle (EV) technologies – using batteries, hydrogen fuel cells, or electrified roads and railways – are regarded as a primary means to decarbonize transport. EVs can have very low GHG emissions – depending, of course, on the carbon intensity of the electricity. ML is vital for a range of different problems related to EVs. Rigas et al. [220] 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 [221] or optimization in hybrid EVs [205]), and to detect faults and lateral misalignment in wireless charging of EVs [222].

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) [223]. 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 [224].

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 [225] (see §1.1.1,1.2). There is much potential for ML to improve such vehicle-to-grid technology, for example with reinforcement learning [226], which can reduce GHG emissions from electricity generation. Vehicle-to-grid technology comes with private and social financial benefits. However, consumers are expected to be reluctant to agree to such services, as they might not want to compromise their driving range [227].

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 [228–235]. However, many models developed in academia are based on laboratory data that do not account for real-world factors such as environmental conditions [228–230]. 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 *Long-term*

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 [155]. Electrofuels [236], solar fuels 1.1.1, biofuels [237], hydrogen [238, 239], and perhaps natural gas [240] 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 [155]. 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 [159]. Fuels such as electrofuels and hydrogen can be produced using electricity-intensive processes and can be stored at lower cost than electricity. Thus, as a form of energy storage, these fuels could provide services to the electricity grid by enabling flexible power use 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 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 [241–243]. Seo et al. propose to conduct long-term travel surveys with online learning, which reduces the demand on respondents, while obtaining high data quality [244]. Sun et al. [245] 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. [246, 247]. Also in the freight sector, ML has been applied to analyze modal trade-offs, for example by imputing data on counterfactual mode choices [248].

Enabling low-carbon options ***High Leverage***

In order to incentivize 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 [249] and predicting track degradation [250].

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 [251]. ML can also help solve the bike sharing rebalancing problem, where shared bikes accumulate in one location and are lacking in other locations, by improving forecasts of bike demand and inventory [252]. Singla et al. [253] 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 [254]. Many emerging bike and scooter sharing services are dockless, which means that they are parked anywhere in public space and can block sidewalks [255]. ML has been applied to monitor public sentiment about such bike shares via tweets [256]. ML could also provide tools and information for regulators to ensure that public space can be used by everyone [257].

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 [258, 259]. 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 [159]. ML can contribute by improving predictions of the estimated time of arrival (for example of freight trains [260]) or the weight or volume of expected freight (for example for roll-on/roll-off transport – often abbreviated as Ro-Ro [261]). Intelligent transport systems of different modes could be combined and enable more efficient multimodal freight transportation [159].

Some modes with high GHG emissions, such as trucks, can be particularly cost-competitive in regions with lax enforcement of regulation, as they can benefit from overloading and not obeying labor or safety rules [159]. ML can assist public institutions with enforcing their regulations. For example, image recognition can help law enforcement detect overloading of trucks [262].

2.5 Discussion

Decarbonizing transport is essential to a low-carbon society, and there are numerous applications where ML can make an impact. This is because transportation causes a large share of 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 decarbonize transportation 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 strategies[15] could reduce emissions for existing buildings by up to 90% [264]. It is possible today for buildings to consume almost no energy [265].[16] Many of these energy efficiency measures actually result in overall cost savings [266] 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 [267, 268].

Two major challenges are heterogeneity and inertia. Buildings vary according to age, construction, usage, and ownership, so optimal strategies 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 [269]. Urban planning and public policy can play a major role in reducing emissions by providing infrastructure, financial incentives, or energy standards for buildings.[17]

Machine learning provides critical tools for supporting both building managers and policy makers in their efforts to reduce GHG emissions (Fig. 3). At the level of building management, ML can help select strategies that are tailored to individual buildings, and can also contribute to implementing those strategies 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). Finally, we consider how ML can help cities as a whole to transition to low-carbon futures (§3.3).

3.1 Optimizing buildings

In designing new buildings and improving existing ones, there are numerous technologies that can reduce GHG emissions, often saving money in the process [263–266, 270]. ML can accelerate these strategies by (i) modeling data on energy consumption and (ii) optimizing energy use (in *smart buildings*).

Modeling building energy

An essential step towards energy efficiency is making sense of the increasing amounts of data produced by meters and home energy monitors (see for example [271]). This can take the form of energy demand forecasts for specific buildings, which are useful for power companies (§1.1.1) and in evaluating building design and operation strategies [272]. Traditionally, energy demand forecasts are based on models of the physical structure of a building that are essentially massive thermodynamics computations. ML has the potential to speed up these computations greatly, either by ignoring physical knowledge of the building entirely [273, 274], by incorporating it into the computation [275], or by learning to approximate the physical model to reduce the need for expensive simulation (*surrogate models*) [276]. Learning how to transfer the knowledge gained from modeling one building to another can make it easier to render precise estimations of more

[15] 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) [263].

[16] There are even high-rise buildings, e.g. the Tower Raiffeisen-Holding NÖ-Vienna office, or large university buildings, e.g. the Technical University also in Vienna, that achieve such performance.

[17] 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`.

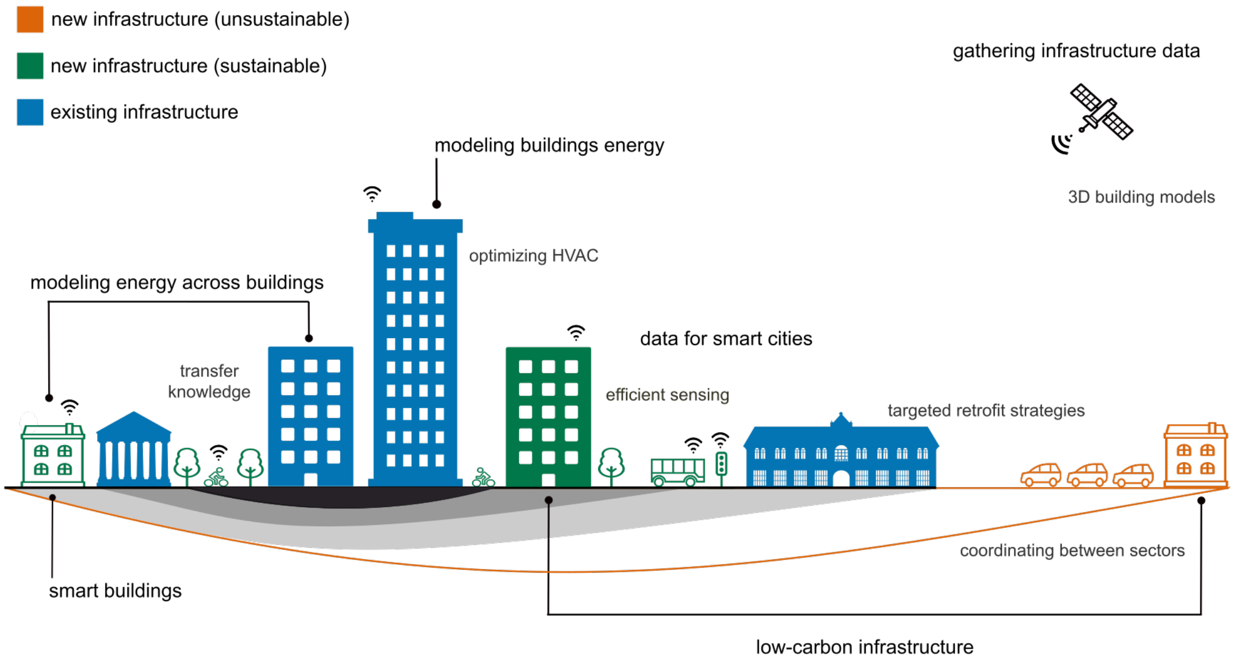

Figure 3: Selected strategies to mitigate GHG emissions from buildings and cities using machine learning.

buildings. For instance, Mocanu et al. [277] 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 can motivate behavioral changes. Promising ML approaches to this problem include hidden Markov models [278], sparse coding algorithms for structured prediction [279], harmonic analysis that picks out the "signatures" of individual appliances [280], and deep neural networks [281].

To verify the success or failure of energy efficiency interventions, statistical ML offers methods for causal inference. For example, Burlig et al. [282] 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 [283].

Smart buildings *High Leverage*

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 [284]. 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 [263]. There are several promising ways to enhance HVAC operating performance, each providing substantial 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 [285, 286] (see also §4.3). For control, Kazmi et al. [287] 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 [288]; ML approaches are well-suited to detect faults in these systems. Wang et al. [289] 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 [290].

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 [291]. ML can help these systems dynamically adapt to changes in occupancy patterns [292]. Moreover, occupancy detection itself represents an opportunity for ML algorithms, ranging from decision trees [293, 294] to deep neural networks [295] that take input from occupancy sensors [293], WiFi signals [295, 296], or appliance power consumption data [294].

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 [297]. For this, automated demand-side response [298] can respond to electricity prices, smart meter signals, or learned user preferences [299]. 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 [300].

While smart building technologies have the capability to significantly increase efficiency, we should note that there are potential drawbacks [301]. 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 [302]. Second, rebound effects are likely to happen in certain cases [303], leading to additional building energy consumption typically ranging between 10 and 20% [304]. 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 the actions of individual building managers [263]. Another concern in the case of widespread adoption of smart meters is the impact on mineral use and embodied energy use arising from their production [305]. Finally, smart home applications present security and privacy risks [306] 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. Where energy-use data on individual buildings exist, ML can be used to derive higher-level patterns. However, 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.

Modeling energy use across buildings

Urban Building Energy Models provide simplified information on the energy use of all buildings across a city. These are different from individual-building models, which model energy use of only specific buildings, but with finer details and temporal granularity (§3.1). While UBEMs have yet to be adopted at scale, they are expected to become fundamental for enabling localized action by city planners [307].[18] UBEMs can for example be used for planning and operating *district heating and cooling*, where a central plant supplies many households in a district. In turn, district heating and cooling reduces HVAC energy consumption and can provide flexible load [309], but it needs large amounts of data at the district level for implementation and operation.

UBEMs include features such as the location, geometries, and various other attributes of interest like building footprint, usage, material, roof type, immediate surroundings etc. ML can be used to held predict energy consumption from such features. 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 [310]. 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 [311], analyzed the effect of energy disclosure on the demand [312], and developed a system for ranking buildings based on energy efficiency [313]. Zhang et al. [314] matched residential energy consumption survey data with public use microdata samples to estimate residential energy consumption at the neighborhood level. Using five commonly accessible features of buildings and climate, Robinson et al. predict commercial building energy use across large American cities [315].

Beyond energy prediction, buildings' features can be used by ML algorithms to pinpoint which buildings have the highest retrofit potential. Simple building characteristics and surrounding environmental factors – both potentially available at scale – can be used [316, 317].

There have also been attempts to upscale individual-building energy models to the district scale. Using deep neural networks for hybrid ML-physical modelling, Nutkiewicz et al. provided precise energy demand forecasts that account for inter-building energy dynamics and urban microclimate factors for all buildings on a campus [318].

Gathering infrastructure data *High Leverage*

Specifics about building infrastructure can often be predicted using ML techniques. Remote sensing is key to inferring infrastructure data [105, 319–323] as satellite data[19] present a source of information that is globally available and largely consistent worldwide. For example, using remote sensing data, Geiß et al. [325] clustered buildings into types to assess the potential of district heat in a German town.

The resolution of infrastructure data ranges from coarse localization of all buildings at the global scale [319], to precise 3D reconstruction of a neighborhood [323]. It is possible to produce a global map of human settlement footprints at meter-level resolution from satellite radar images [319]. 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 [320]. Energy-relevant building attributes, such as the presence of photovoltaic panels, can also be retrieved from these images [105] (see §1.1.1). To generate 3D models, LiDAR data is often used to retrieve heights or classify buildings at city scale [321, 322], but its collection is expensive. Recent research showed that heights can be predicted even without such elevation data, as demonstrated by [326], 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 [323].

[18] The startup nam.R is developing a database of all school buildings in France to help inform retrofitting decisions, harmonizing vast amounts of open and proprietary data with ML [308].

[19] See [324] for a review of different sources of data and deep learning methods for processing them.

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) transportation, buildings, and economic activity. They handle such diverse issues as energy, water, waste, 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 [327], here we discuss only applications that are relevant to reducing GHG emissions.

Data for smart cities ***High Leverage***

Increasingly, important aspects of city life come with digital information that can make the city function in a more coordinated way. Habibzadeh et al. [328] 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.[20] Urban computing [330] is an emerging field looking at data analytics in urban spaces, and aiming to yield insights for data-driven policies. For example, clustering anonymized credit card payments makes it possible to model different communities and lifestyles – of which the sustainability can be assessed [331]. Jiang et al. provides a review of urban computing from mobile phone traces [332].[21] Relevant information on the urban space can also be learned from social media activity, e.g. on Twitter, as reviewed in [333, 334]. There are, however, numerous challenges in making sense of this wealth of data (see [335]), and privacy considerations are of paramount importance when collecting or working with many of these data sources.

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 [336]. ML can also distill information on urban issues related to climate change through web-scraping and text-mining, e.g. [256]. As discussed above (§3.2), ML can also be used to infer infrastructure data.

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 [337–339]. Similar techniques also help to reduce the amount of energy consumed during transmission itself [340].

Third, urban policy-making based on intelligent infrastructure faces major challenges with data management [341]. 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 [342, 343], data fusion [344], and ensemble learning [345].

Low-emissions infrastructure

When smart city projects are properly integrated into urban planning, they can make cities more sustainable and foster low-carbon lifestyles (see [340, 346, 347] for extensive reviews on this topic). Different types of infrastructure interact, meaning that planning strategies should be coordinated to achieve mitigation goals. For instance, urban sprawl influences the energy use from transport, as wider cities tend to be more car-oriented [348–350]. ML-based analysis has shown that the development of efficient public transportation

[20] Note that management of any such private data, even if they are anonymized, poses challenges [329].

[21] See `https://www.microsoft.com/en-us/research/project/urban-computing/` for more applications of urban computing.

is dependent on both the extent of urban sprawl and the local development around transportation hubs [351, 352].

Cities can reduce GHG emissions by coordinating between infrastructure sectors and better adapting services to the needs of the inhabitants. ML and AI can help, for example, to coordinate district heating and cooling networks, solar power generation, and charging stations for electric vehicles and bikes [347], and can improve public lighting systems by regulating light intensity based on historical patterns of foot traffic [353]. 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 [347].

Currently, most smart city projects for urban climate change mitigation are implemented in wealthier regions such as the United States, China, and the EU.[22] The literature on city-scale mitigation strategies is also strongly biased towards the Global North [354], while key mitigation challenges are expected to arise from the Global South [355]. Infrastructure models described in §3.2 could be used to plan low-carbon neighborhoods without relying on advanced smart city technologies. To transfer strategies across cities, it is possible to cluster similar cities based on climate-relevant dimensions [356, 357]. Creutzig et al. [349] 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 [356] using a latent class choice model, or of street networks [357] 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). Realizing the potential of data-driven urban infrastructure can advance mitigation goals while improving the well-being of citizens [264, 269, 358].

[22] See for example the European Union H2020 smart cities project https://ec.europa.eu/inea/en/horizon-2020/smart-cities-communities.

4 Industry *by Anna Waldman-Brown*

Industrial production, logistics, and building materials are leading causes of difficult-to-eliminate GHG emissions [155]. Fortunately for ML researchers, the global industrial sector spends billions of dollars annually gathering data on factories and supply chains [359] – 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 demonstrates considerable potential for reducing industrial GHG emissions under the following circumstances:

- When there is enough 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, and there are clear objective functions.
- When firms' incentives align with reducing emissions (for example, through efficiency gains, regulatory compliance, or high GHG prices).

In particular, ML can potentially reduce global emissions (Fig. 4) 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 [360–363]. However, it is worth noting that greater efficiency may increase the production of goods and thus GHG emissions (via the Jevons paradox) unless industrial actors have sufficient incentives to reduce overall emissions [364].

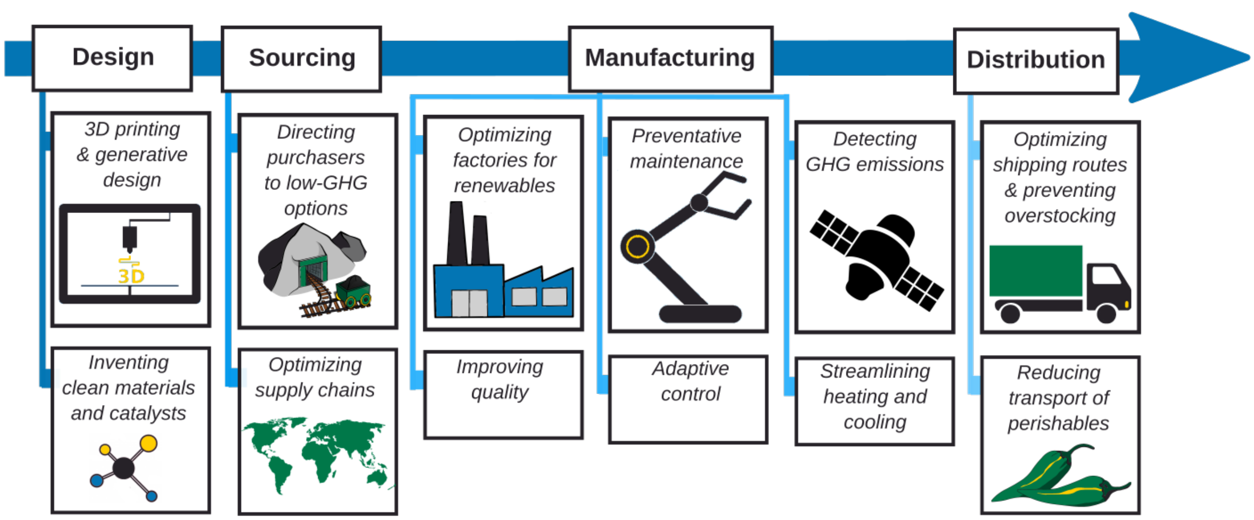

Figure 4: Selected opportunities to use machine learning to reduce greenhouse gas emissions in industry.

4.1 Optimizing 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 [365]. This how to prioritize low-carbon products for producers

indicates the complexity of today's globalized *supply chains*, i.e., the organizational processes and shipping networks that are required to bring a product from producer to final consumer. ML can help 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.) However, for many of these applications to reduce emissions, firms' financial incentives must also align with climate change mitigation through carbon pricing or other policy mechanisms.

Reducing overproduction *Uncertain Impact*

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 [366]. Global excess inventory in 2011 amounted to about $8 trillion worth of goods, according to the Council of Supply Chain Management Professionals [367]. This excess may be in part due to mis-estimation of demand, as the same organization noted that corporate sales estimates diverged from actual sales by an average of 40% [367]. ML may be able to mitigate these issues of overproducing and/or overstocking goods by improving demand forecasting [368, 369]. 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 [370]. 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 [371]. 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 [371, 372].

Recommender systems

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 [373, 374]. 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.

Reducing food waste *High Leverage*

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 [375]. In developing countries, 40% of food waste occurs 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 [375]. ML can help reduce food waste by optimizing delivery routes and improving demand forecasting at the point of sale (see §4.1), as well as improving refrigeration systems [376] (see §4.3). 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 [377].

4.2 Improving materials

Climate-friendly construction *High Leverage* *Long-term*

Cement and steel production together account for over 10% of all global GHG emissions [378]; the cement industry alone emits more GHGs than every country except the US and China [379]. ML can help minimize these emissions by reducing the need for carbon-intensive materials, by transforming industrial processes to run on low-carbon energy, and even by redesigning the chemistry of 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 reducing the resulting GHG emissions [360]. 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; ML and finite element modeling have been used to simulate the physical processes of 3D printing in order to improve the quality of finished products [380].

Assuming future advances in materials science, ML research could potentially draw upon open databases such as the Materials Project [381] and the UCI Machine Learning Repository [382] to invent new, climate-friendly materials [383]. Using semi-supervised generative models and concrete compression data, for example, Ge et al. proposed novel, low-emission concrete formulas that could satisfy desired structural characteristics [382].

Climate-friendly chemicals *High Leverage* *Long-term*

Researchers are also experimenting with supervised learning and thermal imaging systems to rapidly identify promising catalysts and chemical reactions [384, 385], 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.

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 [386]. To develop cleaner ammonia, chemists may be able to invent electrochemical strategies for lower-temperature ammonia production [386, 387]. Given the potential of ML for predicting chemical reactions [385], ML may also be able to help with the discovery of new materials for electrocatalysts and/or proton conductors to facilitate ammonia production.

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 emissions, 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 firms' cost-benefit calculations may dissuade them from retrofitting existing factories to run using low-carbon electricity or to save a few kilowatts [388–390]. Given the heterogeneity across industrial sectors and the secrecy of industrial data, firms will also need to tailor the requisite sensors and data analysis systems to their individual processes. ML will become a much more viable option for industry when factory workers can identify, develop, implement, and monitor their own solutions internally instead of relying upon outside experts [391]. The ML community can assist by building accessible, customizable industry tools tailored for people without a strong background in data science.

Adaptive control *High Leverage*

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. To reduce GHG emissions from HVAC systems, researchers have suggested combining optimization-based control algorithms with ML techniques such as image recognition, regression trees, and time delay neural networks [392, 393] (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 HFC emissions and reducing cooling costs [361, 394]. Deep neural networks could also be used for adaptive control in a variety of industrial networking applications [395], enabling energy savings through self-learning about devices' surroundings.

Predictive maintenance
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 [396, 397]. 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.

Using cleaner electricity ***High Leverage***
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 [362] and powder-coating [398] 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 [84, 399]. 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 [363].

4.4 Discussion

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. As we have seen, there are a number of highly impactful applications where ML can help reduce GHG emissions in industry, with several caveats. First, incentives for cleaner production and distribution are not always aligned with reduced costs, though policies can play a role in aligning these incentives. Second, 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 [359, 400]. 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. Finally, misjudgments can be very costly for manufacturers and retailers, leading most managers to adopt risk-averse strategies towards relatively untested technologies such as ML [391]. For this reason, ML algorithms that determine industrial activities should be robust enough to guarantee both performance and safety, along with providing both interpretable and reproducible results [401].

5 Farms & Forests *by Alexandre Lacoste*

Plants, microbes, and other organisms have been drawing CO_2 from the atmosphere for millions of years. Most of this carbon is continually broken down and recirculated through the carbon cycle, and some is stored deep underground as coal and oil, but a large amount of carbon is sequestered in the biomass of trees, peat bogs, and soil. Our current economy encourages practices that are freeing much of this sequestered carbon through deforestation and unsustainable agriculture. On top of these effects, cattle and rice farming generate methane, a greenhouse gas far more potent than CO_2 itself. Overall, land use by humans is estimated to be responsible for about a quarter of global GHG emissions [26] (and this may be an underestimate [402]). In addition to this direct release of carbon through human actions, the permafrost is now melting, peat bogs are drying, and forest fires are becoming more frequent as a consequence of climate change itself – all of which release yet more carbon [403].

The large scale of this problem allows for a similar scale of positive impact. According to one estimate [404], about a third of GHG emissions reductions could come from better land management and agriculture. ML can play an important role in some of these areas. Precision agriculture could reduce carbon release from the soil and improve crop yield, which in turn could reduce the need for deforestation. Satellite images make it possible to estimate the amount of carbon sequestered in a given area of land, as well as track GHG emissions from it. ML can help monitor the health of forests and peatlands, predict the risk of fire, and contribute to sustainable forestry (Fig. 5). These areas represent highly impactful applications, in particular, of sophisticated computer vision tools, though care must be taken in some cases to avoid negative consequences via the Jevons paradox.

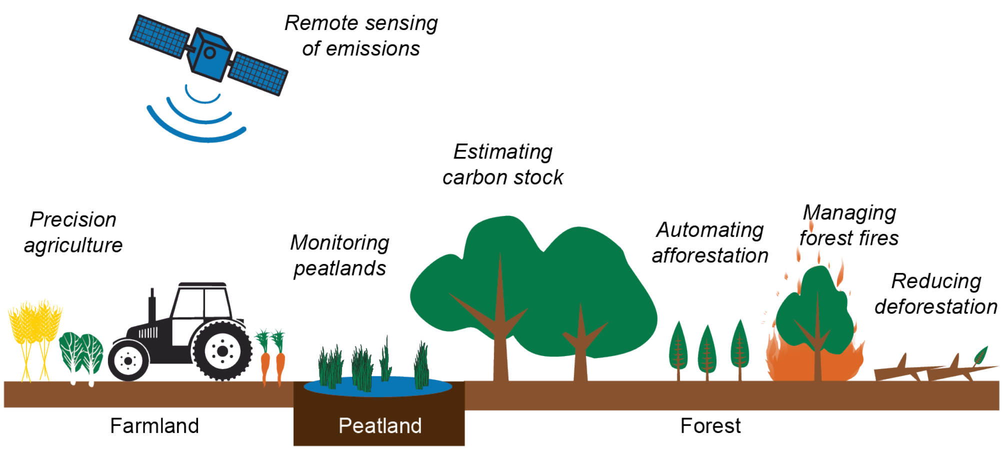

Figure 5: Selected strategies to mitigate GHG emissions from land use using machine learning.

5.1 Remote sensing of emissions *High Leverage*

Having real-time maps of GHGs could help us quantify emissions from agriculture and forestry practices, as well as monitor emissions from other sectors (§1.2).

Such information would be valuable in guiding regulations or incentives that could lead to better land use practices. For example, data on emissions make it possible to set effective targets, and pinpointing the sources of emissions makes it possible to enforce regulations.

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 [405, 406]. These cameras can record up to several hundred wavelengths (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 CO_2, CH_4 (methane), H_2O, and N_2O (nitrous oxide) emissions [407, 408]. 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.[23] Some preliminary work [407] has studied this possibility, but there are no clear results as of yet. This is therefore an open problem with high potential impact.

5.2 Precision agriculture

High Leverage ***Uncertain Impact***

Agriculture is responsible for about 14% of GHG emissions [26]. This might come as a surprise, since plants take up CO_2 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 and disrupting organisms in the soil that contribute to sequestration. Finally, because such farming practices strip soil of nutrients, nitrogen-based fertilizers must be added back to the system. Synthesizing these fertilizers consumes massive amounts of energy, about 2% of global energy consumption [386] (see §4.2). Moreover, while some of this nitrogen is absorbed by plants or retained in the soil, some is converted to nitrous oxide,[24] a greenhouse gas that is about 300 times more potent than CO_2.

Such industrial agriculture approaches are ultimately 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."

Smarter robotic tools can help enable precision agriculture. RIPPA [410], a robot under development at the University of Sydney, is equipped with a hyperspectral camera and has the capacity to perform mechanical weeding, targeted pesticide application, and vacuuming of pests. It can cover 5 acres per day on solar energy and collect large datasets [411] for continual improvement. Many other robotic platforms[25] 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.

There are many additional ways in which ML can contribute to precision agriculture. Intelligent irrigation systems can save large amounts of water while reducing pests that thrive under excessive moisture [404]. ML can also help in disease detection, weed detection, and soil sensing [412–414]. ML can guide crop yield prediction [415] and even macroeconomic models that help farmers predict crop demand and decide what to plant at the beginning of the season [416]. 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 [417], and there is already a significant economic incentive to increase efficiency. However, efficiency gains do not necessarily translate into reduced GHG emissions (e.g. 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

[23]Microsatellites with higher resolution hyperspectral cameras are expected to launch over the coming years, including a proposal by Bluefield Technologies that would provide methane detection at 20-meter spatial resolution with daily refresh. Even once this technology comes online, ML will remain useful to cover the daily gaps and to estimate emissions of other GHGs.

[24]Some fertilizer additionally often ends up in waterways, which can contaminate drinking water and induce blooms of toxic algae [409].

[25]Examples include `sagarobotics.com`, `ecorobotix.com`, and `farm.bot`.

intercropping [404]. ML tools for policy makers and agronomists [418] 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 Monitoring peatlands

High Leverage

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 [419]. When peat dries, however, it releases carbon through decomposition and also becomes susceptible to fire [419, 420]. A single peat fire in Indonesia in 1997 is reported to have released emissions comparable to 20-50% of global fossil fuel emissions during the same year [421].

Monitoring peatlands and protecting them from artificial drainage or droughts is essential to preserve the carbon sequestered in them [422, 423]. In [424], 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 to be made by 2020 using specialized satellites [425]. Advanced ML could potentially help develop precise monitoring tools at low cost and predict the risk of fire.

5.4 Managing forests

Estimating carbon stock

High Leverage

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 [426], tree species and heights are a good indicator of the carbon stock.

The height of trees can be estimated fairly accurately with LiDAR devices mounted on UAVs, but 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 [426, 427]. 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 data is often not equally distributed across regions or seasons. Hence domain adaptation and transfer learning techniques may help algorithms to generalize better.

Automating afforestation

Long-term ***Uncertain Impact***

Planting trees, also called *afforestation*, can be a means of sequestering CO_2 over the long term. According to one estimate, up to 0.9 billion hectares of extra canopy cover could theoretically be added [428] globally. However, care must be taken when planting trees to ensure a positive impact. Afforestation that comes at the expense of farmland (or ecosystems such as peat bogs) could result in a net increase of GHG emissions. Moreover, planting trees without regard for local conditions and native species can reduce the climate impact of afforestation as well as negatively affecting biodiversity.

ML can be helpful in automating large-scale afforestation by locating appropriate planting sites, monitoring plant health, assessing weeds, and analyzing trends. Startups like BioCarbon Engineering[26] and Droneseed[27] are even developing UAVs that are capable of planting seed packets more quickly and cheaply than traditional methods [429].

Managing forest fires

Besides their potential for harming people and property, forest fires release CO_2 into the atmosphere (which

[26] www.biocarbonengineering.com
[27] www.droneseed.co

in turn increases the rate of forest fires [430]). 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 [431]. Drought forecasting [432] is helpful in predicting regions that are more at risk, as is estimating the water content in the tree canopy [433]. In [434, 435], 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 [436]. With good tools to evaluate regions that are more at risk, firefighters can perform controlled burns and cut select areas to prevent the progression of fires.

Reducing deforestation ***High Leverage***

Only 17% of the world's forests are legally protected [437]. 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.

Tools for tracking deforestation can provide valuable data for informing policy makers, as well as law enforcement in cases where deforestation may be conducted illegally. ML can be used to differentiate selective cutting from clearcutting using remote sensing imagery [438–441]. Another approach is to install (old) smartphones powered by solar panels in the forest; ML can then be used to detect and report chainsaw sounds within a one-kilometer radius [442].

Logistics and transport still dominate the cost of wood harvesting, which often motivates clearcutting. Increasingly, ML tools [443] 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 by increasing the amount of wood harvested. 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. The scope of the problem is highly globalized, but the necessary actions are highly localized. Many applications also involve a diversity of stakeholders. 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 [444, 445]. Although there has been significant progress in negative emissions research [446–450], the actual CO_2 removal industry is still in its infancy. As such, many of the ML applications we outline in this section are either speculative or in the early stages of development or commercialization.

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) [446, 451, 452]. 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 [453]. These processes can both be massively accelerated by human activity to achieve necessary scales of CO_2 removal [446]. However, although these biomass, mineral, and 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.

→ plants, biochar, minerals + oceanic

6.1 Direct air capture

Long-term

Another approach is to build facilities to extract CO_2 from power plant exhaust, industrial processes, or even ambient air [454]. While this "direct air capture" (DAC) approach faces technical hurdles, it requires little land and has, according to current understanding, minimal negative environmental impacts [455]. The basic idea behind DAC 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 [446, 447]. Several companies have recently been started to pilot these methods.[28,29,30]

look into pilot DAC methods?

While CO_2 sorbents are improving significantly [456, 457], 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 materials discovery and process engineering workflows [87, 92, 93, 458] to maximize sorbent reusability and CO_2 uptake while minimizing the energy 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 [459]).

ML used to optimize tech

material opt.

6.2 Sequestering CO_2

High Leverage *Long-term* *Uncertain Impact*

Once CO_2 is captured, it must be sequestered or stored, securely and at scale, to prevent re-release back into the atmosphere. 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 [446]. A Norwegian oil company has successfully sequestered CO_2 from an offshore natural gas field in a saline aquifer for more than twenty

[28] https://carbonengineering.com/
[29] https://www.climeworks.com/
[30] https://globalthermostat.com/

ML to find sites

years [460]. Another promising option is to sequester CO_2 in volcanic basalt formations, which is being piloted in Iceland [461].

Machine learning may be able to help with many aspects of CO_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 [462]. These models and the data behind them could likely be repurposed to help trap CO_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_2 flow and remaining injection capacity [463]; recently, [464] found success using convolutional image-to-image regression techniques for uncertainty quantification in a global CO_2 storage simulation study. Additionally, it is important to monitor for CO_2 leaks [465]. ML techniques have recently been applied to monitoring potential CO_2 leaks from wells [466]; computer vision approaches for emissions detection (see [467] and §5.1) may also be applicable.

6.3 Discussion

Given limits on how much more CO_2 humanity can safely emit and the difficulties associated with eliminating emissions entirely, CO_2 removal may have a critical role to play in tackling climate change. Promising applications for ML in CO_2 removal include informing research and development of novel component materials, characterizing geologic resource availability, and monitoring underground CO_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.

Adaptation

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°C. The fundamental physics underlying those 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)*.[31] These models inform local and national government decisions (see IPCC reports [4, 26, 469]), help people calculate their climate risks (see §10 and §8) and allow us to estimate the potential impacts of solar geoengineering (see §9).

Recent trends have created opportunities for ML to advance the state-of-the-art in climate prediction (Fig. 6). First, new and cheaper satellites are creating petabytes of climate observation data.[32] Second, massive climate modeling projects are generating petabytes of simulated climate data.[33] Third, climate forecasts are computationally expensive [473] (the simulations in [472] took three weeks to run on NCAR supercomputers), while ML methods are becoming increasingly fast to train and run, especially on next-generation computing hardware. 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.

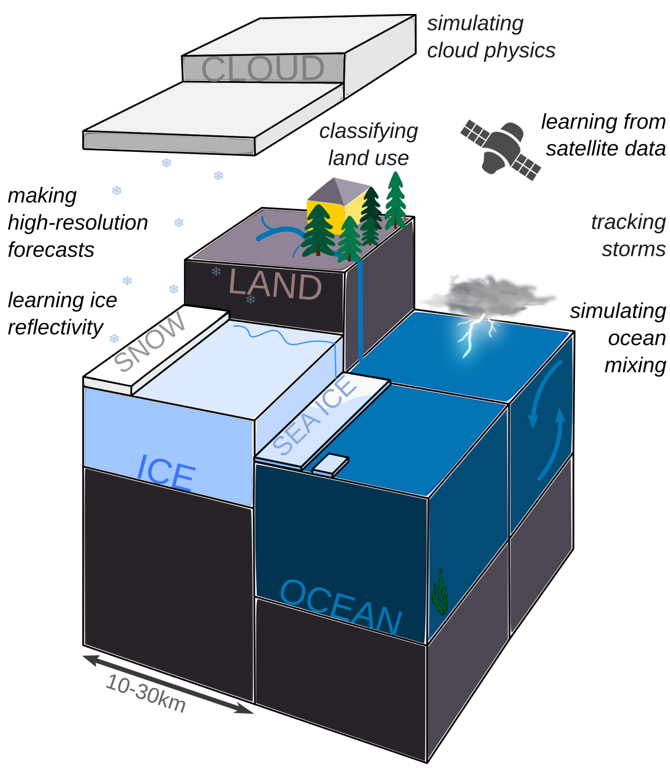

Figure 6: Schematic of a climate model, with selected strategies to improve climate change predictions using machine learning.

[31]Learn about climate modeling from `climate.be/textbook` [468] or Climate Literacy, `youtu.be/XGi2a0tNjOo`

[32]e.g. NASA's Earth Science Data Systems program, `earthdata.nasa.gov`, and ESA's Earth Online, `earth.esa.int`

[33]e.g. the Coupled Model Intercomparison Project, `cmip.llnl.gov` [470, 471] and Community Earth System Model Large Ensemble [472]

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.

7.1.1 Data for climate models

When data are plentiful, climate scientists build data-driven models. In these areas, ML techniques may solve many problems that were previously challenging. These include black box problems, for instance sensor calibration [474], and classification of observational data, for instance classifying crop cover or identifying pollutant sources in satellite imagery [475, 476]. More applications like these are likely to appear as satellite databases grow. The authors of [13] describe many opportunities for data scientists to assimilate data from diverse field and remote sensing sources, many of which have since been explored by climate informatics researchers.

Numerous authors, such as [477], have identified geoscience problems that would be aided by the development of benchmark datasets. Efforts to develop such datasets include EnviroNet [478], the IS-GEO benchmark datasets [479], and ExtremeWeather [480]. We expect the collection of curated geoscience datasets to continue to grow; this process might even be accelerated by ML optimizations in data collection systems [477]. We strongly encourage modellers to dive into the data in collaboration with domain experts. We also recommend that modellers who seek to learn directly from data see [481] for specific advice on fitting and over-fitting climate data.

7.1.2 Accelerating climate models

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. These models are structured in terms of coupled partial differential equations that represent physical processes like cloud formation, ice sheet flow, and permafrost melt. ML models provide new techniques for solving such systems efficiently.

Clouds and aerosols *High Leverage*

Recent work has shown how deep neural networks could be combined with existing thermodynamics knowledge 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 [469, 482]. 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. Gentine 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 [483] and was stable in a simplified global model [484]. Existing scientific model structures do not always offer great trade-offs between cost and accuracy. 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 identify more climate model components that could be replaced by neural networks (we highlight other impactful components below), to optimize those models, and to automate their training workflows (see examples in [485]).

Ice sheets and sea level rise *High Leverage*

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.[34] 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 [486] and models of the extent of Antarctic sea ice do not match reality well [487]. 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 [481, 486, 488]. 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 to predict short-term changes in the sea ice extent.

7.1.3 Working with climate models

ML could also be used to identify and leverage relationships between climate variables. 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. 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 [489].

The best climate predictions are synthesized from ensembles of 20+ climate models [490]. Making good ensemble predictions is an excellent ML problem. Monteleoni et al. proposed that online ML algorithms could create better predictions of one or more target variables in a multi-model ensemble of climate models [491]; this idea has been refined in [492, 493]. More recently, Anderson and Lucas used random forests to make high-resolution predictions from a mix of high- and low-resolution models, which could reduce the costs of building multi-model ensembles [494].

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 [495]. The proposed model would be written in Julia, in contrast to existing models which are mostly written in C++ and 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.

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. 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 [496]).[35] 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

[34] See e.g. `icebridge.gsfc.nasa.gov` and `pgc.umn.edu/data/arcticdem`.

[35] This is one of several reasons why climate models produce accurate long-term predictions in spite of atmospheric chaos.

exceptionally difficult to predict, although Cohen et al. [497] argue that machine learning could bridge that gap by making good predictions on four to six week timescales [498]. 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 [499].

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 [500]. The authors describe ML systems that correct bias, recognize patterns, and predict storms. Moving forward, they envision human experts working alongside automated forecasts.

7.2.1 Storm tracking

Climate models cannot predict the specific dates of future events, but they can predict changes in long-term trends like drought frequency and storm intensity. Information about these trends helps individuals, corporations and towns 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. Researchers have used deep learning to classify [501], detect [480] and segment [502] cyclones and atmospheric rivers, as well as tornadoes [503], in historical climate datasets. Tools for more event types would be useful, as would online tools that work within climate models, labelled datasets for predicting future events, and statistical tools that quantify the uncertainty in new extreme event forecasts.

7.2.2 Local forecasts

High Leverage

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. [504]). 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 [505], which could inform individuals buying insurance or homes. Since ML methods like neural networks are effective at predicting local flooding during extreme weather events [506], these could be used to update local flood risk estimates to benefit individuals. The start-up Jupiter Intelligence is working to make climate predictions more actionable by translating climate forecasts into localised flood and temperature risk scores.

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 [507], including physics-informed neural networks [508] and neural ordinary differential equations [151]. In the further future, researchers are developing ML approaches for a wide range of scientific modeling challenges, including crash prediction [509], adaptive numerical meshing [510], uncertainty quantification [511, 512] and performance optimization [513]. If these strategies 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 [477, 484, 485, 495, 500, 514]. 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 *by Kris Sankaran*

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 strategies have been put forward, from robust power grids to food shortage prediction (Fig. 7), and while this is good news for society, it can be overwhelming for an ML practitioner hoping to contribute. Fortunately, a few critical needs tend to recur across strategies – it is by meeting these needs that ML has the greatest potential to support societal adaptation [8, 16, 515]. 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 (Fig. 7).

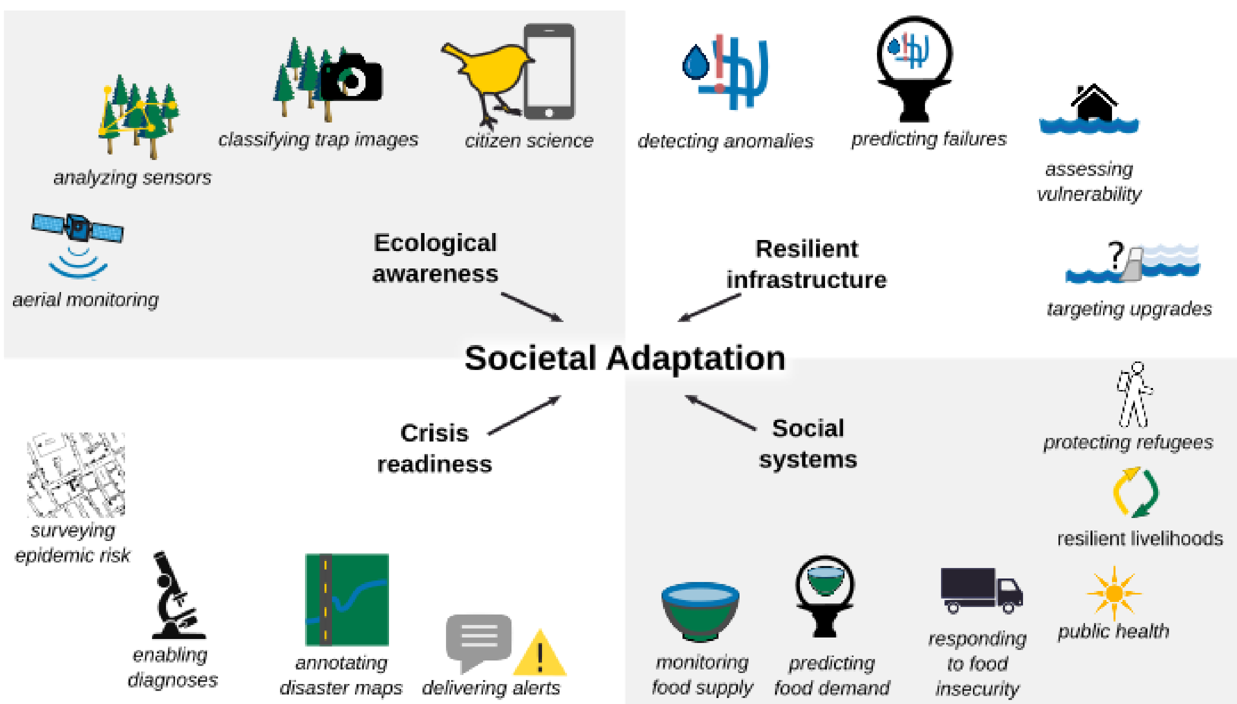

Figure 7: Selected strategies to accelerate societal adaptation to climate change using machine learning.

We note that the projects involved vary in scale from local to global, from infrastructure upgrades and crisis preparedness planning to international ecosystem monitoring and disease surveillance. Hence, we anticipate valuable contributions by researchers who have the flexibility to formulate experimental approaches, by industrial engineers and entrepreneurs who have the expertise to translate prototypes into wide-reaching systems, and by civil servants who lead many existing climate adaptation efforts.

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.

Monitoring ecosystems ***High Leverage***

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 [516–519] (see also §5.1). For example, tree cover can be automatically extracted from aerial imagery to characterize deforestation [520, 521]. At the scale of regions or biomes, analysis of large-scale simulations can illuminate the evolution of ecosystems across potential climate futures [522, 523]. A more direct source of data is offered by environmental sensor networks, made from densely packed but low-cost devices [12, 524, 525]. To monitor ocean ecosystems, marine robots are useful, because they can be used to survey large areas on demand [526, 527].

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 [528, 529]. In marine robotics, improved techniques for sampling regions to explore and automatic summarization of expedition results would both provide value [530, 531]. 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, 14, 532].

Monitoring biodiversity ***High Leverage***

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. ML can help infer biodiversity counts from image-based sensors. For instance, camera traps take photos 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 [533–535]. It is also possible to use aerial imagery to estimate the size of large herds [536] or count birds [537]. In underwater ecosystems, ML has been used to identify plankton automatically from underwater cameras [538] and to infer fish populations from the structure of coral reefs [539].

Citizen science can also enable dataset collection at a scale impossible in individual studies [540–543]. For example, by leveraging public enthusiasm for birdwatching, eBird has logged more than 140 million observations [540], which have been used for population and migration studies [544]. Computer vision algorithms that can classify species from photographs have furthered such citizen science efforts by making identifications easier and more accurate [545, 546], though these face challenges such as class imbalances in training data [547]. Work with citizen science data poses the additional challenge that researchers have no control over where samples come from. To incentivize observations from undersampled regions, mechanisms from game theory can be applied [548], and even when sampling biases persist, estimates of dataset shift can minimize their influence [549].

Monitoring biodiversity may be paired with interventions to protect rare species or control invasive pests. Machine learning is providing new solutions to assess the impact of ecological interventions [550–552] and prevent poaching [548].

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 it is likely that there will be an increased risk of flooding caused by heavy rain or coastal inundations, resulting in property damage and traffic blockages[553].

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 [554, 555]. From this more general perspective of improving social processes, ML can support two types of activities: Design and maintenance.

Designing infrastructure ***Long-term***

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 [556]. 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 [557, 558]. Beyond drawing from flood hazard maps, it is possible to use data from real-world flooding events [559], and to send localized predictions to those at risk [560]. 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 [561–564]. 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 [565].

Infrastructure is long-lived, but the future is uncertain, and planners must weigh immediate resource costs against future societal risks [566]. One area that urgently needs adaptation strategies is the consistent access to drinking water, which can be jeopardized by climate variability [567, 568]. 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 [569–571].

Maintaining infrastructure ***High Leverage***

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 [137, 138, 572, 573]. 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 [574, 575].

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, due to drought and other factors, will pose a threat to food security, as already evidenced by long periods of drought in North America, West Africa and East Asia [576, 577]. 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. See also §11 for perspective on how ML can support the function and analysis of complex social environments.

Food security ***High Leverage***

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 [578–580]. 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 [581, 582]. 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 [583–585], presenting another opportunity for blending large scale simulation with ML [586, 587].

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 [588–591].

Resilient livelihoods

Individuals whose livelihoods depend on one activity, and who have less access to community resources, are those who are most at risk [592, 593]. 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 [594]. Mobile money efforts can increase access to liquid purchasing power; they can also be used to monitor economic health [595, 596]. Skill-matching programs and online training are often driven by data, with some programs specifically aiming to benefit refugees [597–599] (see also §12).

Supporting displaced people *Long-term* *Uncertain Impact*

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 [600–603]. More than quantifying migration patterns, there have been efforts directly aimed at protecting refugees, either through improving rescue operations [604, 605] or monitoring negative public sentiment [606]. 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.

Assessing health risks

Climate change will affect exposure to health hazards, and machine learning can play a role in measuring and mitigating their impacts across subpopulations. Two of the most relevant expected shifts are (1) heat waves will become more frequent and (2) outdoor and indoor air quality will deteriorate [607, 608]. These exposures have either direct or indirect effects on health. For example, prolonged heat episodes both directly cause heat stroke and can trigger acute episodes in chronic conditions, like heart or respiratory disease [609, 610].

Careful data collection and analysis have played a leading role in epidemiology and public health efforts for generations. It should be no surprise that ML has emerged as an important tool in these disciplines, supporting a variety of research efforts, from increasing the efficiency of disease simulators to supporting the fine-grained measurement of exposures and their health impacts [611, 612].

These disciplines are increasingly focused on the risks posed by climate change specifically. For example, new sources of data have enabled detailed sensing of urban heat islands [613–615], water quality [616, 617], and air pollution [618, 619]. Further, data on health indicators, which are already collected, can quantitatively characterize observed impacts across regions as well as illuminate which populations are most at risk to climate-change induced health hazards [620]. For example, it is known that the young, elderly, and socially isolated are especially vulnerable during heat waves, and finer-grained risk estimates could potentially drive outreach [621, 622].

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 [623–625].

Managing epidemics

Climate change will increase the range of vector and water-borne diseases, elevating the likelihood that these new environments experience epidemics [607]. Disease surveillance and outbreak forecasting systems can be built from web data and specially-designed apps, in addition to traditional surveys [626–628]. While non-survey proxies are observational and self-reported, current research attempts to address these issues [629, 630]. Beyond surveillance, point-of-care diagnostics have enjoyed a renaissance, thanks in part to ML [515, 631]. 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 [632]. Ensuring that these systems reliably and transparently augment extension workers, guiding data collection and route planning when appropriate, are active areas of study [633, 634].

Disaster response ***High Leverage***

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 [635, 636]. Further, this imagery can assist damage assessment, by comparing scenes immediately pre- and post-disaster [637, 638]. 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 [624, 639, 640].

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 *by Andrew S. Ross*

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 [641–644]. 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 CO_2 removal, which we discuss in §5 and §6).

Solar geoengineering generally comes with a host of potential side effects and governance challenges. Moreover, unlike CO_2 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) [645]. 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 [646], and it has received increasing attention in the wake of societal inaction on mitigation. Although [644] 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.

Overview

The primary candidate methods for geoengineering are marine cloud brightening [647] (making low-lying clouds more reflective), cirrus thinning [648] (making high-flying clouds trap less heat), and stratospheric aerosol injection [649] (which we discuss below). Other candidates (which are either less effective or harder to implement) include "white-roof" methods [650] and even launching sunshades into space [651].

Injecting sulfate aerosols into the stratosphere is considered a leading candidate for solar geoengineering both because of its economic and technological feasibility [652, 653] 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 [654].) 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 [655], and they make sunlight slightly more diffuse, which can impact agriculture [656].

9.1 Understanding and improving aerosols

Design *Long-term*

The effects and side-effects of aerosols in the stratosphere (or at slightly lower altitudes for cirrus thinning [657]) 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 [658]. 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 [87, 92, 93, 458]. 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.

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, [659] 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).

9.2 Engineering a planetary control system

High Leverage | *Long-term* | *Uncertain Impact*

Efficient emulations and error-bars will be essential for what MacMartin and Kravitz [660] call "The Engineering of Climate Engineering." According to [660], 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 [661–663], which could be useful for fine-tuning geoengineering interventions such as deciding where and when to release aerosols. For an initial attempt at analyzing stratospheric aerosol injection as a reinforcement learning problem (using a neural network climate model emulator), see [664].

9.3 Modeling impacts

Long-term

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. There has been some recent work in applying ML to assess the impacts of geoengineering. For example, [665] use deep neural networks to estimate the effects of aerosols on human health, while [666] use them to estimate the effects of solar geoengineering on agriculture. References [667, 668] use relatively simple local and polynomial regression techniques but applied to extensive empirical data to estimate the past and future effects of temperature change on economic production. More generally, the field of *Integrated Assessment Modeling* [669, 670] aims to map the outputs of a climate model to societal impacts; for a general discussion of potential opportunities for applying ML to IAMs, see §11.2.

9.4 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 [671]. Because of these issues, there has been significant debate about whether it is ethically responsible to research this topic [672, 673]. However, although it creates new risks, solar geoengineering could actually be a moderating force against the terrifying uncertainties climate change already introduces [646, 674], and ultimately many environmental groups and governmental bodies have come down on the side of supporting further research.[36,37,38] In this section, we have attempted to outline some of the technical challenges in implementing and evaluating solar geoengineering. We hope the ML community can help geoengineering researchers tackle these challenges.

[36] https://www.edf.org/climate/our-position-geoengineering

[37] https://www.nrdc.org/media/2015/150210

[38] https://www.ucsusa.org/sites/default/files/attach/2019/gw-position-Solar-Geoengineering-022019.pdf

Tools for Action

10 Individual Action *by Natasha Jaques*

Individuals may worry that they are powerless to affect climate change, or lack clarity on which of their behaviors are most important to change. In fact, there are actions which can meaningfully reduce each person's carbon footprint, and, if widely adopted, could have a significant impact on mitigating global emissions [404, 675]. AI can help to identify those behaviors, inform individuals, and provide constructive opportunities by modeling individual behavior.

10.1 Understanding personal carbon footprint

We as individuals 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.[39] 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 who wish to 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 [676] or interfaces for purchasing flights. Such information can empower people to understand how they can best help mitigate climate change through behavior change.

Residences are responsible for a large share of GHG emissions [4] (see also §3). A large meta-analysis found that significant residential energy savings can be achieved [677], by targeting the right interventions to the right households [678–680]. ML can predict a household's emissions in transportation, energy, water, waste, foods, goods, and services, as a function of its characteristics [681]. These predictions can be used to tailor customized interventions for high-emissions households [682]. Changing behavior both helps mitigate climate change and benefits individuals; studies have shown that many carbon mitigation strategies also provide cost savings to consumers [681].

Household energy disaggregation breaks down overall electricity consumption into energy use by individual appliances (see also §3.1) [683], which can help facilitate behavior change [684]. 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 [685]. A variety of ML techniques have been used to effectively disaggregate household energy, such as spectral clustering, Hidden Markov Models, and neural networks [683].

ML can also be used to predict the marginal emissions of energy consumption in real time, on a scale of hours,[40] potentially allowing consumers to effectively schedule activities such as charging an electric vehicle when the emissions (and prices [686]) will be lowest [687]. 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. [688, 689]). Methods like reinforcement learning can be used to learn how to optimally schedule household appliances to consume energy more efficiently and sustainably [690, 691]. 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 [80, 83].

[39] See e.g. https://www.tmrow.com/
[40] https://www.watttime.org/

10.2 Facilitating behavior change

High Leverage

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. [692–696]), 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 [678]. Other works have used ML to predict how much consumers are willing to pay to avoid potential environmental harms of energy consumption [697], 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 especially worthwhile; all the more so because data show that the size and composition of household carbon footprints varies dramatically across geographic regions and demographics [681].

Citizens who would like to engage with policy decisions, or explore different options to reduce their personal carbon footprint, can have difficulty understanding existing laws and policies due to their complexity. They may benefit from tools that make policy information more manageable and relevant to the individual (e.g. based on where the individual lives). There is the potential for natural language processing to derive understandable insights from policy texts for these applications, similar to automated compliance checking [698, 699].

Understanding individual behavior 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 [700]. 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 [701]. To make the effects of climate change more real for consumers, and thus help motivate those who wish to act, image generation techniques such as CycleGANs have been used to visualize the potential consequences of extreme weather events on houses and cities [702]. Gamification via deep learning has been proposed to further allow individuals to explore their personal energy usage [703]. 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 [704, 705].

10.3 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 Collective Decisions *by Tegan Maharaj and Nikola Milojevic-Dupont*

Addressing climate change requires swift and effective decision-making by groups at multiple levels – communities, unions, NGOs, businesses, governments, intergovernmental organizations, and many more. Such collective decision-making encompasses many kinds of action – for example, negotiating international treaties to reduce GHG emissions, designing carbon markets, building resilient infrastructure, and establishing community-owned solar farms. These decisions often involve multiple stakeholders with different goals and priorities, requiring difficult trade-offs. The economic and societal systems involved are often extremely complex, and the impacts of climate-related decisions can play out globally across long time horizons. To address some of these challenges, researchers are using empirical and mathematical methods from fields such as policy analysis, operations research, economics, game theory, and computational social science; there are many opportunities for ML to support and supplement these methods.

11.1 Modeling social interactions

When designing climate change strategies, it is critical to understand how organizations and individuals act and interact in response to different incentives and constraints. Agent-based models (ABMs) [706, 707] represent one approach used in simulating the actions and interactions of *agents* (people, companies, etc.) in their environment. ABMs have been applied to a multitude of problems relevant to climate change, in particular to study low-carbon technology adoption [708–711]. For example, when modeling solar PV adoption [712], agents may represent individuals who act based on factors such as financial interest and the behavior of their peers [713, 714]; the goal is then to study how these agents interact in response to different conditions, such as electricity rates, subsidy programs, and geographical considerations. Other applications of ABMs include modeling how behavior under social norms changes with external pressures [715], how the economy and climate may evolve given a diversity of political and economic beliefs [716], and how individuals may migrate in response to environmental changes [717]. While agent and environment models in ABMs are often hand-designed by experts, ML can help integrate data-driven insights into these models [718], for example by learning rules or models for agents based on observational data [712, 719], or by using unsupervised methods such as VAEs or GANs to discover salient features useful in modeling a complex environment. While the hope of learning or tuning behavior from data is promising for generalization, many data-driven approaches lose the interpretability for which ABMs are valued; work in interpretable ML methods could potentially help with this.

In addition to ABMs, techniques from game theory can be valuable in modeling behavior, e.g. to explore cooperation in the face of a depleting resource [720]. Multi-agent reinforcement learning can also be applied to understand the behavior of groups of agents who need to cooperate; see [721] for an overview and [722, 723] for recent examples. Combined with mechanism design,[41] such approaches can be used to design methods for cooperation that lead to mutually beneficial outcomes, for example when formalizing procedures around international climate agreements [724, 725].

11.2 Informing policy

The actions required to address climate change, both in mitigation and adaptation, require making policies[42] at the local, national, and international levels [726]. Various institutions act as policy-makers: for instance, governments, international organizations, non-governmental organizations, standards committees, and professional institutions. Tools from *policy analysis* – the process of evaluating the outcomes of past policies

[41]Mechanism design is often called "inverse game theory" – rather than determining optimal strategies for players, mechanism design seeks to design games such that certain strategies are incentivized.

[42]*Policy* can refer, for example, to laws, measures, standards, or best practices.

and assessing future policy alternatives[43] – can help inform the choices these institutions make. Policy analysis uses quantitative tools from statistics, economics, and operations research such as cost-benefit analysis, uncertainty analysis, and multi-criteria decision making to inform the policy-making process; see [727, 728] for an introduction. ML can provide data for policy analysis, help improve existing tools for assessing policy options, and provide new tools for evaluating the effects of policies.

Gathering data *High Leverage*

When creating policies, decision-makers must often negotiate fundamental uncertainties in the underlying data. ML can help alleviate some of this uncertainty by providing data. For instance, as detailed elsewhere in this paper, ML can help pinpoint sources of emissions (§1.2,5.1), approximate traffic patterns (§2.1), identify infrastructure at risk (§8.2), and mine information from companies' financial disclosures (§13). Natural language processing, network analysis, and clustering techniques can also be used to analyze social media data to understand public opinions and discourse around climate change [729–731]. These data can then be used to identify areas of intervention, compute the benefits and costs of a project, or evaluate the effectiveness of a policy after it has been implemented.

Assessing policy options

Decision-makers often construct mathematical models to help them assess or trade off between different policy alternatives. ML is particularly relevant to approaches that model large and complex socio-economic systems to assess outcomes of particular strategies, as well as optimization-based tools that help with navigating the decision.

Policy-makers often wish to analyze how different policy alternatives may contribute to achieving a particular objective. Computational approaches such as simulation and (partial) equilibrium models can be used to compare different policy options, assess the effects of underlying assumptions, or propose strategies that are consistent with the objectives of decision-makers. Of particular relevance to climate change mitigation are *integrated assessment models* (IAMs), which incorporate economic models, climate models, and policy information (see [732] for an overview). IAMs are used to explore future societal pathways that are consistent with climate goals (e.g. 1.5°C mean global temperature increase), and play a prominent role in the IPCC assessments [733]. While these models can simulate interactions between many variables in great detail, this comes at the cost of computational complexity and presents opportunities for machine learning. Much as with Earth system models (§7), ML can be applied within any of the various sub-models that make up an IAM. One set of applications involves deriving results at the appropriate spatial resolution, since different components of an IAM operate at different scales. Outputs with high resolution may be aggregated via clustering methods to provide insights [734], while at coarser resolution, statistical *downscaling* can help to disaggregate data to an appropriate spatial resolution, as seen in applications such as crop yield [735], wind speed [736] or surface temperature [737]. ML also has the potential to help with sensitivity and uncertainty analysis [738], with finding numerical solutions for computational expensive submodels [739, 740], and assessing the validity of the models [741].

In addition to assessing the outcomes of various policies, policy-makers may also employ optimization-based tools to figure out what decisions to make. For example, combinatorial optimization is a powerful tool used widely for decision-making in operations research. See [194] for a survey of how ML can be employed to help solve combinatorial optimization problems.

Tools from the field of *multi-criteria decision-making* can also help policy-makers manage trade-offs between different policies by reconciling competing objectives and minimizing negative side-effects; in particular, in cases where policy objectives and constraints can be mathematically formalized, *multi-objective*

[43] The former is often referred to as *ex-post policy analysis* and the latter as *ex-ante policy analysis*.

optimization can provide a pragmatic approach to making decisions. Here, a decision-maker would formulate their decision-making process as an optimization problem by combining multiple optimization objectives subject to physical or other types of constraints; the goal is to then find a solution (or set of solutions) that is *Pareto-optimal* with respect to all of the objective functions. However, finding these solutions is often computationally expensive. Practitioners have applied bio-inspired algorithms such as particle swarm, genetic, or evolutionary algorithms to search for or compute Pareto-optimal solutions that satisfy the constraints. This approach has been applied in a number of climate change-related fields, including energy and infrastructure planning [111, 742–746], industry [747, 748], land use [749, 750], and more [751–754]. Previous work has also employed parallel surrogate search, assisted by ML, to efficiently solve multi-objective optimization problems [755]. Optimization algorithms which have been successful in the context of hyperparameter tuning (e.g. Bayesian optimization [756, 757]) or guided search algorithms (e.g. tree search algorithms [758]) could also potentially be applied to this problem.

Evaluating policy effects ***High Leverage***

When creating new policies, decision-makers may wish to understand previous policies (e.g. from other jurisdictions) and how these policies performed. ML can help analyze previous policy actions automatically and at scale by improving computational text analysis. In particular, natural language processing methods are already used in the field of political science to analyze political texts and legislation [759]; these approaches could be promising for systematically studying climate change policies. Causal inference techniques can also help assess the effect of a particular policy or climate-related event from observed outcomes. ML can play a role in causal inference [760–762], including in the context of policy problems [763, 764] and in climate-relevant scenarios such as estimating the effects of temperature on human mortality [765] and the effects of World Bank projects on vegetative cover [766].

11.3 Designing markets

In economics, GHG emissions can be seen as a *negative externality*: while a changing climate results in a cost for society, this cost is often not reflected in the market price of goods or services that cause GHG emissions. This is problematic, since organizations and individuals making decisions solely on the basis of market prices will tend to favor cheaper goods, even if those goods emit a large amount of GHGs. Market-based tools[44] such as carbon taxes aim to enforce prices reflecting the societal cost of GHGs and thus encourage socially beneficial behavior through market forces. ML can help in understanding the impacts of market instruments; assessing their effectiveness at reducing emissions; and supporting a swift, effective and fair implementation.[45]

Predicting carbon prices

There are several approaches to pricing GHG emissions. Carbon taxes and quotas aim to influence the behavior of organizations by shaping supply and demand within an existing market. By contrast, cap-and-trade approaches such as those within the European Union involve a completely new market, an *Emissions Trading Scheme*, within which companies can buy and sell a limited number of GHG emissions permits. Prices within such cap-and-trade markets are highly sensitive to control elements such as the number of permits released at a given time. ML can be used to analyze prices within these markets, for example by predicting prices via supervised learning [771–774] or analyzing the main drivers of prices via hierarchical clustering [775].

[44] For background on market-based strategies, see [767–769].

[45] For a review on ML for energy economics and finance, see [770].

Non-carbon markets
Market design can influence GHG emissions even in settings where such emissions are not directly penalized. For instance, dynamic pricing in electricity markets – varying the price of electricity to consumers based on, e.g., how much wind power is available – can shape demand for low-carbon energy sources (see §1.1.1). Following seminal research on modeling pricing in markets as a bandit problem [776], many works have applied bandit and other reinforcement learning (RL) algorithms to determine prices or other market values. For example, RL has been applied to predict bids [777] and market power [778] in electricity markets, and to set dynamic prices in more general settings [79]. ML can also help solve auctions in supply chains [196].

Assessing market effects
When designing market-based strategies, it is necessary to understand how effectively each strategy will reduce emissions, as well as how the underlying socio-technical system may be affected. Studies have considered effects of carbon pricing on economic growth and energy intensity [779, 780], or on electricity prices [781]. Effects of pricing mechanisms can also be indirect, as companies' strategic decisions can have longer-term effects. ML can be useful in analyzing these effects. For example, self-organizing maps have been used to analyze how R&D investment in green technologies changes in response to fuel prices [782], while a game theoretical framework using neural networks has been used to study the optimal production strategies for companies under carbon quotas [783].

To ensure that market-based strategies are effective and equitable, it is important to understand their distributional effects, as certain social groups or classes of stakeholders may be affected more than others. For example, a flat carbon tax on gasoline will have a larger effect on lower-income populations, as fuel expenses are a bigger share of their total budget. Here, clustering can help identify permit allocation schemes that maximize social welfare [784], and supervised learning has been used to predict winners and losers from changing electricity tariff schemes [785]. *Hedonic pricing* can also help identify how much different consumers may be willing to pay for a environmental good or a service, which is a noisy measure for the monetary value of that good or service; these values are typically inferred using regression or ML techniques on historical market data [786–789]. It is also important to analyze which organizations or individuals can actually participate in a given market. For example, carbon markets can be more flexible if viable offsets exist, including those offered by landowners who sequester carbon through forest conservation and management; ML has been used to examine the factors influencing the financial viability of such projects [790].

11.4 Discussion

The complexity, scale, and fundamental uncertainty inherent in the problems of climate change can pose challenges for collective decision-making. ML can help supplement existing mathematical frameworks that are employed to alleviate some of these challenges, including agent-based models, integrated assessment models, multi-objective optimization, and market design. Interpretable and fair ML techniques may be of particular importance in this context, as they may enable decision-makers to more effectively and equitably employ insights from ML models. While these quantitative assessment tools can provide useful input to the decision-making process, it is worth noting that decisions regarding climate change may ultimately depend on qualitative discussions around norms, values, or equity considerations that may not be captured in quantitative models.

12 Education *by Alexandra Luccioni*

Access to quality education is a key part of sustainable development, with significant benefits for climate and society at large. Education contributes to improving quality of life, helps individuals make informed decisions, and trains the next generation of innovators. Education is also paramount in helping people across societies understand and address the causes and consequences of climate change and provides the skills and tools necessary for adapting to its impacts. For instance, education can both improve the resilience of communities, particularly in developing countries that will be disproportionately affected by climate change [791], and empower individuals, especially from developed countries, to adopt more sustainable lifestyles [792]. As climate change itself may diminish educational outcomes for some populations, due to its negative effects on agricultural productivity and household income [793, 794], this makes providing high-quality educational interventions globally all the more important.

AI in Education *Long-term*

There are a number of ways that AI and ML can contribute to education and teaching – for instance by improving access to educational opportunities, helping personalize the teaching process, and stepping in when teachers have limited time. The field of AIED (Artificial Intelligence in EDucation) has existed for over 30 years, and until recently relied on explicitly modeling content, learners, and tutoring strategies based on psychological theories of learning. However, AIED is increasingly incorporating data-driven insights derived from ML techniques.

One important area of AIED research has been Intelligent Tutoring Systems (ITSs), which can adapt their behavior in real time according to the needs of individuals or to support collaborative learning [795]. While ITSs have traditionally been defined and constructed by hand, recent approaches have applied ML techniques such as multi-armed bandit techniques to adaptively personalize sequences of learning activities [796], LSTMs to generate questions to evaluate language comprehension [797], and reinforcement learning to improve the strategies used within the ITS [798, 799]. However, there remains much work to be done to bridge the performance gap between digital and human tutors, and ML-based approaches have an important role to play in this endeavor – for example, via natural language processing techniques for creating conversational agents [800], learner analytics for classifying student profiles, [801], and adaptive learning approaches to propose relevant educational activities and exercises [802]. [46]

While ITSs generally focus on individualized or small-group instruction, AIED can also help provide tools that improve educational outcomes at scale for larger groups of learners. For instance, scalable, adaptive online courses could give hundreds of thousands of learners access to learning resources that they would not usually have in their local educational facilities [806]. Furthermore, giving teachers guidance derived from computational teaching algorithms or heuristics could help them design better educational curricula and improve student learning outcomes [807]. In this context, AIED applications 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. Key considerations for creating AIED tools that can be applied across the globe include adapting to local technological and cultural needs, addressing barriers such as access to electricity and internet [142, 143], and taking into account students' computing skills, language, and culture [808, 809].

Learning about climate

Research has shown that educational activities centered on climate change and carbon footprints can engage learners in understanding the connection between personal and collective actions and their impact on

[46] For further background on this area, see [803–805].

global climate, and can enable individuals to make climate-friendly lifestyle choices such as reducing energy use [810]. There have also been proposals for interactive websites explaining climate science as well as educational interventions focusing on local and actionable aspects of sustainable development [811]. In these contexts, ML can help create personalized educational tools, for instance by generating images of future impacts of extreme weather events based on a learner's address [702] or by anchoring an individual's learning experience in a digital replica of their real-life location and allowing them to explore the way that climate change will impact a specific location [812].

13 Finance *by Alexandra Luccioni*

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. Climate change poses a substantial financial risks to global assets measured in the trillions of dollars [813], and it is hard to forecast where, how, or when climate change will 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 physical impacts on assets or distribution chains, legislative impacts on profit generation, and indirect market consequences such as supply and demand.[47]

Climate investment

Climate investment, the current dominant approach in climate finance, involves investing money in low-carbon assets [817]. 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 [818] or by designing carbon-neutral investment portfolios that remove or under-weight companies with relatively high carbon footprints [819]. 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 [820]. 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.

Climate analytics *High Leverage*

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 [817]. Since this is a predictive approach to addressing climate change from a financial perspective, it is one where ML can potentially have 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 sources polluted by shifting landscapes. Approaches used in this field include: natural language processing techniques for identifying climate risks and investment opportunities in disclosures made by companies [821] as well as for analyzing the evolution of climate coverage in the media to dynamically hedge climate change risk [822]; econometric approaches for developing arbitrage strategies that take advantage of the carbon risk factor in financial markets [823]; and ML approaches for forecasting the price of carbon in emission exchanges[48] [825, 826].

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; (2) in-depth modeling of variables linked to climate risk; (3) designing a statistical climate factor that can be used to project the variation of stock prices given a compound set of events; and (4) identifying direct and indirect climate risk exposure in annual company reports. ML plays a central role in these strategies, and can be a powerful tool in leveraging the financial sector to mitigate climate change and in reducing the financial impacts of climate change on society.

[47] For further reading regarding the impact of climate change on financial markets, see [814–816].

[48] 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 [824] and also §11.3.

Conclusion

Machine learning, like any 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 enable automatic monitoring through remote sensing (e.g. by pinpointing deforestation, gathering data on buildings, and assessing damage after disasters). It can accelerate the process of scientific discovery (e.g. by suggesting new materials for batteries, construction, and carbon capture). ML can optimize systems to improve efficiency (e.g. by consolidating freight, designing carbon markets, and reducing food waste). And it can accelerate computationally expensive physical simulations through hybrid modeling (e.g. climate models and energy scheduling models). These and other cross-cutting themes are shown in Table 2. We emphasize that in each application, ML is only one part of the solution; it is a tool that enables other tools across fields.

Applying machine learning to tackle climate change has the potential both to benefit society and to advance the field of machine learning. Many of the problems we have discussed here highlight cutting-edge areas of ML, such as interpretability, causality, and uncertainty quantification. Moreover, meaningful action on climate problems requires dialogue with fields within and outside computer science and can lead to interdisciplinary methodological innovations, such as improved physics-constrained ML techniques.

The nature of climate-relevant data poses challenges and opportunities. For many of the applications we identify, data can be proprietary or include sensitive personal information. Where datasets exist, they may not be organized with a specific task in mind, unlike typical ML benchmarks that have a clear objective. Datasets may include information from heterogeneous sources, which must be integrated using domain knowledge. Moreover, the available data may not be representative of global use cases. For example, forecasting weather or electricity demand in the US, where data are abundant, is very different from doing so 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; 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 is a starting point.
- **Collaborate.** Find collaborators, who may be researchers, entrepreneurs, established companies, or policy makers. Every domain discussed here has experts who understand its opportunities and pitfalls, even if they do not necessarily understand ML.
- **Listen.** Listen to what your collaborators and other stakeholders say is needed. 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.

	Accelerated experimentation	Control systems	Forecasting	Human interaction	Hybrid physical models	Predictive maintenance	Remote sensing	System optimization
1 Electricity systems								
Enabling low-carbon electricity	•	•	•		•	•	•	•
Reducing current-system impacts			•			•	•	
Ensuring global impact			•		•		•	
2 Transportation								
Reducing transport activity	•	•	•				•	•
Improving vehicle efficiency	•	•						•
Alternative fuels & electrification	•	•	•					•
Modal shift		•	•	•		•	•	•
3 Buildings and cities								
Optimizing buildings		•	•		•	•		•
Urban planning							•	
The future of cities								•
4 Industry								
Optimizing supply chains		•	•					•
Improving materials	•							
Production & energy		•				•		•
5 Farms & forests								
Remote sensing of emissions							•	
Precision agriculture		•	•				•	
Monitoring peatlands							•	
Managing forests		•	•				•	
6 Carbon dioxide removal								
Direct air capture	•				•			
Sequestering CO_2					•			
7 Climate prediction								
Uniting data, ML & climate science			•		•		•	
Forecasting extreme events			•		•		•	
8 Societal impacts								
Ecology							•	
Infrastructure						•		•
Social systems			•	•			•	•
Crisis			•				•	
9 Solar geoengineering								
Understanding & improving aerosols		•			•			
Engineering a planetary control system		•			•			
Modeling impacts					•			
10 Individual action								
Understanding personal footprint			•	•				
Facilitating behavior change				•				
11 Collective decisions								
Modeling social interactions				•				
Informing policy				•				
Designing markets			•					•
12 Education				•				
13 Finance			•	•				

Table 2: Cross-cutting objectives that are relevant to many climate change domains.

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), Carl Elkin (DeepMind), Jim Gao (DeepMind), Muhammad Hasan (DeepMind), Guannan He (CMU), Jeremy Keen (CMU), Zico Kolter (CMU), Luke Lavin (CMU), Sanam Mirzazad (EPRI), David Pfau (DeepMind), Crystal Qian (DeepMind), Juliet Rothenberg (DeepMind), Sims Witherspoon (DeepMind) 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 Érika Mata (IVL - Swedish Environmental Research Institute, IPCC Lead Author Buildings section), Duccio Piovani (nam.R) and Jack Kelly (Open Climate Fix) for feedback and ideas.

Industry. We appreciate all the constructive feedback from Angela Acocella (MIT), Kevin McCloskey (Google), and Bill Tubbs (University of British Columbia), and we are grateful to Kipp Bradford (Yale) for his recommendations around 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.

Farms & Forests. 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 Ghaleb Abdulla (LLNL), Ben Kravitz (PNNL) and David John Gagne II (UCAR) for enlightening conversations; Goodwin Gibbins (Imperial College London) and Ben Kravitz (PNNL) for detailed editing and feedback; and Claire Monteleoni (CU Boulder) and Prabhat (LBL) for feedback which improved the quality of this manuscript.

Societal adaptation. We thank Loubna Benabbou (UQAR), Mike Schäfer (University of Zurich), Andrea Garcia Tapia (Stevens Tech), Slava Jankin Mikhaylov (Hertie School Berlin), 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.

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

Collective Decisions. We thank Sebastian Sewerin (ETH Zürich), D. Cale Reeves (UT Austin), and Rahul Ladhania (UPenn).

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 constructive discussions and the valuable feedback.

The authors gratefully acknowledge support from National Science Foundation grant 1803547, the Center for Climate and Energy Decision Making through a cooperative agreement between the National Science Foundation and Carnegie Mellon University (SES-00949710), US Department of Energy contract DE-FG02-97ER25308, the Natural Sciences and Engineering Research Council of Canada, and the MIT Media Lab Consortium.

References

[1] Joseph Romm. *Climate Change: What Everyone Needs to Know*. Oxford University Press, 2018.

[2] David Archer and Stefan Rahmstorf. *The climate crisis: An introductory guide to climate change*. Cambridge University Press, 2010.

[3] Christopher B Field, Vicente Barros, Thomas F Stocker, and Qin Dahe. *Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change*. Cambridge University Press, 2012.

[4] IPCC. *Global warming of 1.5°C. An IPCC special report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty* [V. Masson-Delmotte, P. Zhai, H. O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, Y. Chen, S. Connors, M. Gomis, E. Lonnoy, J. B. R. Matthews, W. Moufouma-Okia, C. Péan, R. Pidcock, N. Reay, M. Tignor, T. Waterfield, X. Zhou (eds.)]. 2018.

[5] Gregory D Hager, Ann Drobnis, Fei Fang, Rayid Ghani, Amy Greenwald, Terah Lyons, David C Parkes, Jason Schultz, Suchi Saria, Stephen F Smith, et al. Artificial intelligence for social good. *Preprint arXiv:1901.05406*, 2019.

[6] Bettina Berendt. AI for the common good?! pitfalls, challenges, and ethics pen-testing. *Paladyn, Journal of Behavioral Robotics*, 10(1):44–65, 2019.

[7] Maria De-Arteaga, William Herlands, Daniel B Neill, and Artur Dubrawski. Machine learning for the developing world. *ACM Transactions on Management Information Systems (TMIS)*, 9(2):9, 2018.

[8] Carla Gomes, Thomas Dietterich, Bistra Dilkina, Ermon Stefano, Fei Fang, Alan Farnsworth, Alan Fern, Xioali Fern, Daniel Fink, Douglas Fisher, Alexander Flecker, Daniel Freund, Angela Fuller, John Gregoire, John Hopcroft, Zico Kolter, Warren Powell, Nicole Santov, John Selker, Bart Selman, Daniel Shelcon, David Shmoys, Milind Tambe, Christopher Wood, Weng-Keen Wong, Xiaojian Wu, Steve Kelling, Yexiang Xue, Amulya Yadav, Aziz Yakubu, and Mary Lou Zeeman. Computational sustainability: Computing for a better world and a sustainable future. *Communications of ACM (in the press)*, 2019.

[9] Lucas N Joppa. The case for technology investments in the environment. *Nature*, pages 325 – 328, 2017.

[10] Jörg Lässig, Kristian Kersting, and Katharina Morik. *Computational Sustainability*, volume 645. Springer, 2016.

[11] Carla P Gomes. Computational sustainability: Computational methods for a sustainable environment, economy, and society. *The Bridge*, 39(4):5–13, 2009.

[12] Thomas G Dietterich. Machine learning in ecosystem informatics and sustainability. In *Twenty-First International Joint Conference on Artificial Intelligence*, 2009.

[13] C. Monteleoni, G.A. Schmidt, F. Alexander, A. Niculescu-Mizil, K. Steinhaeuser, M. Tippett, A. Banerjee, M.B. Blumenthal, A.R. Ganguly, J.E. Smerdon, and M. Tedesco. Climate informatic. In T. Yu, N. Chawla, and S. Simoff, editors, *Computational Intelligent Data Analysis for Sustainable Development; Data Mining and Knowledge Discovery Series*, chapter 4, pages 81–126. CRC Press, Taylor & Francis Group, 2013.

[14] James H Faghmous and Vipin Kumar. A big data guide to understanding climate change: The case for theory-guided data science. *Big data*, 2(3):155–163, 2014.

[15] Lynn Helena Kaack. *Challenges and Prospects for Data-Driven Climate Change Mitigation*. PhD thesis, Carnegie Mellon University, Pittsburgh, PA, 2019.

[16] James D Ford, Simon E Tilleard, Lea Berrang-Ford, Malcolm Araos, Robbert Biesbroek, Alexandra C Lesnikowski, Graham K MacDonald, Angel Hsu, Chen Chen, and Livia Bizikova. Opinion: Big data has big potential for applications to climate change adaptation. *Proceedings of the National Academy of Sciences*, 113(39):10729–10732, 2016.

[17] Stanford Graduate School of Business. Andrew Ng: Artificial intelligence is the new electricity. `https://www.youtube.com/watch?v=21EiKfQYZXc`, Feb 2017.

[18] Sarvapali Ramchurn, Perukrishnen Vytelingum, Alex Rogers, and Nicholas R Jennings. Putting the "smarts" into the smart grid: A grand challenge for artificial intelligence. *Communications of the ACM*, 55(4):86–97, 2012.

[19] Kasun S Perera, Zeyar Aung, and Wei Lee Woon. Machine learning techniques for supporting renewable energy generation and integration: a survey. In *International Workshop on Data Analytics for Renewable Energy Integration*, pages 81–96. Springer, 2014.

[20] David G. Victor. How artificial intelligence will affect the future of energy and climate. `https://www.brookings.edu/research/how-artificial-intelligence-will-affect-the-future-of-energy-and-climate/`, 2019.

[21] T. Bruckner, I.A. Bashmakov, Y. Mulugetta, H. Chum, A. de la Vega Navarro, J. Edmonds, A. Faaij, B. Fungtammasan, A. Garg, E. Hertwich, D. Honnery, D. Infield, M. Kainuma, S. Khennas, S. Kim, H.B. Nimir, K. Riahi, N. Strachan, R. Wiser, and X. Zhang. *Energy Systems, in IPCC, Working Group III contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Climate Change 2014: Mitigation of Climate Change, chapter 8. Geneva* [O. Edenhofer, R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel, J.C. Minx, (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2014.

[22] Federal Energy Regulatory Commission et al. Energy primer: A handbook of energy market basics. *Federal Energy Regulatory Commission: Washington, DC, USA*, 2015.

[23] Alexandra Von Meier. *Electric Power Systems: A Conceptual Introduction*. Wiley Online Library, 2006.

[24] Daniel Sadi Kirschen and Goran Strbac. *Fundamentals of Power System Economics*, volume 1. Wiley Online Library, 2004.

[25] Allen J Wood, Bruce F Wollenberg, and Gerald B Sheblé. *Power Generation, Operation, and Control*. John Wiley & Sons, 2013.

[26] IPCC. *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [O. Edenhofer, R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel, J.C. Minx, (eds.)]. 2014.

[27] Felix Creutzig. Economic and ecological views on climate change mitigation with bioenergy and negative emissions. *GCB bioenergy*, 8(1):4–10, 2016.

[28] Alexey Lokhov. Technical and economic aspects of load following with nuclear power plants. *NEA, OECD, Paris, France*, 2011.

[29] Annette Evans, Vladimir Strezov, and Tim J Evans. Assessment of utility energy storage options for increased renewable energy penetration. *Renewable and Sustainable Energy Reviews*, 16(6):4141–4147, 2012.

[30] Eric S Hittinger and Inês ML Azevedo. Bulk energy storage increases United States electricity system emissions. *Environmental science & technology*, 49(5):3203–3210, 2015.

[31] Oytun Babacan, Ahmed Abdulla, Ryan Hanna, Jan Kleissl, and David G Victor. Unintended effects of residential energy storage on emissions from the electric power system. *Environmental science & technology*, 52(22):13600–13608, 2018.

[32] Johan Mathe, Nina Miolane, Nicolas Sebastien, and Jeremie Lequeux. Pvnet: A lrcn architecture for spatio-temporal photovoltaic powerforecasting from numerical weather prediction. *Preprint arXiv:1902.01453*, 2019.

[33] Utpal Kumar Das, Kok Soon Tey, Mehdi Seyedmahmoudian, Saad Mekhilef, Moh Yamani Idna Idris, Willem Van Deventer, Bend Horan, and Alex Stojcevski. Forecasting of photovoltaic power generation and model optimization: A review. *Renewable and Sustainable Energy Reviews*, 81:912–928, 2018.

[34] Cyril Voyant, Gilles Notton, Soteris Kalogirou, Marie-Laure Nivet, Christophe Paoli, Fabrice Motte, and Alexis Fouilloy. Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, 105:569–582, 2017.

[35] Can Wan, Jian Zhao, Yonghua Song, Zhao Xu, Jin Lin, and Zechun Hu. Photovoltaic and solar power forecasting for smart grid energy management. *CSEE Journal of Power and Energy Systems*, 1(4):38–46, 2015.

[36] Yuchi Sun, Gergely Szűcs, and Adam R Brandt. Solar pv output prediction from video streams using convolutional neural networks. *Energy & Environmental Science*, 11(7):1811–1818, 2018.

[37] Alasdair Bruce and Lyndon Ruff. Deep learning solar PV and carbon intensity forecasts. `http://powerswarm.co.uk/wp-content/uploads/2018/10/2018.10.18-Bruce-National-Grid-ESO-Deep-Learning-Solar-PV-and-Carbon-Intensity.pdf`.

[38] Ahmad Alzahrani, Pourya Shamsi, Cihan Dagli, and Mehdi Ferdowsi. Solar irradiance forecasting using deep neural networks. *Procedia Computer Science*, 114:304–313, 2017.

[39] Jiaming Li, John K Ward, Jingnan Tong, Lyle Collins, and Glenn Platt. Machine learning for solar irradiance forecasting of photovoltaic system. *Renewable energy*, 90:542–553, 2016.

[40] Navin Sharma, Pranshu Sharma, David Irwin, and Prashant Shenoy. Predicting solar generation from weather forecasts using machine learning. In *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 528–533. IEEE, 2011.

[41] Aoife M Foley, Paul G Leahy, Antonino Marvuglia, and Eamon J McKeogh. Current methods and advances in forecasting of wind power generation. *Renewable Energy*, 37(1):1–8, 2012.

[42] Machine learning can boost the value of wind energy. `https://deepmind.com/blog/machine-learning-can-boost-value-wind-energy/`.

[43] Can Wan, Zhao Xu, Pierre Pinson, Zhao Yang Dong, and Kit Po Wong. Probabilistic forecasting of wind power generation using extreme learning machine. *IEEE Transactions on Power Systems*, 29(3):1033–1044, 2014.

[44] Da Liu, Dongxiao Niu, Hui Wang, and Leilei Fan. Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm. *Renewable Energy*, 62:592–597, 2014.

[45] Pierre Pinson and GN Kariniotakis. Wind power forecasting using fuzzy neural networks enhanced with on-line prediction risk assessment. In *2003 IEEE Bologna Power Tech Conference Proceedings,*, volume 2, pages 8–pp. IEEE, 2003.

[46] Tao Hong and Shu Fan. Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, 32(3):914–938, 2016.

[47] Soliman Abdel-hady Soliman and Ahmad Mohammad Al-Kandari. *Electrical load forecasting: modeling and model construction*. Elsevier, 2010.

[48] Hesham K Alfares and Mohammad Nazeeruddin. Electric load forecasting: literature survey and classification of methods. *International journal of systems science*, 33(1):23–34, 2002.

[49] Henrique Steinherz Hippert, Carlos Eduardo Pedreira, and Reinaldo Castro Souza. Neural networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on power systems*, 16(1):44–55, 2001.

[50] Romain Juban, Henrik Ohlsson, Mehdi Maasoumy, Louis Poirier, and J Zico Kolter. A multiple quantile regression approach to the wind, solar, and price tracks of gefcom2014. *International Journal of Forecasting*, 32(3):1094–1102, 2016.

[51] Matt Wytock and Zico Kolter. Sparse Gaussian conditional random fields: Algorithms, theory, and application to energy forecasting. In *International conference on machine learning*, pages 1265–1273, 2013.

[52] Alexander Kell, A Stephen McGough, and Matthew Forshaw. Segmenting residential smart meter data for short-term load forecasting. In *Proceedings of the Ninth International Conference on Future Energy Systems*, pages 91–96. ACM, 2018.

[53] Christian Beckel, Leyna Sadamori, and Silvia Santini. Automatic socio-economic classification of households using electricity consumption data. In *Proceedings of the Fourth International Conference on Future Energy Systems*, pages 75–86. ACM, 2013.

[54] James Anderson, Fengyu Zhou, and Steven H Low. Disaggregation for networked power systems. In *2018 Power Systems Computation Conference (PSCC)*, pages 1–7. IEEE, 2018.

[55] Emre C Kara, Ciaran M Roberts, Michaelangelo Tabone, Lilliana Alvarez, Duncan S Callaway, and Emma M Stewart. Disaggregating solar generation from feeder-level measurements. *Sustainable Energy, Grids and Networks*, 13:112–121, 2018.

[56] Gregory S Ledva, Laura Balzano, and Johanna L Mathieu. Real-time energy disaggregation of a distribution feeder's demand using online learning. *IEEE Transactions on Power Systems*, 33(5):4730–4740, 2018.

[57] Lynn H Kaack, Jay Apt, M Granger Morgan, and Patrick McSharry. Empirical prediction intervals improve energy forecasting. *Proceedings of the National Academy of Sciences*, 114(33):8752–8757, 2017.

[58] Priya Donti, Brandon Amos, and J Zico Kolter. Task-based end-to-end model learning in stochastic optimization. In *Advances in Neural Information Processing Systems*, pages 5484–5494, 2017.

[59] Adam N Elmachtoub and Paul Grigas. Smart "predict, then optimize". *Preprint arXiv:1710.08005*, 2017.

[60] Bryan Wilder, Bistra Dilkina, and Milind Tambe. Melding the data-decisions pipeline: Decision-focused learning for combinatorial optimization. *Preprint arXiv:1809.05504*, 2018.

[61] Carlo Brancucci Martinez-Anido, Benjamin Botor, Anthony R Florita, Caroline Draxl, Siyuan Lu, Hendrik F Hamann, and Bri-Mathias Hodge. The value of day-ahead solar power forecasting improvement. *Solar Energy*, 129:192–203, 2016.

[62] KS Pandya and SK Joshi. A survey of optimal power flow methods. *Journal of Theoretical & Applied Information Technology*, 4(5), 2008.

[63] Neel Guha, Zhecheng Wang, Matt Wytock, and Arun Majumdar. Machine learning for AC optimal power flow. `http://www.neelguha.com/opf.pdf`, Jun 2019.

[64] Dimitris Bertsimas and Bartolomeo Stellato. Online mixed-integer optimization in milliseconds. *Preprint arXiv:1907.02206*, 2019.

[65] Ahmed Zamzam and Kyri Baker. Learning optimal solutions for extremely fast ac optimal power flow. *arXiv preprint arXiv:1910.01213*, 2019.

[66] Mahdi Jamei, Letif Mones, Alex Robson, Lyndon White, James Requeima, and Cozmin Ududec. Meta-optimization of optimal power flow. `https://www.climatechange.ai/CameraReady/43/CameraReadySubmission/icml_invenia_cameraready.pdf`, Jun 2019.

[67] Benjamin Donnot, Isabelle Guyon, Marc Schoenauer, Patrick Panciatici, and Antoine Marot. Introducing machine learning for power system operation support. *Preprint arXiv:1709.09527*, 2017.

[68] Andreas Essl, André Ortner, Reinhard Haas, and Peter Hettegger. Machine learning analysis for a flexibility energy approach towards renewable energy integration with dynamic forecasting of electricity balancing power. In *2017 14th International Conference on the European Energy Market (EEM)*, pages 1–6. IEEE, 2017.

[69] Nicholas Moehle, Enzo Busseti, Stephen Boyd, and Matt Wytock. Dynamic energy management. *Preprint arXiv:1903.06230*, 2019.

[70] Roel Dobbe, Oscar Sondermeijer, David Fridovich-Keil, Daniel Arnold, Duncan Callaway, and Claire Tomlin. Towards distributed energy services: Decentralizing optimal power flow with machine learning. *Preprint arXiv:1806.06790*, 2019.

[71] Stavros Karagiannopoulos, Roel Dobbe, Petros Aristidou, Duncan Callaway, and Gabriela Hug. Data-driven control design schemes in active distribution grids: Capabilities and challenges. In *Proceedings of the 2019 IEEE PowerTech conference*. IEEE, 2019.

[72] Stavros Karagiannopoulos, Petros Aristidou, and Gabriela Hug. Data-driven local control design for active distribution grids using off-line optimal power flow and machine learning techniques. *IEEE Transactions on Smart Grid*, 2019.

[73] Roel Dobbe, David Fridovich-Keil, and Claire Tomlin. Fully decentralized policies for multi-agent systems: An information theoretic approach. In *Advances in Neural Information Processing Systems*, pages 2941–2950, 2017.

[74] Jianming Lian, Wei Zhang, Y Sun, Laurentiu D Marinovici, Karanjit Kalsi, and Steven E Widergren. Transactive system: Part i: Theoretical underpinnings of payoff functions, control decisions, information privacy, and solution concepts. Technical report, Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2018.

[75] Jianming Lian, Y Sun, Karanjit Kalsi, Steven E Widergren, Di Wu, and Huiying Ren. Transactive system: Part ii: Analysis of two pilot transactive systems using foundational theory and metrics. Technical report, Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2018.

[76] Lu Zhang, Jianjun Tan, Dan Han, and Hao Zhu. From machine learning to deep learning: progress in machine intelligence for rational drug discovery. *Drug discovery today*, 22(11):1680–1685, 2017.

[77] Camus Energy. `https://camus.energy/`, 2019.

[78] Christian Borgs, Ozan Candogan, Jennifer Chayes, Ilan Lobel, and Hamid Nazerzadeh. Optimal multiperiod pricing with service guarantees. *Management Science*, 60(7):1792–1811, 2014.

[79] Roberto Maestre, Juan Ramón Duque, Alberto Rubio, and Juan Arévalo. Reinforcement learning for fair dynamic pricing. *CoRR*, abs/1803.09967, 2018.

[80] Sarvapali D Ramchurn, Perukrishnen Vytelingum, Alex Rogers, and Nick Jennings. Agent-based control for decentralised demand side management in the smart grid. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, pages 5–12. International Foundation for Autonomous Agents and Multiagent Systems, 2011.

[81] Sarvapali D Ramchurn, Perukrishnen Vytelingum, Alex Rogers, and Nicholas R Jennings. Agent-based homeostatic control for green energy in the smart grid. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(4):35, 2011.

[82] Matthias Deindl, Carsten Block, Rustam Vahidov, and Dirk Neumann. Load shifting agents for automated demand side management in micro energy grids. In *2008 Second IEEE International Conference on Self-Adaptive and Self-Organizing Systems*, pages 487–488. IEEE, 2008.

[83] Fredrik Ygge, JM Akkermans, Arne Andersson, Marko Krejic, and Erik Boertjes. The homebots system and field test: A multi-commodity market for predictive power load management. In *Proceedings Fourth International Conference on the Practical Application of Intelligent Agents and Multi-Agent Technology*, volume 1, pages 363–382, 1999.

[84] Niv Buchbinder, Navendu Jain, and Ishai Menache. Online job-migration for reducing the electricity bill in the cloud. In *International Conference on Research in Networking*, pages 172–185. Springer, 2011.

[85] Daniel F Salas and Warren B Powell. Benchmarking a scalable approximate dynamic programming algorithm for stochastic control of grid-level energy storage. *INFORMS Journal on Computing*, 30(1):106–123, 2018.

[86] Powertac. `https://powertac.org/`, 2019.

[87] Keith T Butler, Daniel W Davies, Hugh Cartwright, Olexandr Isayev, and Aron Walsh. Machine learning for molecular and materials science. *Nature*, 559(7715):547, 2018.

[88] Junwen Bai, Yexiang Xue, Johan Bjorck, Ronan Le Bras, Brendan Rappazzo, Richard Bernstein, Santosh K Suram, Robert Bruce van Dover, John M Gregoire, and Carla P Gomes. Phase mapper: Accelerating materials discovery with ai. *AI Magazine*, 39(1):15–26, 2018.

[89] Carla P Gomes, Junwen Bai, Yexiang Xue, Johan Björck, Brendan Rappazzo, Sebastian Ament, Richard Bernstein, Shufeng Kong, Santosh K Suram, R Bruce van Dover, et al. Crystal: a multi-agent ai system for automated mapping of materials' crystal structures. *MRS Communications*, pages 1–9, 2019.

[90] Santosh K Suram, Yexiang Xue, Junwen Bai, Ronan Le Bras, Brendan Rappazzo, Richard Bernstein, Johan Bjorck, Lan Zhou, R Bruce van Dover, Carla P Gomes, et al. Automated phase mapping with agilefd and its application to light absorber discovery in the v–mn–nb oxide system. *ACS combinatorial science*, 19(1):37–46, 2016.

[91] Koji Fujimura, Atsuto Seko, Yukinori Koyama, Akihide Kuwabara, Ippei Kishida, Kazuki Shitara, Craig AJ Fisher, Hiroki Moriwake, and Isao Tanaka. Accelerated materials design of lithium superionic conductors based on first-principles calculations and machine learning algorithms. *Advanced Energy Materials*, 3(8):980–985, 2013.

[92] Yue Liu, Tianlu Zhao, Wangwei Ju, and Siqi Shi. Materials discovery and design using machine learning. *Journal of Materiomics*, 3(3):159–177, 2017.

[93] Rafael Gómez-Bombarelli, Jennifer N Wei, David Duvenaud, José Miguel Hernández-Lobato, Benjamín Sánchez-Lengeling, Dennis Sheberla, Jorge Aguilera-Iparraguirre, Timothy D Hirzel, Ryan P Adams, and Alán Aspuru-Guzik. Automatic chemical design using a data-driven continuous representation of molecules. *ACS central science*, 4(2):268–276, 2018.

[94] Mitsutaro Umehara, Helge S Stein, Dan Guevarra, Paul F Newhouse, David A Boyd, and John M Gregoire. Analyzing machine learning models to accelerate generation of fundamental materials insights. *npj Computational Materials*, 5(1):34, 2019.

[95] Subhashini Venugopalan and Varun Rai. Topic based classification and pattern identification in patents. *Technological Forecasting and Social Change*, 94:236–250, 2015.

[96] David Abel, Edward C Williams, Stephen Brawner, Emily Reif, and Michael L Littman. Bandit-based solar panel control. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[97] Hany Abdelrahman, Felix Berkenkamp, Jan Poland, and Andreas Krause. Bayesian optimization for maximum power point tracking in photovoltaic power plants. In *2016 European Control Conference (ECC)*, pages 2078–2083. IEEE, 2016.

[98] J Zico Kolter, Zachary Jackowski, and Russ Tedrake. Design, analysis, and learning control of a fully actuated micro wind turbine. In *2012 American Control Conference (ACC)*, pages 2256–2263. IEEE, 2012.

[99] Srinivasan Iyengar, Stephen Lee, Daniel Sheldon, and Prashant Shenoy. Solarclique: Detecting anomalies in residential solar arrays. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, page 38. ACM, 2018.

[100] Bistra Dilkina, Jayant R Kalagnanam, and Elena Novakovskaia. Method for designing the layout of turbines in a windfarm, November 17 2015. US Patent 9,189,570.

[101] Rafał Weron. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International journal of forecasting*, 30(4):1030–1081, 2014.

[102] Jesus Lago, Fjo De Ridder, and Bart De Schutter. Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms. *Applied Energy*, 221:386–405, 2018.

[103] Hao Wang and Baosen Zhang. Energy storage arbitrage in real-time markets via reinforcement learning. In *2018 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2018.

[104] Jordan M Malof, Kyle Bradbury, Leslie M Collins, and Richard G Newell. Automatic detection of solar photovoltaic arrays in high resolution aerial imagery. *Applied energy*, 183:229–240, 2016.

[105] Jiafan Yu, Zhecheng Wang, Arun Majumdar, and Ram Rajagopal. DeepSolar: A machine learning framework to efficiently construct a solar deployment database in the United States. *Joule*, 2(12):2605–2617, 2018.

[106] Priya L Donti, Liu Yajing, Andreas J Schmitt, Andrey Bernstein, Rui Yang, and Yingchen Zhang. Matrix completion for low-observability voltage estimation. *Preprint arXiv:1801.09799*, 2018.

[107] Huaiguang Jiang and Yingchen Zhang. Short-term distribution system state forecast based on optimal synchrophasor sensor placement and extreme learning machine. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2016.

[108] Michael Pertl, Kai Heussen, Oliver Gehrke, and Michel Rezkalla. Voltage estimation in active distribution grids using neural networks. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2016.

[109] William Steinhurst, Patrick Knight, and Melissa Schultz. Hydropower greenhouse gas emissions. *Conservation Law Foundation*, 24:6, 2012.

[110] Energy department awards $5.5 million to apply machine learning to geothermal exploration. `https://www.energy.gov/eere/articles/energy-department-awards-55-million-apply-machine-learning-geothermal-exploration`.

[111] Xiaojian Wu, Jonathan Gomes-Selman, Qinru Shi, Yexiang Xue, Roosevelt Garcia-Villacorta, Elizabeth Anderson, Suresh Sethi, Scott Steinschneider, Alexander Flecker, and Carla Gomes. Efficiently approximating the pareto frontier: hydropower dam placement in the amazon basin. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[112] Fu-Chen Chen and Mohammad R Jahanshahi. NB-CNN: Deep learning-based crack detection using convolutional neural network and naïve Bayes data fusion. *IEEE Transactions on Industrial Electronics*, 65(5):4392–4400, 2018.

[113] Francesco Calivá, Fabio Sousa De Ribeiro, Antonios Mylonakis, Christophe Demazière, Paolo Vinai, Georgios Leontidis, and Stefanos Kollias. A deep learning approach to anomaly detection in nuclear reactors. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2018.

[114] Nicholas W Touran, John Gilleland, Graham T Malmgren, Charles Whitmer, and William H Gates III. Computational tools for the integrated design of advanced nuclear reactors. *Engineering*, 3(4):518–526, 2017.

[115] Nature Physics. Insight: Nuclear fusion. `https://www.nature.com/collections/bccqhmkbyw`.

[116] Steven C Cowley. The quest for fusion power. *Nature Physics*, 12(5):384, 2016.

[117] EA Baltz, E Trask, M Binderbauer, M Dikovsky, H Gota, R Mendoza, JC Platt, and PF Riley. Achievement of sustained net plasma heating in a fusion experiment with the optometrist algorithm. *Scientific reports*, 7(1):6425, 2017.

[118] Barbara Cannas, Alessandra Fanni, E Marongiu, and P Sonato. Disruption forecasting at jet using neural networks. *Nuclear fusion*, 44(1):68, 2003.

[119] A Murari, G Vagliasindi, P Arena, L Fortuna, O Barana, M Johnson, JET-EFDA Contributors, et al. Prototype of an adaptive disruption predictor for jet based on fuzzy logic and regression trees. *Nuclear Fusion*, 48(3):035010, 2008.

[120] Jesús Vega, Sebastián Dormido-Canto, Juan M López, Andrea Murari, Jesús M Ramírez, Raúl Moreno, Mariano Ruiz, Diogo Alves, Robert Felton, JET-EFDA Contributors, et al. Results of the jet real-time disruption predictor in the iter-like wall campaigns. *Fusion Engineering and Design*, 88(6-8):1228–1231, 2013.

[121] CG Windsor, G Pautasso, C Tichmann, RJ Buttery, TC Hender, JET EFDA Contributors, et al. A cross-tokamak neural network disruption predictor for the jet and asdex upgrade tokamaks. *Nuclear fusion*, 45(5):337, 2005.

[122] D Wroblewski, GL Jahns, and JA Leuer. Tokamak disruption alarm based on a neural network model of the high-beta limit. *Nuclear Fusion*, 37(6):725, 1997.

[123] Julian Kates-Harbeck, Alexey Svyatkovskiy, and William Tang. Predicting disruptive instabilities in controlled fusion plasmas through deep learning. *Nature*, 2019.

[124] Justin E Barton, William P Wehner, Eugenio Schuster, Federico Felici, and Olivier Sauter. Simultaneous closed-loop control of the current profile and the electron temperature profile in the tcv tokamak. In *2015 American Control Conference (ACC)*, pages 3316–3321. IEEE, 2015.

[125] F Felici, O Sauter, S Coda, BP Duval, TP Goodman, JM Moret, JI Paley, TCV Team, et al. Real-time physics-model-based simulation of the current density profile in tokamak plasmas. *Nuclear Fusion*, 51(8):083052, 2011.

[126] F Felici and O Sauter. Non-linear model-based optimization of actuator trajectories for tokamak plasma profile control. *Plasma Physics and Controlled Fusion*, 54(2):025002, 2012.

[127] Gregorij V Pereverzev and PN Yushmanov. Astra. automated system for transport analysis in a tokamak. 2002.

[128] JF Artaud, V Basiuk, F Imbeaux, Martin Schneider, J Garcia, G Giruzzi, P Huynh, T Aniel, F Albajar, JM Ané, et al. The cronos suite of codes for integrated tokamak modelling. *Nuclear Fusion*, 50(4):043001, 2010.

[129] RV Budny, R Andre, G Bateman, F Halpern, CE Kessel, A Kritz, and D McCune. Predictions of h-mode performance in iter. Technical report, Princeton Plasma Physics Lab.(PPPL), Princeton, NJ (United States), 2008.

[130] Integrated plasma simulator (ips) v2.1 documentation. `http://ipsframework.sourceforge.net/doc/html/`.

[131] Stiffi Zukhrufany. The utilization of supervised machine learning in predicting corrosion to support preventing pipelines leakage in oil and gas industry. Master's thesis, University of Stavanger, Norway, 2018.

[132] Tim Edward and Rob Salkowitz. How machine learning contributes to smarter pipeline maintenance. `https://www.oilandgaseng.com/articles/how-machine-learning-contributes-to-smarter-pipeline-maintenance/`, Apr 2018.

[133] Jiangwen Wan, Yang Yu, Yinfeng Wu, Renjian Feng, and Ning Yu. Hierarchical leak detection and localization method in natural gas pipeline monitoring sensor networks. *Sensors*, 12(1):189–214, 2012.

[134] SwRI developing methane leak detection system for DOE. `https://www.swri.org/press-release/swri-developing-methane-leak-detection-system-doe`, Oct 2016.

[135] Bluefield Technologies. http://bluefield.co/, 2016.

[136] Biswarup Bhattacharya and Abhishek Sinha. Deep fault analysis and subset selection in solar power grids. *Preprint arXiv:1711.02810*, 2017.

[137] Cynthia Rudin, David Waltz, Roger N Anderson, Albert Boulanger, Ansaf Salleb-Aouissi, Maggie Chow, Haimonti Dutta, Philip N Gross, Bert Huang, Steve Ierome, et al. Machine learning for the New York City power grid. *IEEE transactions on pattern analysis and machine intelligence*, 34(2):328–345, 2012.

[138] Van Nhan Nguyen, Robert Jenssen, and Davide Roverso. Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning. *International Journal of Electrical Power & Energy Systems*, 99:107–120, 2018.

[139] WattTime. https://www.watttime.org/, 2019.

[140] electricitymap. https://www.electricitymap.org, 2019.

[141] Carbon intensity api. https://carbonintensity.org.uk/, 2019.

[142] Shahidur R Khandker, Douglas F Barnes, and Hussain A Samad. *Welfare impacts of rural electrification: a case study from Bangladesh*. The World Bank, 2009.

[143] Shahidur R Khandker, Douglas F Barnes, Hussain Samad, and Nguyen Huu Minh. *Welfare impacts of rural electrification: evidence from Vietnam*. The World Bank, 2009.

[144] Douglas Douglas Austin Ellman. *The reference electrification model: a computer model for planning rural electricity access*. PhD thesis, Massachusetts Institute of Technology, 2015.

[145] Martin Cenek, Rocco Haro, Brandon Sayers, and Jifeng Peng. Climate change and power security: Power load prediction for rural electrical microgrids using long short term memory and artificial neural networks. *Applied Sciences*, 8(5):749, 2018.

[146] Fred Otieno, Nathan Williams, and Patrick McSharry. Forecasting energy demand for microgrids over multiple horizons. In *2018 IEEE PES/IAS PowerAfrica*, pages 457–462. IEEE, 2018.

[147] Hongyu Ren, Russell Stewart, Jiaming Song, Volodymyr Kuleshov, and Stefano Ermon. Learning with weak supervision from physics and data-driven constraints. *AI Magazine*, 39(1), 2018.

[148] Dimitry Gershenson, Brandon Rohrer, and Anna Lerner. A new predictive model for more accurate electrical grid mapping. https://code.fb.com/connectivity/electrical-grid-mapping/.

[149] Carbon tracker to measure world's power plant emissions from space with support from google.org. https://www.carbontracker.org/carbon-tracker-to-measure-worlds-power-plant-emissions-from-space-with-support-from-google-org/, May 2019.

[150] Andrey Bogomolov, Bruno Lepri, Roberto Larcher, Fabrizio Antonelli, Fabio Pianesi, and Alex Pentland. Energy consumption prediction using people dynamics derived from cellular network data. *EPJ Data Science*, 5(1):13, 2016.

[151] Tian Qi Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*, pages 6571–6583, 2018.

[152] Connor Schenck and Dieter Fox. Spnets: Differentiable fluid dynamics for deep neural networks. *Preprint arXiv:1806.06094*, 2018.

[153] Filipe de Avila Belbute-Peres, Kevin Smith, Kelsey Allen, Josh Tenenbaum, and J Zico Kolter. End-to-end differentiable physics for learning and control. In *Advances in Neural Information Processing Systems*, pages 7178–7189, 2018.

[154] Felix Creutzig, Patrick Jochem, Oreane Y. Edelenbosch, Linus Mattauch, Detlef P. van Vuuren, David McCollum, and Jan Minx. Transport: A roadblock to climate change mitigation? *Science*, 350(6263):911–912, 2015.

[155] Steven J. Davis, Nathan S. Lewis, Matthew Shaner, Sonia Aggarwal, Doug Arent, Inês L. Azevedo, Sally M. Benson, Thomas Bradley, Jack Brouwer, Yet-Ming Chiang, Christopher T. M. Clack, Armond Cohen, Stephen Doig, Jae Edmonds, Paul Fennell, Christopher B. Field, Bryan Hannegan, Bri-Mathias Hodge, Martin I. Hoffert, Eric Ingersoll, Paulina Jaramillo, Klaus S. Lackner, Katharine J. Mach, Michael Mastrandrea, Joan Ogden, Per F. Peterson, Daniel L. Sanchez, Daniel Sperling, Joseph Stagner, Jessika E. Trancik, Chi-Jen Yang, and Ken Caldeira. Net-zero emissions energy systems. *Science*, 360(6396), 2018.

[156] R. Schaeffer, R. Sims, J. Corfee-Morlot, F. Creutzig, X. Cruz-Nunez, D. Dimitriu, and M. et al. D'Agosto. *Transport, in IPCC, Working Group III contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Climate Change 2014: Mitigation of Climate Change, chapter 8. Geneva* [O. Edenhofer, R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel, J.C. Minx, (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2014.

[157] Maria Figueroa, Oliver Lah, Lewis M Fulton, Alan McKinnon, and Geetam Tiwari. Energy for transport. *Annual Review of Environment and Resources*, 39:295–325, 2014.

[158] Jacob Teter, Pierpaolo Cazzola, and Timur Gül. *The Future of Trucks*. International Energy Agency, 2017.

[159] Lynn H Kaack, Parth Vaishnav, M Granger Morgan, Inês L Azevedo, and Srijana Rai. Decarbonizing intraregional freight systems with a focus on modal shift. *Environmental Research Letters*, 13(8):083001, 2018.

[160] Zia Wadud, Don MacKenzie, and Paul Leiby. Help or hindrance? the travel, energy and carbon impacts of highly automated vehicles. *Transportation Research Part A: Policy and Practice*, 86:1 – 18, 2016.

[161] Weiliang Zeng, Tomio Miwa, and Takayuki Morikawa. Application of the support vector machine and heuristic k-shortest path algorithm to determine the most eco-friendly path with a travel time constraint. *Transportation Research Part D: Transport and Environment*, 57:458 – 473, 2017.

[162] M.H. Zaki and T. Sayed. Automated cyclist data collection under high density conditions. *IET Intelligent Transport Systems*, 10(5):361–369, 2016.

[163] Robert Krile, Fred Todt, and Jeremy Schroeder. Assessing roadway traffic count duration and frequency impacts on annual average daily traffic estimation. Technical Report FHWA-PL-16-012, Federal Highway Administration, Washington, D.C., United States, 2016.

[164] Ioannis Tsapakis and William H Schneider. Use of support vector machines to assign short-term counts to seasonal adjustment factor groups. *Transportation Research Record: Journal of the Transportation Research Board*, (2527):8–17, 2015.

[165] Massimiliano Gastaldi, Riccardo Rossi, Gregorio Gecchele, and Luca Della Lucia. Annual average daily traffic estimation from seasonal traffic counts. *Procedia-Social and Behavioral Sciences*, 87:279–291, 2013.

[166] Lars Wilko Sommer, Tobias Schuchert, and Jürgen Beyerer. Fast deep vehicle detection in aerial images. In *Applications of Computer Vision (WACV), 2017 IEEE Winter Conference on*, pages 311–319. IEEE, 2017.

[167] Qiling Jiang, Liujuan Cao, Ming Cheng, Cheng Wang, and Jonathan Li. Deep neural networks-based vehicle detection in satellite images. In *Bioelectronics and Bioinformatics (ISBB), 2015 International Symposium on*, pages 184–187. IEEE, 2015.

[168] T Nathan Mundhenk, Goran Konjevod, Wesam A Sakla, and Kofi Boakye. A large contextual dataset for classification, detection and counting of cars with deep learning. In *European Conference on Computer Vision*, pages 785–800. Springer, 2016.

[169] Zhipeng Deng, Hao Sun, Shilin Zhou, Juanping Zhao, and Huanxin Zou. Toward fast and accurate vehicle detection in aerial images using coupled region-based convolutional neural networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2017.

[170] Lynn H. Kaack, George H. Chen, and M. Granger Morgan. Truck traffic monitoring with satellite images. In *Proceedings of the 2Nd ACM SIGCAS Conference on Computing and Sustainable Societies*, COMPASS '19, pages 155–164, New York, NY, USA, 2019. ACM.

[171] Silvio Nocera, Cayetano Ruiz-Alarcón-Quintero, and Federico Cavallaro. Assessing carbon emissions from road transport through traffic flow estimators. *Transportation Research Part C: Emerging Technologies*, 95:125 – 148, 2018.

[172] H. M. Abdul Aziz and Satish V. Ukkusuri. A novel approach to estimate emissions from large transportation networks: Hierarchical clustering-based link-driving-schedules for EPA-MOVES using dynamic time warping measures. *International Journal of Sustainable Transportation*, 12(3):192–204, 2018.

[173] M. Yin, M. Sheehan, S. Feygin, J. Paiement, and A. Pozdnoukhov. A generative model of urban activities from cellular data. *IEEE Transactions on Intelligent Transportation Systems*, 19(6):1682–1696, June 2018.

[174] Wei Ma and Zhen (Sean) Qian. Estimating multi-year 24/7 origin-destination demand using high-granular multi-source traffic data. *Transportation Research Part C: Emerging Technologies*, 96:96 – 121, 2018.

[175] Alireza Ermagun and David Levinson. Spatiotemporal traffic forecasting: review and proposed directions. *Transport Reviews*, 38(6):786–814, 2018.

[176] Liang Tang, Chenfeng Xiong, and Lei Zhang. Spatial transferability of neural network models in travel demand modeling. *Journal of Computing in Civil Engineering*, 32(3):04018010, 2018.

[177] Xiaoqing Dai, Lijun Sun, and Yanyan Xu. Short-term origin-destination based metro flow prediction with probabilistic model selection approach. *Journal of Advanced Transportation*, 2018:15, 2018.

[178] Peyman Noursalehi, Haris N. Koutsopoulos, and Jinhua Zhao. Real time transit demand prediction capturing station interactions and impact of special events. *Transportation Research Part C: Emerging Technologies*, 97:277 – 300, 2018.

[179] Ed Manley, Chen Zhong, and Michael Batty. Spatiotemporal variation in travel regularity through transit user profiling. *Transportation*, 45(3):703–732, May 2018.

[180] Mohammad Sajjad Ghaemi, Bruno Agard, Martin Trépanier, and Vahid Partovi Nia. A visual segmentation method for temporal smart card data. *Transportmetrica A: Transport Science*, 13(5):381–404, Jan 2017.

[181] Calvin P Tribby, Harvey J Miller, Barbara B Brown, Carol M Werner, and Ken R Smith. Analyzing walking route choice through built environments using random forests and discrete choice techniques. *Environment and Planning B: Urban Analytics and City Science*, 44(6):1145–1167, 2017.

[182] Alexandre Jacquillat and Amedeo R. Odoni. A roadmap toward airport demand and capacity management. *Transportation Research Part A: Policy and Practice*, 114:168 – 185, 2018.

[183] Hanbong Lee, Waqar Malik, Bo Zhang, Balaji Nagarajan, and Yoon C Jung. Taxi time prediction at Charlotte airport using fast-time simulation and machine learning techniques. In *15th AIAA Aviation Technology, Integration, and Operations Conference*, page 2272, 2015.

[184] J. Wen, J. Zhao, and P. Jaillet. Rebalancing shared mobility-on-demand systems: A reinforcement learning approach. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pages 220–225, Oct 2017.

[185] Anissa Yuniashaesa Suatmadi, Felix Creutzig, and Ilona Otto. On-demand motorcycle taxis improve mobility, not sustainability. *Case Studies on Transport Policy*, 2019.

[186] Edgar G Hertwich, Saleem Ali, Luca Ciacci, Tomer Fishman, Niko Heeren, Eric Masanet, Farnaz Nojavan Asghari, Elsa Olivetti, Stefan Pauliuk, Qingshi Tu, and Paul Wolfram. Material efficiency strategies to reducing greenhouse gas emissions associated with buildings, vehicles, and electronics—a review. *Environmental Research Letters*, 14(4):043004, apr 2019.

[187] Xiqun (Michael) Chen, Majid Zahiri, and Shuaichao Zhang. Understanding ridesplitting behavior of on-demand ride services: An ensemble learning approach. *Transportation Research Part C: Emerging Technologies*, 76:51 – 70, 2017.

[188] Aditi Moorthy, Robert De Kleine, Gregory Keoleian, Jeremy Good, and Geoff Lewis. Shared autonomous vehicles as a sustainable solution to the last mile problem: A case study of ann arbor-detroit area, mar 2017.

[189] Namwoo Kang, Fred M. Feinberg, and Panos Y. Papalambros. Autonomous electric vehicle sharing system design. *Journal of Mechanical Design*, 139(1):011402–011402–10, 10 2016.

[190] Miaojia Lu, Morteza Taiebat, Ming Xu, and Shu-Chien Hsu. Multiagent spatial simulation of autonomous taxis for urban commute: Travel economics and environmental impacts. *Journal of Urban Planning and Development*, 144(4):04018033, 2018.

[191] T. Donna Chen, Kara M. Kockelman, and Josiah P. Hanna. Operations of a shared, autonomous, electric vehicle fleet: Implications of vehicle & charging infrastructure decisions. *Transportation Research Part A: Policy and Practice*, 94:243 – 254, 2016.

[192] Jonn Axsen and Benjamin K. Sovacool. The roles of users in electric, shared and automated mobility transitions. *Transportation Research Part D: Transport and Environment*, 71:1 – 21, 2019.

[193] W. Brian Arthur. Competing technologies, increasing returns, and lock-in by historical events. *The Economic Journal*, 99(394):116–131, 1989.

[194] Yoshua Bengio, Andrea Lodi, and Antoine Prouvost. Machine Learning for Combinatorial Optimization: a Methodological Tour d'Horizon. *arXiv e-prints*, page arXiv:1811.06128, Nov 2018.

[195] Ben Gesing and D. Peterson, S. and Michelsen. Artificial intelligence in logistics: a collaborative report by DHL and IBM on implications and use cases for the logistics industry. *DHL Trend Research, Troisdorf*, 2018.

[196] Tuomas Sandholm. Very-large-scale generalized combinatorial multi-attribute auctions: Lessons from conducting $60 billion of sourcing.

[197] Tian Xie and Jeffrey C. Grossman. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys. Rev. Lett.*, 120:145301, Apr 2018.

[198] Kyriacos Shiarlis, Joao Messias, Maarten van Someren, Shimon Whiteson, Jaebok Kim, Jered Vroon, Gwenn Englebienne, Khiet Truong, Vanessa Evers, Noé Pérez-Higueras, et al. TERESA: A socially intelligent semi-autonomous telepresence system. In *Workshop on machine learning for social robotics at ICRA-2015 in Seattle*, 2015.

[199] Peter Arnfalk, Ulf Pilerot, Per Schillander, and Pontus Grönvall. Green IT in practice: virtual meetings in Swedish public agencies. *Journal of Cleaner Production*, 123:101–112, 2016.

[200] Jeffrey Marlow, Chiara Borrelli, Sean P. Jungbluth, Colleen Hoffman, Jennifer Marlow, and Peter R. Girguis. Opinion: Telepresence is a potentially transformative tool for field science. *Proceedings of the National Academy of Sciences*, 114(19):4841–4844, 2017.

[201] Yolande Strengers. Meeting in the global workplace: Air travel, telepresence and the body. *Mobilities*, 10(4):592–608, 2015.

[202] Andreas W. Schäfer, Antony D. Evans, Tom G. Reynolds, and Lynnette Dray. Costs of mitigating CO2 emissions from passenger aircraft. *Nature Climate Change*, 6:412 EP –, 11 2015.

[203] Alex Burnap, Yanxin Pan, Ye Liu, Yi Ren, Honglak Lee, Richard Gonzalez, and Panos Y. Papalambros. Improving design preference prediction accuracy using feature learning. *Journal of Mechanical Design*, 138(7):071404–071404–12, 05 2016.

[204] Vijay Manikandan Janakiraman, XuanLong Nguyen, and Dennis Assanis. Stochastic gradient based extreme learning machines for stable online learning of advanced combustion engines. *Neurocomputing*, 177:304 – 316, 2016.

[205] Ahmed M. Ali and Dirk Söffker. Towards optimal power management of hybrid electric vehicles in real-time: A review on methods, challenges, and state-of-the-art solutions. *Energies*, 11(3), 2018.

[206] Raul Yondo, Esther Andrés, and Eusebio Valero. A review on design of experiments and surrogate models in aircraft real-time and many-query aerodynamic analyses. *Progress in Aerospace Sciences*, 96:23 – 61, 2018.

[207] Y-C Lai, C P L Barkan, J Drapa, N Ahuja, J M Hart, P J Narayanan, C V Jawahar, A Kumar, L R Milhon, and M P Stehly. Machine vision analysis of the energy efficiency of intermodal freight trains. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 221(3):353–364, 2007.

[208] L. Scime and J. Beuth. Anomaly detection and classification in a laser powder bed additive manufacturing process using a trained computer vision algorithm. *Additive Manufacturing*, 19:114–126, 2018.

[209] S.A. Shevchik, C. Kenel, C. Leinenbach, and K. Wasmer. Acoustic emission for in situ quality monitoring in additive manufacturing using spectral convolutional neural networks. *Additive Manufacturing*, 21:598–604, 2018.

[210] G.X. Gu, C.-T. Chen, D.J. Richmond, and M.J. Buehler. Bioinspired hierarchical composite design using machine learning: Simulation, additive manufacturing, and experiment. *Materials Horizons*, 5(5):939–945, 2018.

[211] 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.

[212] Austin Brown, Jeffrey Gonder, and Brittany Repac. *An Analysis of Possible Energy Impacts of Automated Vehicles*, pages 137–153. Springer International Publishing, Cham, 2014.

[213] P. A. Hancock, Illah Nourbakhsh, and Jack Stewart. On the future of transportation in an era of automated and autonomous vehicles. *Proceedings of the National Academy of Sciences*, 116(16):7684–7691, 2019.

[214] Joshuah K. Stolaroff, Constantine Samaras, Emma R. O'Neill, Alia Lubers, Alexandra S. Mitchell, and Daniel Ceperley. Energy use and life cycle greenhouse gas emissions of drones for commercial package delivery. *Nature Communications*, 9(1):409, 2018.

[215] Mason Marks. Robots in space: Sharing our world with autonomous delivery vehicles. *SSRN Electronic Journal*, 01 2019.

[216] Matthew Guttenberg, Shashank Sripad, and Venkatasubramanian Viswanathan. Evaluating the potential of platooning in lowering the required performance metrics of li-ion batteries to enable practical electric semi-trucks. *ACS Energy Letters*, 2(11):2642–2646, Oct 2017.

[217] Cathy Wu, Aboudy Kreidieh, Eugene Vinitsky, and Alexandre M. Bayen. Emergent behaviors in mixed-autonomy traffic. In *CoRL*, 2017.

[218] Cathy Wu, Aboudy Kreidieh, Kanaad Parvate, Eugene Vinitsky, and Alexandre M Bayen. Flow: Architecture and benchmarking for reinforcement learning in traffic control. *Preprint arXiv:1710.05465*, 2017.

[219] David Jiménez, Sara Hernández, Jesús Fraile-Ardanuy, Javier Serrano, Rubén Fernández, and Federico Álvarez. Modelling the effect of driving events on electrical vehicle energy consumption using inertial sensors in smartphones. *Energies*, 11(2), 2018.

[220] E. S. Rigas, S. D. Ramchurn, and N. Bassiliades. Managing electric vehicles in the smart grid using artificial intelligence: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 16(4):1619–1635, Aug 2015.

[221] Terry Hansen and Chia-Jiu Wang. Support vector based battery state of charge estimator. *Journal of Power Sources*, 141(2):351 – 358, 2005.

[222] R. Tavakoli and Z. Pantic. ANN-based algorithm for estimation and compensation of lateral misalignment in dynamic wireless power transfer systems for EV charging. In *2017 IEEE Energy Conversion Congress and Exposition (ECCE)*, pages 2602–2609, Oct 2017.

[223] Shuangyuan Wang, Ran Li, Adrian Evans, and Furong Li. Electric vehicle load disaggregation based on limited activation matching pursuits. *Energy Procedia*, 158:2611 – 2616, 2019. Innovative Solutions for Energy Transitions.

[224] Ye Tao, Miaohua Huang, and Lan Yang. Data-driven optimized layout of battery electric vehicle charging infrastructure. *Energy*, 150:735 – 744, 2018.

[225] Matthias D. Galus, Marina González Vayá, Thilo Krause, and Göran Andersson. The role of electric vehicles in smart grids. *Wiley Interdisciplinary Reviews: Energy and Environment*, 2(4):384–400, 2013.

[226] José Vázquez-Canteli and Zoltán Nagy. Reinforcement learning for demand response: A review of algorithms and modeling techniques. *Applied Energy*, 235:1072 – 1089, 2019.

[227] Michael K. Hidrue and George R. Parsons. Is there a near-term market for vehicle-to-grid electric vehicles? *Applied Energy*, 151:67 – 76, 2015.

[228] Lifeng Wu, Xiaohui Fu, and Yong Guan. Review of the remaining useful life prognostics of vehicle lithium-ion batteries using data-driven methodologies. *Applied Sciences*, 6(6):166, 2016.

[229] Wladislaw Waag, Christian Fleischer, and Dirk Uwe Sauer. Critical review of the methods for monitoring of lithium-ion batteries in electric and hybrid vehicles. *Journal of Power Sources*, 258:321–339, 2014.

[230] Xiao-Sheng Si, Wenbin Wang, Chang-Hua Hu, and Dong-Hua Zhou. Remaining useful life estimation–a review on the statistical data driven approaches. *European journal of operational research*, 213(1):1–14, 2011.

[231] Kristen A Severson, Peter M Attia, Norman Jin, Nicholas Perkins, Benben Jiang, Zi Yang, Michael H Chen, Muratahan Aykol, Patrick K Herring, Dimitrios Fraggedakis, et al. Data-driven prediction of battery cycle life before capacity degradation. *Nature Energy*, page 1, 2019.

[232] LD Ellis, S Buteau, Samuel G Hames, LM Thompson, DS Hall, and JR Dahn. A new method for determining the concentration of electrolyte components in lithium-ion cells, using fourier transform infrared spectroscopy and machine learning. *Journal of The Electrochemical Society*, 165(2):A256–A262, 2018.

[233] Xiaosong Hu, Shengbo Eben Li, and Yalian Yang. Advanced machine learning approach for lithium-ion battery state estimation in electric vehicles. *IEEE Transactions on Transportation electrification*, 2(2):140–149, 2016.

[234] Steven K Kauwe, Trevor David Rhone, and Taylor D Sparks. Data-driven studies of li-ion-battery materials. *Crystals*, 9(1):54, 2019.

[235] Samuel Buteau and J R. Dahn. Analysis of thousands of electrochemical impedance spectra of lithium-ion cells through a machine learning inverse model. *Journal of The Electrochemical Society*, 166:A1611–A1622, 01 2019.

[236] Selma Brynolf, Maria Taljegard, Maria Grahn, and Julia Hansson. Electrofuels for the transport sector: A review of production costs. *Renewable and Sustainable Energy Reviews*, 81:1887 – 1905, 2018.

[237] International Energy Agency. *Biofuels for Transport*. 2011.

[238] U.S. Department of Energy. Fuel cell technologies office multi-year research, development, and demonstration plan. `https://www.energy.gov/eere/fuelcells/downloads/fuel-cell-technologies-office-multi-year-research-development-and-22`.

[239] Zachary P. Cano, Dustin Banham, Siyu Ye, Andreas Hintennach, Jun Lu, Michael Fowler, and Zhongwei Chen. Batteries and fuel cells for emerging electric vehicle markets. *Nature Energy*, 3(4):279–289, 2018.

[240] Fan Tong, Paulina Jaramillo, and Inês M. L. Azevedo. Comparison of life cycle greenhouse gases from natural gas pathways for medium and heavy-duty vehicles. *Environmental Science & Technology*, 49(12):7123–7133, 06 2015.

[241] Hichem Omrani. Predicting travel mode of individuals by machine learning. *Transportation Research Procedia*, 10:840–849, 2015.

[242] Daisik Nam, Hyunmyung Kim, Jaewoo Cho, and R Jayakrishnan. A model based on deep learning for predicting travel mode choice. In *Proceedings of the Transportation Research Board 96th Annual Meeting Transportation Research Board, Washington, DC, USA*, pages 8–12, 2017.

[243] Julian Hagenauer and Marco Helbich. A comparative study of machine learning classifiers for modeling travel mode choice. *Expert Systems with Applications*, 78:273 – 282, 2017.

[244] Toru Seo, Takahiko Kusakabe, Hiroto Gotoh, and Yasuo Asakura. Interactive online machine learning approach for activity-travel survey. *Transportation Research Part B: Methodological*, 2017.

[245] Yanshuo Sun, Zhibin Jiang, Jinjing Gu, Min Zhou, Yeming Li, and Lei Zhang. Analyzing high speed rail passengers' train choices based on new online booking data in china. *Transportation Research Part C: Emerging Technologies*, 97:96 – 113, 2018.

[246] Sina Dabiri and Kevin Heaslip. Inferring transportation modes from gps trajectories using a convolutional neural network. *Transportation Research Part C: Emerging Technologies*, 86:360 – 371, 2018.

[247] Wei Tu, Jinzhou Cao, Yang Yue, Shih-Lung Shaw, Meng Zhou, Zhensheng Wang, Xiaomeng Chang, Yang Xu, and Qingquan Li. Coupling mobile phone and social media data: a new approach to understanding urban functions and diurnal patterns. *International Journal of Geographical Information Science*, 31(12):2331–2358, 2017.

[248] Amir Samimi, Kazuya Kawamura, and Abolfazl Mohammadian. A behavioral analysis of freight mode choice decisions. *Transportation Planning and Technology*, 34(8):857–869, 2011.

[249] Ali Jamshidi, Siamak Hajizadeh, Zhou Su, Meysam Naeimi, Alfredo Núñez, Rolf Dollevoet, Bart De Schutter, and Zili Li. A decision support approach for condition-based maintenance of rails based on big data analysis. *Transportation Research Part C: Emerging Technologies*, 95:185 – 206, 2018.

[250] Iman Soleimanmeigouni, Alireza Ahmadi, and Uday Kumar. Track geometry degradation and maintenance modelling: A review. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 232(1):73–102, 2018.

[251] Michael Hyland, Zihan Hong, Helen Karla Ramalho de Farias Pinto, and Ying Chen. Hybrid cluster-regression approach to model bikeshare station usage. *Transportation Research Part A: Policy and Practice*, 115:71 – 89, 2018. Smart urban mobility.

[252] Robert Regue and Will Recker. Proactive vehicle routing with inferred demand to solve the bikesharing rebalancing problem. *Transportation Research Part E: Logistics and Transportation Review*, 72:192 – 209, 2014.

[253] Adish Singla, Marco Santoni, Gÿbor Bartók, Pratik Mukerji, Moritz Meenen, and Andreas Krause. Incentivizing users for balancing bike sharing systems. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, AAAI'15, pages 723–729. AAAI Press, 2015.

[254] A. Ghanem, M. Elhenawy, M. Almannaa, H. I. Ashqar, and H. A. Rakha. Bike share travel time modeling: San Francisco bay area case study. In *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pages 586–591, June 2017.

[255] Kirstin Anderson-Hall, Brandon Bordenkircher, Riley O'Neil, and Smith C Scott. Governing micro-mobility: A nationwide assessment of electric scooter regulations. Technical report, 2019.

[256] Ali Rahim Taleqani, Jill Hough, and Kendall E. Nygard. Public opinion on dockless bike sharing: A machine learning approach. *Transportation Research Record*, 0(0):0361198119838982, 0.

[257] Andrew Small and Laura Bliss. The race to code the curb. *Citylab*, April 2019. Available at `https://www.citylab.com/transportation/2019/04/smart-cities-maps-curb-data-coord-sidewalk-tech-street-design/586177/`.

[258] Mehmet Altinkaya and Metin Zontul. Urban bus arrival time prediction: A review of computational models. *International Journal of Recent Technology and Engineering (IJRTE)*, 2:164–169, 01 2013.

[259] Ehsan Mazloumi, Geoff Rose, Graham Currie, and Sara Moridpour. Prediction intervals to account for uncertainties in neural network predictions: Methodology and application in bus travel time prediction. *Engineering Applications of Artificial Intelligence*, 24(3):534 – 542, 2011.

[260] William Barbour, Juan Carlos Martinez Mori, Shankara Kuppa, and Daniel B. Work. Prediction of arrival times of freight traffic on US railroads using support vector regression. *Transportation Research Part C: Emerging Technologies*, 93:211 – 227, 2018.

[261] José Antonio Moscoso-López, Ignacio Turias, Maria Jesús Jiménez-Come, Juan Jesús Ruiz-Aguilar, and María del Mar Cerbán. A two-stage forecasting approach for short-term intermodal freight prediction. *International Transactions in Operational Research*, 26(2):642–666, 2019.

[262] L. Zhou and G. Wu. An overload behavior detection system for engineering transport vehicles based on deep learning. In *American Institute of Physics Conference Series*, volume 1955 of *American Institute of Physics Conference Series*, page 040038, April 2018.

[263] O. Lucon, D. Ürge Vorsatz, A. Zain Ahmed, P. Bertoldi, L.F. Cabeza, N. Eyre, A. Gadgil, L. D. D. Harvey, Y. Jiang, S. Liphoto, S. Mirasgedis, S. Murakami, J. Parikh, C. Pyke, and M.V. Vilariño. Buildings. In *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel and J.C. Minx (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.* 2014.

[264] Diana Urge-Vorsatz, Ksenia Petrichenko, Maja Staniec, and Jiyong Eom. Energy use in buildings in a long-term perspective. *Current Opinion in Environmental Sustainability*, 5(2):141–151, 2013.

[265] Mark Olsthoorn, Joachim Schleich, and Corinne Faure. Exploring the diffusion of low-energy houses: An empirical study in the european union. *Energy Policy*, 129:1382 – 1393, 2019.

[266] Janet Stephenson, Barry Barton, Gerry Carrington, Daniel Gnoth, Rob Lawson, and Paul Thorsnes. Energy cultures: A framework for understanding energy behaviours. *Energy Policy*, 38(10):6120 – 6129, 2010.

[267] Camilo Mora, Chelsie WW Counsell, Coral R Bielecki, and Leo V Louis. Twenty-seven ways a heat wave can kill you: deadly heat in the era of climate change. *Circulation: Cardiovascular Quality and Outcomes*, 10(11):e004233, 2017.

[268] Camilo Mora, Bénédicte Dousset, Iain R Caldwell, Farrah E Powell, Rollan C Geronimo, Coral R Bielecki, Chelsie WW Counsell, Bonnie S Dietrich, Emily T Johnston, Leo V Louis, et al. Global risk of deadly heat. *Nature Climate Change*, 7(7):501, 2017.

[269] Felix Creutzig, Peter Agoston, Jan C. Minx, Josep G. Canadell, Robbie M. Andrew, Corinne Le Quéré, Glen P. Peters, Ayyoob Sharifi, Yoshiki Yamagata, and Shobhakar Dhakal. Urban infrastructure choices structure climate solutions. *Nature Climate Change*, 6(12):1054–1056, December 2016.

[270] +Neil Gershenfeld, +Stephen Samouhos, and +Bruce Nordman. Intelligent infrastructure for energy efficiency. *Science*, 327(5969):1086–1088, 2010.

[271] Sense. `https://sense.com`.

[272] Kadir Amasyali and Nora M. El-Gohary. A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*, 81:1192 – 1205, 2018.

[273] JF Kreider, DE Claridge, P Curtiss, R Dodier, JS Haberl, and M Krarti. Building energy use prediction and system identification using recurrent neural networks. *Journal of solar energy engineering*, 117(3):161–166, 1995.

[274] Nikolaos G Paterakis, Elena Mocanu, Madeleine Gibescu, Bart Stappers, and Walter van Alst. Deep learning versus traditional machine learning methods for aggregated energy demand prediction. In *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pages 1–6. IEEE, 2017.

[275] Bing Dong, Zhaoxuan Li, SM Mahbobur Rahman, and Rolando Vega. A hybrid model approach for forecasting future residential electricity consumption. *Energy and Buildings*, 117:341–351, 2016.

[276] Liesje Van Gelder, Payel Das, Hans Janssen, and Staf Roels. Comparative study of metamodelling techniques in building energy simulation: Guidelines for practitioners. *Simulation Modelling Practice and Theory*, 49:245 – 257, 2014.

[277] Elena Mocanu, Phuong H. Nguyen, Wil L. Kling, and Madeleine Gibescu. Unsupervised energy prediction in a Smart Grid context using reinforcement cross-building transfer learning. *Energy and Buildings*, 116:646–655, March 2016.

[278] J Zico Kolter and Tommi Jaakkola. Approximate inference in additive factorial hmms with application to energy disaggregation. In *Artificial Intelligence and Statistics*, pages 1472–1482, 2012.

[279] J Zico Kolter, Siddharth Batra, and Andrew Y Ng. Energy disaggregation via discriminative sparse coding. In *Advances in Neural Information Processing Systems*, pages 1153–1161, 2010.

[280] D. Srinivasan, W. S. Ng, and A. C. Liew. Neural-network-based signature recognition for harmonic source identification. *IEEE Transactions on Power Delivery*, 21(1):398–405, Jan 2006.

[281] Jack Kelly and William Knottenbelt. Neural nilm: Deep neural networks applied to energy disaggregation. In *Proceedings of the 2Nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, BuildSys '15, pages 55–64, New York, NY, USA, 2015. ACM.

[282] Fiona Burlig, Christopher Knittel, David Rapson, Mar Reguant, and Catherine Wolfram. Machine learning from schools about energy efficiency. Technical report, National Bureau of Economic Research, 2017.

[283] Jason Hartford, Greg Lewis, Kevin Leyton-Brown, and Matt Taddy. Deep IV: A flexible approach for counterfactual prediction. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1414–1423. JMLR. org, 2017.

[284] Neil Gershenfeld, Stephen Samouhos, and Bruce Nordman. Intelligent infrastructure for energy efficiency. *Science*, 327(5969):1086–1088, 2010.

[285] Zakia Afroz, GM Shafiullah, Tania Urmee, and Gary Higgins. Modeling techniques used in building HVAC control systems: A review. *Renewable and Sustainable Energy Reviews*, 83:64–84, 2018.

[286] Guoyin Fu. Deep belief network based ensemble approach for cooling load forecasting of air-conditioning system. *Energy*, 148:269–282, 2018.

[287] Hussain Kazmi, Fahad Mehmood, Stefan Lodeweyckx, and Johan Driesen. Gigawatt-hour scale savings on a budget of zero: Deep reinforcement learning based optimal control of hot water systems. *Energy*, 144:159–168, February 2018.

[288] Woohyun Kim and James E. Braun. Evaluation of the impacts of refrigerant charge on air conditioner and heat pump performance. *International Journal of Refrigeration*, 35(7):1805 – 1814, 2012.

[289] Zhanwei Wang, Zhiwei Wang, Suowei He, Xiaowei Gu, and Zeng Feng Yan. Fault detection and diagnosis of chillers using Bayesian network merged distance rejection and multi-source non-sensor information. *Applied Energy*, 188:200–214, February 2017.

[290] Feng Jia, Yaguo Lei, Jing Lin, Xin Zhou, and Na Lu. Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mechanical Systems and Signal Processing*, 72:303–315, 2016.

[291] June Young Park, Thomas Dougherty, Hagen Fritz, and Zoltan Nagy. Lightlearn: An adaptive and occupant centered controller for lighting based on reinforcement learning. *Building and Environment*, 147:397–414, 2019.

[292] Parisa Rashidi and Diane J Cook. Keeping the resident in the loop: Adapting the smart home to the user. *IEEE Trans. Systems, Man, and Cybernetics, Part A*, 39(5):949–959, 2009.

[293] Simona D'Oca and Tianzhen Hong. Occupancy schedules learning process through a data mining framework. *Energy and Buildings*, 88:395–408, 2015.

[294] Jie Zhao, Bertrand Lasternas, Khee Poh Lam, Ray Yun, and Vivian Loftness. Occupant behavior and schedule modeling for building energy simulation through office appliance power consumption data mining. *Energy and Buildings*, 82:341–355, October 2014.

[295] Han Zou, Yuxun Zhou, Jianfei Yang, and Costas J Spanos. Towards occupant activity driven smart buildings via wifi-enabled iot devices and deep learning. *Energy and Buildings*, 177:12–22, 2018.

[296] Han Zou, Yuxun Zhou, Jianfei Yang, and Costas J Spanos. Unsupervised wifi-enabled iot device-user association for personalized location-based service. *IEEE Internet of Things Journal*, 6(1):1238–1245, 2019.

[297] Ana Carolina Riekstin, Antoine Langevin, Thomas Dandres, Ghyslain Gagnon, and Mohamed Cheriet. Time series-based GHG emissions prediction for smart homes. *IEEE Transactions on Sustainable Computing*, 2018.

[298] Qinran Hu and Fangxing Li. Hardware design of smart home energy management system with dynamic price response. *IEEE Transactions on Smart Grid*, 4(4):1878–1887, 2013.

[299] Xin Jin, Kyri Baker, Dane Christensen, and Steven Isley. Foresee: A user-centric home energy management system for energy efficiency and demand response. *Applied Energy*, 205:1583 – 1595, 2017.

[300] Yi Liu, Chao Yang, Li Jiang, Shengli Xie, and Yan Zhang. Intelligent edge computing for IoT-based energy management in smart cities. *IEEE Network*, 33(2):111–117, 2019.

[301] Eric Hittinger and Paulina Jaramillo. Internet of Things: Energy boon or bane? *Science*, 364(6438):326–328, 2019.

[302] Muhammad Ateeq, Farruh Ishmanov, Muhammad Khalil Afzal, and Muhammad Naeem. Multi-parametric analysis of reliability and energy consumption in IoT: A deep learning approach. *Sensors*, 19(2):309, 2019.

[303] Inês ML Azevedo. Consumer end-use energy efficiency and rebound effects. *Annual Review of Environment and Resources*, 39:393–418, 2014.

[304] Pei-Luen Patrick Rau. *Cross-Cultural Design. Applications in Cultural Heritage, Creativity and Social Development: 10th International Conference, CCD 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15-20, 2018, Proceedings*, volume 10912. Springer, 2018.

[305] Glenn Gregory Sias. *Characterization of the Life Cycle Environmental Impacts and Benefits of Smart Electric Meters and Consequences of their Deployment in California*. PhD thesis, UCLA, 2017.

[306] Nick Couldry and Ulises A Mejias. Data colonialism: rethinking big data's relation to the contemporary subject. *Television & New Media*, 20(4):336–349, 2019.

[307] Christoph F. Reinhart and Carlos Cerezo Davila. Urban building energy modeling – A review of a nascent field. *Building and Environment*, 97:196–202, February 2016.

[308] nam.R. Digital twin. `https://namr.com/#`.

[309] Annelies Vandermeulen, Bram van der Heijde, and Lieve Helsen. Controlling district heating and cooling networks to unlock flexibility: A review. *Energy*, 151:103–115, 2018.

[310] J. Zico Kolter and Joseph Ferreira. A Large-Scale Study on Predicting and Contextualizing Building Energy Usage. In *Twenty-Fifth AAAI Conference on Artificial Intelligence*, August 2011.

[311] Constantine E Kontokosta and Christopher Tull. A data-driven predictive model of city-scale energy use in buildings. *Applied energy*, 197:303–317, 2017.

[312] Sokratis Papadopoulos, Bartosz Bonczak, and Constantine E Kontokosta. Pattern recognition in building energy performance over time using energy benchmarking data. *Applied Energy*, 221:576–586, 2018.

[313] Sokratis Papadopoulos and Constantine E. Kontokosta. Grading buildings on energy performance using city benchmarking data. *Applied Energy*, 233-234:244 – 253, 2019.

[314] Wenwen Zhang, Caleb Robinson, Subhrajit Guhathakurta, Venu M. Garikapati, Bistra Dilkina, Marilyn A. Brown, and Ram M. Pendyala. Estimating residential energy consumption in metropolitan areas: A microsimulation approach. *Energy*, 155:162 – 173, 2018.

[315] Caleb Robinson, Bistra Dilkina, Jeffrey Hubbs, Wenwen Zhang, Subhrajit Guhathakurta, Marilyn A Brown, and Ram M Pendyala. Machine learning approaches for estimating commercial building energy consumption. *Applied energy*, 208:889–904, 2017.

[316] Fazel Khayatian, Luca Sarto, et al. Building energy retrofit index for policy making and decision support at regional and national scales. *Applied energy*, 206:1062–1075, 2017.

[317] Erwan Bocher, Gwendall Petit, Jérémy Bernard, and Sylvain Palominos. A geoprocessing framework to compute urban indicators: The MApUCE tools chain. *Urban climate*, 24:153–174, 2018.

[318] Alex Nutkiewicz, Zheng Yang, and Rishee K Jain. Data-driven urban energy simulation (DUE-S): A framework for integrating engineering simulation and machine learning methods in a multi-scale urban energy modeling workflow. *Applied energy*, 225:1176–1189, 2018.

[319] Thomas Esch, Wieke Heldens, Andreas Hirner, Manfred Keil, Mattia Marconcini, Achim Roth, Julian Zeidler, Stefan Dech, and Emanuele Strano. Breaking new ground in mapping human settlements from space–the global urban footprint. *ISPRS Journal of Photogrammetry and Remote Sensing*, 134:30–42, 2017.

[320] Microsoft. Computer generated building footprints for the United States. `https://github.com/Microsoft/USBuildingFootprints`.

[321] Zhenyu Lu, Jungho Im, Jinyoung Rhee, and Michael Hodgson. Building type classification using spatial and landscape attributes derived from LiDAR remote sensing data. *Landscape and Urban Planning*, 130:134 – 148, 2014.

[322] André Henn, Christoph Römer, Gerhard Gröger, and Lutz Plümer. Automatic classification of building types in 3D city models. *GeoInformatica*, 16(2):281–306, Apr 2012.

[323] Maros Blaha, Christoph Vogel, Audrey Richard, Jan D Wegner, Thomas Pock, and Konrad Schindler. Large-scale semantic 3D reconstruction: an adaptive multi-resolution model for multi-class volumetric labeling. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3176–3184, 2016.

[324] Xiao Xiang Zhu, Devis Tuia, Lichao Mou, Gui-Song Xia, Liangpei Zhang, Feng Xu, and Friedrich Fraundorfer. Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geoscience and Remote Sensing Magazine*, 5(4):8–36, 2017.

[325] Christian Geiß, Hannes Taubenböck, Michael Wurm, Thomas Esch, Michael Nast, Christoph Schillings, and Thomas Blaschke. Remote sensing-based characterization of settlement structures for assessing local potential of district heat. *Remote Sensing*, 3(7):1447–1471, Jul 2011.

[326] Filip Biljecki, Hugo Ledoux, and Jantien Stoter. Generating 3D city models without elevation data. *Computers, Environment and Urban Systems*, 64:1 – 18, 2017.

[327] Paolo Neirotti, Alberto De Marco, Anna Corinna Cagliano, Giulio Mangano, and Francesco Scorrano. Current trends in smart city initiatives: Some stylised facts. *Cities*, 38:25–36, 2014.

[328] H. Habibzadeh, A. Boggio-Dandry, Z. Qin, T. Soyata, B. Kantarci, and H. T. Mouftah. Soft sensing in smart cities: Handling 3vs using recommender systems, machine intelligence, and data analytics. *IEEE Communications Magazine*, 56(2):78–86, Feb 2018.

[329] Felix Creutzig, Martina Franzen, Rolf Moeckel, Dirk Heinrichs, Kai Nagel, and Helga Weisz. Leveraging Digitalization for Sustainability in Urban Transport (in print). *Global Sustainability*, 2019.

[330] Yu Zheng, Licia Capra, Ouri Wolfson, and Hai Yang. Urban computing: Concepts, methodologies, and applications. *ACM Transaction on Intelligent Systems and Technology*, October 2014.

[331] Riccardo Di Clemente, Miguel Luengo-Oroz, Matias Travizano, Sharon Xu, Bapu Vaitla, and Marta C González. Sequences of purchases in credit card data reveal lifestyles in urban populations. *Nature communications*, 9, 2018.

[332] Shan Jiang, Gaston A Fiore, Yingxiang Yang, Joseph Ferreira Jr, Emilio Frazzoli, and Marta C González. A review of urban computing for mobile phone traces: current methods, challenges and opportunities. In *Proceedings of the 2nd ACM SIGKDD international workshop on Urban Computing*, page 2. ACM, 2013.

[333] Rositsa T Ilieva and Timon McPhearson. Social-media data for urban sustainability. *Nature Sustainability*, 1(10):553, 2018.

[334] Derek Ruths and Jürgen Pfeffer. Social media for large studies of behavior. *Science*, 346(6213):1063–1064, 2014.

[335] Farnaz Mosannenzadeh, Maria Rosaria Di Nucci, and Daniele Vettorato. Identifying and prioritizing barriers to implementation of smart energy city projects in Europe: An empirical approach. *Energy Policy*, 105:191–201, 2017.

[336] City of Los Angeles. Mobility data specification. `https://github.com/CityOfLosAngeles/mobility-data-specification.git`, 2018.

[337] Songnian Li, Suzana Dragicevic, Francesc Antón Castro, Monika Sester, Stephan Winter, Arzu Coltekin, Christopher Pettit, Bin Jiang, James Haworth, Alfred Stein, et al. Geospatial big data handling theory and methods: A review and research challenges. *ISPRS journal of Photogrammetry and Remote Sensing*, 115:119–133, 2016.

[338] Lorenzo Valerio, Andrea Passarella, and Marco Conti. Hypothesis transfer learning for efficient data computing in smart cities environments. In *2016 IEEE International Conference on Smart Computing (SMARTCOMP)*, pages 1–8. IEEE, 2016.

[339] Daniele Ravi, Charence Wong, Benny Lo, and Guang-Zhong Yang. A deep learning approach to on-node sensor data analytics for mobile or wearable devices. *IEEE journal of biomedical and health informatics*, 21(1):56–64, 2017.

[340] Khan Muhammad, Jaime Lloret, and Sung Wook Baik. Intelligent and energy-efficient data prioritization in green smart cities: Current challenges and future directions. *IEEE Communications Magazine*, 57(2):60–65, 2019.

[341] Sarah Giest. Big data analytics for mitigating carbon emissions in smart cities: opportunities and challenges. *European Planning Studies*, 25(6):941–957, Feb 2017.

[342] Antoine Bordes, Xavier Glorot, Jason Weston, and Yoshua Bengio. A semantic matching energy function for learning with multi-relational data. *Machine Learning*, 94(2):233–259, 2014.

[343] AnHai Doan, Jayant Madhavan, Pedro Domingos, and Alon Halevy. Ontology matching: A machine learning approach. In *Handbook on ontologies*, pages 385–403. Springer, 2004.

[344] Yu Zheng. Methodologies for cross-domain data fusion: An overview. September 2015.

[345] Bartosz Krawczyk, Leandro L Minku, João Gama, Jerzy Stefanowski, and Michał Woźniak. Ensemble learning for data stream analysis: A survey. *Information Fusion*, 37:132–156, 2017.

[346] Jinsong Wu, Song Guo, Jie Li, and Deze Zeng. Big data meet green challenges: Big data toward green applications. *IEEE Systems Journal*, 10(3):888–900, 2016.

[347] Edward O'Dwyer, Indranil Pan, Salvador Acha, and Nilay Shah. Smart energy systems for sustainable smart cities: Current developments, trends and future directions. *Applied Energy*, 237:581–597, 2019.

[348] Reid Ewing and Robert Cervero. "Does Compact Development Make People Drive Less?" The Answer Is Yes. *Journal of the American Planning Association*, 83(1):19–25, January 2017.

[349] Felix Creutzig, Giovanni Baiocchi, Robert Bierkandt, Peter-Paul Pichler, and Karen C. Seto. Global typology of urban energy use and potentials for an urbanization mitigation wedge. *Proceedings of the National Academy of Sciences*, 112(20):6283–6288, May 2015.

[350] Chuan Ding, Xinyu Jason Cao, and Petter Næss. Applying gradient boosting decision trees to examine non-linear effects of the built environment on driving distance in oslo. *Transportation Research Part A: Policy and Practice*, 110:107–117, 2018.

[351] Mafalda Silva, Vítor Leal, Vítor Oliveira, and Isabel M Horta. A scenario-based approach for assessing the energy performance of urban development pathways. *Sustainable cities and society*, 40:372–382, 2018.

[352] Saeed Monajem and Farzan Ekram Nosratian. The evaluation of the spatial integration of station areas via the node place model; an application to subway station areas in Tehran. *Transportation Research Part D: Transport and Environment*, 40:14–27, 2015.

[353] Juan F De Paz, Javier Bajo, Sara Rodríguez, Gabriel Villarrubia, and Juan M Corchado. Intelligent system for lighting control in smart cities. *Information Sciences*, 372:241–255, 2016.

[354] William F Lamb, Felix Creutzig, Max W Callaghan, and Jan C Minx. Learning about urban climate solutions from case studies. *Nature Climate Change*, page 1, 2019.

[355] Harini Nagendra, Xuemei Bai, Eduardo S Brondizio, and Shuaib Lwasa. The urban south and the predicament of global sustainability. *Nature Sustainability*, 1(7):341, 2018.

[356] Yafei Han. Global urban typology discovery with a latent class choice model. *Proceedings of the Transportation Research Board 97th Annual Meeting*, page 5.

[357] R. Louf and M. Barthelemy. A typology of street patterns. *Journal of The Royal Society Interface*, 11(101):20140924–20140924, October 2014.

[358] Xuemei Bai, Richard J Dawson, Diana Ürge-Vorsatz, Gian C Delgado, Aliyu Salisu Barau, Shobhakar Dhakal, David Dodman, Lykke Leonardsen, Valérie Masson-Delmotte, Debra C Roberts, et al. Six research priorities for cities and climate change, 2018.

[359] Mike Gualtieri, Noel Yuhanna, Holger Kisker, Rowan Curran, Brandon Purcell, Sophia Christakis, Shreyas Warrier, and Matthew Izzi. The Forrester Wave: Big data streaming analytics, Q1 2016. *Forrester.com*, January 2016.

[360] Rubaiat Habib Kazi, Tovi Grossman, Hyunmin Cheong, Ali Hashemi, and George W Fitzmaurice. DreamSketch: Early stage 3D design explorations with sketching and generative design. In *UIST*, pages 401–414, 2017.

[361] Richard Evans and Jim Gao. DeepMind AI reduces Google data centre cooling bill by 40%. *DeepMind blog*, 20, 2016.

[362] Xiao Zhang, Gabriela Hug, J Zico Kolter, and Iiro Harjunkoski. Model predictive control of industrial loads and energy storage for demand response. In *2016 IEEE Power and Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2016.

[363] Josep Ll. Berral, Íñigo Goiri, Ramón Nou, Ferran Julià, Jordi Guitart, Ricard Gavaldà, and Jordi Torres. Towards energy-aware scheduling in data centers using machine learning. In *Proceedings of the 1st International Conference on Energy-Efficient Computing and Networking*, e-Energy '10, pages 215–224, New York, NY, USA, 2010. ACM.

[364] Steve Sorrell. Jevons' paradox revisited: The evidence for backfire from improved energy efficiency. *Energy policy*, 37(4):1456–1469, 2009.

[365] Auslan Cramb. 12,000-mile trip to have seafood shelled. *The Telegraph*, November 2006.

[366] Xi Wang, Hua Cai, and H Keith Florig. Energy-saving implications from supply chain improvement: An exploratory study on China's consumer goods retail system. *Energy Policy*, 95:411–420, 2016.

[367] Andrew Winston. Excess inventory wastes carbon and energy, not just money. *Harvard Business Review*, 2011.

[368] A Okay Akyuz, Mitat Uysal, Berna Atak Bulbul, and M Ozan Uysal. Ensemble approach for time series analysis in demand forecasting: Ensemble learning. In *2017 IEEE International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, pages 7–12. IEEE, 2017.

[369] Grigorios Tsoumakas. A survey of machine learning techniques for food sales prediction. *Artificial Intelligence Review*, 52(1):441–447, 2019.

[370] SCM Globe. Zara clothing company supply chain. *SCM Globe*, 2015.

[371] Christophe Rizet, Eric Cornélis, Michael Browne, and Jacques Léonardi. GHG emissions of supply chains from different retail systems in Europe. *Procedia-Social and Behavioral Sciences*, 2(3):6154–6164, 2010.

[372] Gustavo M Ugarte, Jay S Golden, and Kevin J Dooley. Lean versus green: The impact of lean logistics on greenhouse gas emissions in consumer goods supply chains. *Journal of Purchasing and Supply Management*, 22(2):98–109, 2016.

[373] Troy R Hawkins, Bhawna Singh, Guillaume Majeau-Bettez, and Anders Hammer Strømman. Comparative environmental life cycle assessment of conventional and electric vehicles. *Journal of Industrial Ecology*, 17(1):53–64, 2013.

[374] Gerald Rebitzer, Tomas Ekvall, Rolf Frischknecht, Davis Hunkeler, G Norris, Tomas Rydberg, W-P Schmidt, Sangwon Suh, B Pennington Weidema, and David W Pennington. Life cycle assessment: Part 1: Framework, goal and scope definition, inventory analysis, and applications. *Environment international*, 30(5):701–720, 2004.

[375] Jenny Gustavsson, Christel Cederberg, Ulf Sonesson, Robert Van Otterdijk, and Alexandre Meybeck. Global food losses and food waste. 2011.

[376] Antonella Meneghetti and Luca Monti. Greening the food supply chain: an optimisation model for sustainable design of refrigerated automated warehouses. *International Journal of Production Research*, 53(21):6567–6587, 2015.

[377] Guillermo Fuertes, Ismael Soto, Raúl Carrasco, Manuel Vargas, Jorge Sabattin, and Carolina Lagos. Intelligent packaging systems: sensors and nanosensors to monitor food quality and safety. *Journal of Sensors*, 2016, 2016.

[378] Manfred Fischedick, Joyashree Roy, Amr Abdel-Aziz, Adolf Acquaye, Julian Allwood, Jean-Paul Ceron, Yong Geng, Haroon Kheshgi, Alessandro Lanza, Daniel Perczyk, Lynn Price, Estela Santalla, Claudia Sheinbaum, Kanako Tanaka, et al. Industry. In *Climate change 2014: mitigation of climate change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 2014.

[379] Johanna Lehne and Felix Preston. Making concrete change, innovation in low-carbon cement and concrete. *Chatham House Report, Energy Enivronment and Resources Department: London, UK*, pages 1–66, 2018.

[380] Ivanna Baturynska, Oleksandr Semeniuta, and Kristian Martinsen. Optimization of process parameters for powder bed fusion additive manufacturing by combination of machine learning and finite element method: A conceptual framework. *Procedia CIRP*, 67:227–232, 2018.

[381] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, et al. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *Apl Materials*, 1(1):011002, 2013.

[382] Xiou Ge, Richard T Goodwin, Jeremy R Gregory, Randolph E Kirchain, Joana Maria, and Lav R Varshney. Accelerated discovery of sustainable building materials. *Preprint arXiv:1905.08222*, 2019.

[383] Logan Ward, Ankit Agrawal, Alok Choudhary, and Christopher Wolverton. A general-purpose machine learning framework for predicting properties of inorganic materials. *npj Computational Materials*, 2:16028, 2016.

[384] Benjamin A Rizkin, Karina Popovich, and Ryan L Hartman. Artificial neural network control of thermoelectrically-cooled microfluidics using computer vision based on ir thermography. *Computers & Chemical Engineering*, 121:584–593, 2019.

[385] Connor W Coley, Wengong Jin, Luke Rogers, Timothy F Jamison, Tommi S Jaakkola, William H Green, Regina Barzilay, and Klavs F Jensen. A graph-convolutional neural network model for the prediction of chemical reactivity. *Chemical science*, 10(2):370–377, 2019.

[386] Joseph H Montoya, Charlie Tsai, Aleksandra Vojvodic, and Jens K Nørskov. The challenge of electrochemical ammonia synthesis: A new perspective on the role of nitrogen scaling relations. *ChemSusChem*, 8(13):2180–2186, 2015.

[387] SW Wood and Annette Cowie. A review of greenhouse gas emission factors for fertiliser production. 2004.

[388] Kenneth Gillingham and James H Stock. The cost of reducing greenhouse gas emissions. *Journal of Economic Perspectives*, 32(4):53–72, 2018.

[389] Erica L Plambeck. Reducing greenhouse gas emissions through operations and supply chain management. *Energy Economics*, 34:S64–S74, 2012.

[390] Lan Tao, Elizabeth Garnsey, David Probert, and Tom Ridgman. Innovation as response to emissions legislation: revisiting the automotive catalytic converter at Johnson Matthey. *R&d Management*, 40(2):154–168, 2010.

[391] Susan Helper, Raphael Martins, and Robert Seamans. Who profits from industry 4.0? theory and evidence from the automotive industry. *SSRN preprint ssrn.3377771*, 2019.

[392] Muhammad Aftab, Chien Chen, Chi-Kin Chau, and Talal Rahwan. Automatic hvac control with real-time occupancy recognition and simulation-guided model predictive control in low-cost embedded system. *Energy and Buildings*, 154:141–156, 2017.

[393] Ján Drgoňa, Damien Picard, Michal Kvasnica, and Lieve Helsen. Approximate model predictive building control via machine learning. *Applied Energy*, 218:199–216, 2018.

[394] Jim Gao. Machine learning applications for data center optimization. 2014.

[395] Ejaz Ahmed, Ibrar Yaqoob, Arif Ahmed, Abdullah Gani, Muhammad Imran, and Sghaier Guizani. Green industrial networking: recent advances, taxonomy, and open research challenges. *IEEE Communications Magazine*, 54(10):38–45, 2016.

[396] Edward Glaessgen and David Stargel. The digital twin paradigm for future NASA and US Air Force vehicles. In *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*, page 1818, 2012.

[397] Fei Tao, Jiangfeng Cheng, Qinglin Qi, Meng Zhang, He Zhang, and Fangyuan Sui. Digital twin-driven product design, manufacturing and service with big data. *The International Journal of Advanced Manufacturing Technology*, 94(9-12):3563–3576, 2018.

[398] Rockwell Automation. Akzonobel powder coatings saves over €15,000 per month thanks to advanced energy monitoring solution from rockwell automation, 2014.

[399] Nathaniel Horner, Inês Azevedo, Doug Sicker, and Yuvraj Agarwal. Dynamic data center load response to variability in private and public electricity costs. In *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 80–85. IEEE, 2016.

[400] Brendan Coffey. Factory records: GE providing Procter & Gamble greater access to the cloud for analyzing manufacturing data, 2019.

[401] Peter Henderson, Riashat Islam, Philip Bachman, Joelle Pineau, Doina Precup, and David Meger. Deep reinforcement learning that matters. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[402] Natalie M Mahowald, Daniel S Ward, Scott C Doney, Peter G Hess, and James T Randerson. Are the impacts of land use on warming underestimated in climate policy? *Environmental Research Letters*, 12(9):094016, 2017.

[403] Nasa Science. The study of Earth as an integrated system. `https://climate.nasa.gov/nasa_science/science/`, 2019.

[404] Paul Hawken. *Drawdown: The most comprehensive plan ever proposed to reverse global warming*. 2015.

[405] Kevin P Gibbons. Hyperspectral imaging what is it? how does it work? Technical report, 2014.

[406] Rebecca Scafutto and Carlos de Souza Filho. Detection of methane plumes using airborne midwave infrared (3–5 μm) hyperspectral data. *Remote Sensing*, 10(8):1237, 2018.

[407] Daniel J Jacob, Alexander J Turner, Joannes D Maasakkers, Jianxiong Sheng, Kang Sun, Xiong Liu, Kelly Chance, Ilse Aben, Jason McKeever, and Christian Frankenberg. Satellite observations of atmospheric methane and their value for quantifying methane emissions. *Atmospheric Chemistry and Physics*, 16(22):14371–14396, 2016.

[408] Peter F Bernath, Mahdi Yousefi, Eric Buzan, and Chris D Boone. A near-global atmospheric distribution of n2o isotopologues. *Geophysical Research Letters*, 44(20):10–735, 2017.

[409] G Philip Robertson and Peter M Vitousek. Nitrogen in agriculture: balancing the cost of an essential resource. *Annual review of environment and resources*, 34:97–125, 2009.

[410] Salah Sukkarieh. Mobile on-farm digital technology for smallholder farmers. Technical report, 2017.

[411] Asher Bender, Brett Whelan, and Salah Sukkarieh. Ladybird Cobbitty 2017 Brassica Dataset. 2019.

[412] Mirwaes Wahabzada, Anne-Katrin Mahlein, Christian Bauckhage, Ulrike Steiner, Erich-Christian Oerke, and Kristian Kersting. Plant phenotyping using probabilistic topic models: uncovering the hyperspectral language of plants. *Scientific reports*, 6:22482, 2016.

[413] Konstantinos Liakos, Patrizia Busato, Dimitrios Moshou, Simon Pearson, and Dionysis Bochtis. Machine learning in agriculture: A review. *Sensors*, 18(8):2674, 2018.

[414] Raphael A Viscarra Rossel and Johan Bouma. Soil sensing: A new paradigm for agriculture. *Agricultural Systems*, 148:71–74, 2016.

[415] Jiaxuan You, Xiaocheng Li, Melvin Low, David Lobell, and Stefano Ermon. Deep Gaussian process for crop yield prediction based on remote sensing data. In *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.

[416] Wei Ma, Kendall Nowocin, Niraj Marathe, and George H Chen. An interpretable produce price forecasting system for small and marginal farmers in india using collaborative filtering and adaptive nearest neighbors. In *Proceedings of the Tenth International Conference on Information and Communication Technologies and Development*, page 6. ACM, 2019.

[417] Agriculture, forestry, and fishing, value added. `https://data.worldbank.org/indicator/NV.AGR.TOTL.CD`, 2017.

[418] Chi-Hua Chen, Hsu-Yang Kung, and Feng-Jang Hwang. Deep learning techniques for agronomy applications. *Agronomy*, 9(3), 2019.

[419] Faizal Parish, AA Sirin, D Charman, Hans Joosten, T Yu Minaeva, and Marcel Silvius. Assessment on peatlands, biodiversity and climate change. 2008.

[420] Mike Flannigan, Chelene Krezek-Hanes, Mike Wotton, Mike Waddington, Merritt Turetsky, and Brian Benscoter. Peatland fires and carbon emissions (bulletin 50). Technical report, 2012.

[421] Susan E Page, Florian Siegert, John O Rieley, Hans-Dieter V Boehm, Adi Jaya, and Suwido Limin. The amount of carbon released from peat and forest fires in indonesia during 1997. *Nature*, 420(6911):61, 2002.

[422] Hans Joosten, Marja-Liisa Tapio-Biström, and Susanna Tol. *Peatlands: guidance for climate change mitigation through conservation, rehabilitation and sustainable use*. Food and Agriculture Organization of the United Nations, 2012.

[423] Joseph Holden, PJ Chapman, and JC Labadz. Artificial drainage of peatlands: hydrological and hydrochemical process and wetland restoration. *Progress in Physical Geography*, 28(1):95–123, 2004.

[424] Budiman Minasny, Budi Indra Setiawan, Satyanto Krido Saptomo, Alex B McBratney, et al. Open digital mapping as a cost-effective method for mapping peat thickness and assessing the carbon stock of tropical peatlands. *Geoderma*, 313:25–40, 2018.

[425] Claudia Windeck. A new global peatland map expected for 2020. `www.gislounge.com/new-global-peatland-map-expected-2020`, 2018.

[426] Pedro Rodríguez-Veiga, James Wheeler, Valentin Louis, Kevin Tansey, and Heiko Balzter. Quantifying forest biomass carbon stocks from space. *Current Forestry Reports*, 3(1):1–18, 2017.

[427] Tara O'Shea. Developing the world's first indicator of forest carbon stocks & emissions. https://www.planet.com/pulse/developing-the-worlds-first-indicator-of-forest-carbon-stocks-emissions/, 2019.

[428] Jean-Francois Bastin, Yelena Finegold, Claude Garcia, Danilo Mollicone, Marcelo Rezende, Devin Routh, Constantin M. Zohner, and Thomas W. Crowther. The global tree restoration potential. *Science*, 365(6448):76–79, 2019.

[429] Drones planting trees: An interview with BioCarbon engineering. https://medium.com/@ImpakterMag/drones-planting-trees-an-interview-with-biocarbon-engineering-33c536a22d5e.

[430] Anthony LeRoy Westerling. Increasing western US forest wildfire activity: sensitivity to changes in the timing of spring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1696):20150178, 2016.

[431] Claire A Montgomery. An agent and a consequence of land use change. *The Oxford Handbook of Land Economics*, page 281, 2014.

[432] 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.

[433] PG Brodrick, LDL Anderegg, and GP Asner. Forest drought resistance at large geographic scales. *Geophysical Research Letters*, 2019.

[434] Sriram Ganapathi Subramanian and Mark Crowley. Using spatial reinforcement learning to build forest wildfire dynamics models from satellite images. *Frontiers in ICT*, 5:6, 2018.

[435] Sriram Ganapathi Subramanian and Mark Crowley. Combining MCTS and A3C for prediction of spatially spreading processes in forest wildfire settings. In *Advances in Artificial Intelligence: 31st Canadian Conference on Artificial Intelligence, Canadian AI 2018, Toronto, ON, Canada, May 8–11, 2018, Proceedings 31*, pages 285–291. Springer, 2018.

[436] Rachel M Houtman, Claire A Montgomery, Aaron R Gagnon, David E Calkin, Thomas G Dietterich, Sean McGregor, and Mark Crowley. Allowing a wildfire to burn: estimating the effect on future fire suppression costs. *International Journal of Wildland Fire*, 22(7):871–882, 2013.

[437] K MacDicken, Ö Jonsson, L Piña, S Maulo, V Contessa, Y Adikari, M Garzuglia, E Lindquist, G Reams, and R D'Annunzio. *Global forest resources assessment 2015: how are the world's forests changing?* FAO, 2016.

[438] Matthew G Hethcoat, David P Edwards, Joao MB Carreiras, Robert G Bryant, Filipe M Franca, and Shaun Quegan. A machine learning approach to map tropical selective logging. *Remote Sensing of Environment*, 221:569–582, 2019.

[439] Christopher D Lippitt, John Rogan, Zhe Li, J Ronald Eastman, and Trevor G Jones. Mapping selective logging in mixed deciduous forest. *Photogrammetric Engineering & Remote Sensing*, 74(10):1201–1211, 2008.

[440] AGSJ Baccini, SJ Goetz, WS Walker, NT Laporte, M Sun, D Sulla-Menashe, J Hackler, PSA Beck, R Dubayah, MA Friedl, et al. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature climate change*, 2(3):182, 2012.

[441] Ruth S DeFries, Richard A Houghton, Matthew C Hansen, Christopher B Field, David Skole, and John Townshend. Carbon emissions from tropical deforestation and regrowth based on satellite observations for the 1980s and 1990s. *Proceedings of the National Academy of Sciences*, 99(22):14256–14261, 2002.

[442] Rainforest connection. https://rfcx.org.

[443] Silviaterra. https://www.silviaterra.com.

[444] Sabine Fuss, Josep G Canadell, Glen P Peters, Massimo Tavoni, Robbie M Andrew, Philippe Ciais, Robert B Jackson, Chris D Jones, Florian Kraxner, Nebosja Nakicenovic, et al. Betting on negative emissions. *Nature climate change*, 4(10):850, 2014.

[445] T Gasser, Céline Guivarch, K Tachiiri, CD Jones, and P Ciais. Negative emissions physically needed to keep global warming below 2C. *Nature communications*, 6:7958, 2015.

[446] Ocean Studies Board, Engineering National Academies of Sciences, Medicine, et al. *Negative Emissions Technologies and Reliable Sequestration: A Research Agenda*. National Academies Press, 2019.

[447] David Sandalow, Julio Friedmann, and Colin McCormick. Direct air capture of carbon dioxide: ICEF roadmap 2018. 2018.

[448] Jan C Minx, William F Lamb, Max W Callaghan, Sabine Fuss, Jerome Hilaire, Felix Creutzig, Thorben Amann, Tim Beringer, Wagner de Oliveira Garcia, Jens Hartmann, et al. Negative emissions part 1: Research landscape and synthesis. *Environmental Research Letters*, 13(6):063001, 2018.

[449] Sabine Fuss, William F Lamb, Max W Callaghan, Jérôme Hilaire, Felix Creutzig, Thorben Amann, Tim Beringer, Wagner de Oliveira Garcia, Jens Hartmann, Tarun Khanna, et al. Negative emissions part 2: Costs, potentials and side effects. *Environmental Research Letters*, 13(6):063002, 2018.

[450] Gregory F Nemet, Max W Callaghan, Felix Creutzig, Sabine Fuss, Jens Hartmann, Jérôme Hilaire, William F Lamb, Jan C Minx, Sophia Rogers, and Pete Smith. Negative emissions part 3: Innovation and upscaling. *Environmental Research Letters*, 13(6):063003, 2018.

[451] Felix Creutzig, Nijavalli H Ravindranath, Göran Berndes, Simon Bolwig, Ryan Bright, Francesco Cherubini, Helena Chum, Esteve Corbera, Mark Delucchi, Andre Faaij, et al. Bioenergy and climate change mitigation: an assessment. *Gcb Bioenergy*, 7(5):916–944, 2015.

[452] Carmenza Robledo-Abad, Hans-Jörg Althaus, Göran Berndes, Simon Bolwig, Esteve Corbera, Felix Creutzig, John Garcia-Ulloa, Anna Geddes, Jay S Gregg, Helmut Haberl, et al. Bioenergy production and sustainable development: science base for policymaking remains limited. *Gcb Bioenergy*, 9(3):541–556, 2017.

[453] RD Schuiling and P Krijgsman. Enhanced weathering: an effective and cheap tool to sequester CO2. *Climatic Change*, 74(1-3):349–354, 2006.

[454] Edward S. Rubin, John E. Davison, and Howard J. Herzog. The cost of CO2 capture and storage. *International Journal of Greenhouse Gas Control*, 40:378–400, September 2015.

[455] Felix Creutzig, Christian Breyer, Jerome Hilaire, Jan Minx, Glen Peters, and Robert H Socolow. The mutual dependence of negative emission technologies and energy systems. *Energy & Environmental Science*, 2019.

[456] V Zeleňák, M Badaničová, D Halamova, J Čejka, A Zukal, N Murafa, and G Goerigk. Amine-modified ordered mesoporous silica: effect of pore size on carbon dioxide capture. *Chemical Engineering Journal*, 144(2):336–342, 2008.

[457] Veronica B Cashin, Daniel S Eldridge, Aimin Yu, and Dongyuan Zhao. Surface functionalization and manipulation of mesoporous silica adsorbents for improved removal of pollutants: a review. *Environmental Science: Water Research & Technology*, 4(2):110–128, 2018.

[458] 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 discovery using failed experiments. *Nature*, 533(7601):73, 2016.

[459] Geoffrey Holmes and David W Keith. An air–liquid contactor for large-scale capture of CO2 from air. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370(1974):4380–4403, 2012.

[460] M. D. Zoback and S. M. Gorelick. Earthquake triggering and large-scale geologic storage of carbon dioxide. *Proceedings of the National Academy of Sciences*, 109(26):10164–10168, June 2012.

[461] Sandra Ó Snæbjörnsdóttir and Sigurdur R Gislason. CO2 storage potential of basaltic rocks offshore Iceland. *Energy Procedia*, 86:371–380, 2016.

[462] Mauricio Araya-Polo, Joseph Jennings, Amir Adler, and Taylor Dahlke. Deep-learning tomography. *The Leading Edge*, 37(1):58–66, 2018.

[463] MA Celia, S Bachu, JM Nordbotten, and KW Bandilla. Status of CO2 storage in deep saline aquifers with emphasis on modeling approaches and practical simulations. *Water Resources Research*, 51(9):6846–6892, 2015.

[464] 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 Resources Research*, 55(1):703–728, 2019.

[465] Dylan Moriarty, Laura Dobeck, and Sally Benson. Rapid surface detection of CO2 leaks from geologic sequestration sites. *Energy Procedia*, 63:3975–3983, 2014.

[466] Bailian Chen, Dylan R Harp, Youzuo Lin, Elizabeth H Keating, and Rajesh J Pawar. Geologic co 2 sequestration monitoring design: A machine learning and uncertainty quantification based approach. *Applied energy*, 225:332–345, 2018.

[467] Jingfan Wang, Lyne P. Tchapmi, Arvind P. Ravikumar, Mike McGuire, Clay S. Bell, Daniel Zimmerle, Silvio Savarese, and Adam R. Brandt. Machine vision for natural gas methane emissions detection using an infrared camera. *Applied Energy*, 257:113998, 2020.

[468] H Goosse, P Barriat, W Lefebvre, M Loutre, and V Zunz. *Introduction to climate dynamics and climate modeling*. 2008–2010.

[469] IPCC. *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. 2014.

[470] K E Taylor, R J Stouffer, and G A Meehl. An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4):485–498, 2012.

[471] V Eyring, S Bony, G A Meehl, C A Senior, R J Stouffer, and K E Taylor. Overview of the Coupled Model Intercomparison Project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 6(LLNL-JRNL-736881), 2016.

[472] J Kay, C Deser, A Phillips, A Mai, C Hannay, G Strand, J M Arblaster, S C Bates, G Danabasoglu, J Edwards, M Holland, P Kushner, J-F Lamarque, D Lawrence, K Lindsay, A Middleton, E Munoz, R Neale, K Oleson, L Polvani, and M Vertenstein. The Community Earth System Model (CESM) Large Ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bulletin of the American Meteorological Society*, 96(8):1333–1349, 2015.

[473] J Carman, T Clune, F Giraldo, M Govett, B Gross, A Kamrathe, T Lee, D McCarren, J Michalakes, S Sandgathe, and T Whitcomb. Position paper on high performance computing needs in Earth system prediction. National Earth System Prediction Capability. Technical report, 2017.

[474] D J Lary. Artificial intelligence in geoscience and remote sensing. In *Aerospace Technologies Advancements, edited*. 2010.

[475] David J. Lary, Amir H. Alavi, Amir H. Gandomi, and Annette L. Walker. Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, pages 1–9, 2015.

[476] Nataliia Kussul, Mykola Lavreniuk, Sergii Skakun, and Andrii Shelestov. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, 14(5):778–782, 2017.

[477] Y Gil, S. Pierce, Hassan Babaie, Arindam Banerjee, Kirk Borne, Gary Bust, Michelle Cheatham, Imme Ebert-Uphoff, Carla Gomes, Mary Hill, John Horel, Leslie Hsu, Jim Kinter, Craig Knoblock, David Krum, Vipin Kumar, Pierre Lermusiaux, Yan Liu, Chris North, Victor Pankratius, Shanan Peters, Beth Plale, Allen Pope, Sai Ravela, Juan Restrepo, Aaron Ridley, Hanan Samet, and Shashi Shekhar. Intelligent systems for geosciences: An essential research agenda. *Communications of the ACM*, 62:76–84, January 2019.

[478] Surya Karthik Mukkavilli. EnviroNet: ImageNet for environment. In *18th Conference on Artificial and Computational Intelligence and its Applications to the Environmental Sciences*. American Meteorological Society 99th Annual Meeting, 2019.

[479] Imme Ebert-Uphoff, David Thompson, Ibrahim Demir, Yulia Gel, Mary Hill, Anuj Karpatne, Mariana Guereque, Vipin Kumar, Enrique Cabal-Cano, and Padhraic Smyth. A vision for the development of benchmarks to bridge geoscience and data science. *17th International Workshop on Climate Informatics*, 2017.

[480] Evan Racah, Christopher Beckham, Tegan Maharaj, Samira Ebrahimi Kahou, Prabhat, and Chris Pal. ExtremeWeather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events. In *Advances in Neural Information Processing Systems 30*, pages 3402–3413. 2017.

[481] Frederic Hourdin, Thorsten Mauritsen, Andrew Gettelman, Jean-Christophe Golaz, Venkatramani Balaji, Qingyun Duan, Doris Folini, Duoying Ji, Daniel Klocke, Yun Qian, Florian Rauser, Catherine Rio, Lorenzo Tomassini, Masahiro Watanabe, and Daniel Williamson. The art and science of climate model tuning. *Bulletin of the American Meteorological Society*, 98(3):589–602, 2017.

[482] Steven C Sherwood, Sandrine Bony, and Jean-louis Dufresne. Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature*, 505:37–42, 2014.

[483] P Gentine, M Pritchard, S Rasp, G Reinaudi, and G Yacalis. Could machine learning break the convection parameterization deadlock? *Geophysical Research Letters*, 45:5742–5751, 2018.

[484] Stephan Rasp, Michael S Pritchard, and Pierre Gentine. Deep learning to represent subgrid processes in climate models. *Proceedings of the National Academy of Sciences*, 115(39):1–6, 2018.

[485] Markus Reichstein, Gustau Camps-Valls, Bjorn Stevens, Martin Jung, Joachim Denzler, Nuno Carvalhais, and Prabhat. Deep learning and process understanding for data-driven earth system science. *Nature*, 566(7743):195–204, 2019.

[486] Robert E Kopp, Robert M Deconto, Daniel A Bader, Carling C Hay, M Radley, Scott Kulp, Michael Oppenheimer, David Pollard, and Benjamin H Strauss. Evolving understanding of Antarctic ice-sheet physics and ambiguity in probabilistic sea-level projections. *Earth's Future*, 5(12):1217–1233, 2017.

[487] M.-È. Gagné, N. P. Gillett, and J. C. Fyfe. Observed and simulated changes in Antarctic sea ice extent over the past 50 years. *Geophysical Research Letters*, 42:90–95, 2015.

[488] Edward Hanna, Francisco J Navarro, Frank Pattyn, Catia M Domingues, Xavier Fettweis, Erik R Ivins, Robert J Nicholls, Catherine Ritz, Ben Smith, Slawek Tulaczyk, Pippa L Whitehouse, and H Jay Zwally. Ice-sheet mass balance and climate change. *Nature*, 498(7452):51–59, 2013.

[489] Peer Nowack, Peter Braesicke, Joanna Haigh, Nathan Luke Abraham, and John Pyle. Using machine learning to build temperature-based ozone parameterizations for climate sensitivity simulations. *Environmental Research Letters*, 13(104016), 2018.

[490] Claudia Tebaldi and Reto Knutti. The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A*, 365:2053–2075, 2007.

[491] Claire Monteleoni, Gavin A Schmidt, Shailesh Saroha, and Eva Asplund. Tracking climate models. *Statistical Analysis and Data Mining*, 4:372–392, 2011.

[492] Scott Mcquade and Claire Monteleoni. Global climate model tracking using geospatial neighborhoods. *Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.

[493] E Strobach and G Bel. Improvement of climate predictions and reduction of their uncertainties using learning algorithms. *Atmospheric Chemistry and Physics*, 15:8631–8641, 2015.

[494] Gemma Anderson and Donald D Lucas. Machine learning predictions of a multiresolution climate model ensemble. *Geophysical Research Letters*, 45:4273–4280, 2018.

[495] Tapio Schneider, Shiwei Lan, Andrew Stuart, and João Teixeira. Earth system modeling 2.0 : a blueprint for models that learn from observations and targeted high-resolution simulations. *Geophysical Research Letters*, 44:12396–12417, 2017.

[496] J Shukla. Predictability in the midst of chaos: a scientific basis for climate forecasting. *Science*, 282:728–731, 1998.

[497] Judah Cohen, Dim Coumou, Jessica Hwang, Lester Mackey, Paulo Orenstein, Sonja Totz, and Eli Tziperman. S2S reboot: an argument for greater inclusion of machine learning in subseasonal to seasonal forecasts. *WIREs Climate Change*, 10, 2018.

[498] Jessica Hwang, Paulo Orenstein, Judah Cohen, Karl Pfeiffer, and Lester Mackey. Improving subseasonal forecasting in the western U.S. with machine learning. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019.

[499] David John Gagne, Amy McGovern, Sue Ellen Haupt, Ryan A Sobash, John K Williams, and Ming Xue. Storm-based probabilistic hail forecasting with machine learning applied to convection-allowing ensembles. *Weather and forecasting*, 32(5):1819–1840, 2017.

[500] Amy McGovern, Kimberly L Elmore, David John Gagne, Sue Ellen Haupt, Christopher D Karstens, Ryan Lagerquist, Travis Smith, and John K Williams. Using artificial intelligence to improve real-time decision-making for high-impact weather. *Bulletin of the American Meteorological Society*, 98(10):2073–2090, 2017.

[501] Yunjie Liu, Evan Racah, Prabhat, Joaquin Correa, Amir Khosrowshahi, David Lavers, Kenneth Kunkel, Michael Wehner, and William Collins. Application of deep convolutional neural networks for detecting extreme weather in climate datasets. *International Conference on Advances in Big Data Analytics*, 2016.

[502] Thorsten Kurth, Sean Treichler, Joshua Romero, Mayur Mudigonda, Nathan Luehr, Everett Phillips, Ankur Mahesh, Michael Matheson, Jack Deslippe, Massimiliano Fatica, Prabhat, and Michael Houston. Exascale deep learning for climate analytics. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis*, SC '18, pages 51:1–51:12, Piscataway, NJ, USA, 2018. IEEE Press.

[503] Valliappa Lakshmanan and Travis Smith. An objective method of evaluating and devising storm-tracking algorithms. *Weather and Forecasting*, 25:701–709, 2010.

[504] Wan Li, Li Ni, Zhao-liang Li, Si-bo Duan, and Hua Wu. Evaluation of machine learning algorithms in spatial downscaling of MODIS land surface temperature. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12:2299–2307, 2019.

[505] MC Perignon, P Passalacqua, TM Jarriel, JM Adams, and I Overeem. Patterns of geomorphic processes across deltas using image analysis and machine learning. In *AGU Fall Meeting Abstracts*, 2018.

[506] Muhammed Sit and Ibrahim Demir. Decentralized flood forecasting using deep neural networks. *Preprint arXiv:1902.02308*, 2019.

[507] Maziar Raissi and George Em Karniadakis. Hidden physics models: machine learning of nonlinear partial differential equations. *Journal of Computational Physics*, 357:125–141, 2018.

[508] Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics informed deep learning (Part I): data-driven solutions of nonlinear partial differential equations. *Preprint*, 2017.

[509] D. D. Lucas, R. Klein, J. Tannahill, D. Ivanova, S. Brandon, D. Domyancic, and Y. Zhang. Failure analysis of parameter-induced simulation crashes in climate models. *Geoscientific Model Development*, 6:1157–1171, 2013.

[510] M Jiang, B Gallagher, J Kallman, and D Laney. A supervised learning framework for arbitrary Lagrangian-Eulerian simulations. In *15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Anaheim, CA, 2016.

[511] J Ling and J Templeton. Evaluation of machine learning algorithms for prediction of regions of high Reynolds averaged Navier Stokes uncertainty. *Physics of Fluids*, 27(085103), 2015.

[512] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles. *Advances in Neural Information Processing Systems*, 2017.

[513] Jayaraman Thiagarajan, Nikhil Jain, Rushil Anirudh, Alfredo Giminez, Rahul Sridhar, Marathe Aniruddha, Tao Wang, Mural Emani, Abhinav Bhatele, and Todd Gamblin. Bootstrapping parameter space exploration for fast tuning. In *Proceedings of the 2018 International Conference on Supercomputing*, pages 385–395, 2018.

[514] D J Lary, G K Zewdie, X Liu, D Wu, E Levetin, Allee R J, Nabin Malakar, A Walker, H Mussa, Mannino A, and Aurin D. Machine learning for applications for Earth observation. *Earth Observation Open Science and Innovation*, (165), 2018.

[515] John Quinn, Vanessa Frias-Martinez, and Lakshminarayan Subramanian. Computational sustainability and artificial intelligence in the developing world. *AI Magazine*, 35(3):36, 2014.

[516] Christopher Potter, Shyam Boriah, Michael Steinbach, Vipin Kumar, and Steven Klooster. Terrestrial vegetation dynamics and global climate controls. *Climate Dynamics*, 31(1):67–78, 2008.

[517] Shyam Boriah, Vipin Kumar, Michael Steinbach, Christopher Potter, and Steven Klooster. Land cover change detection: a case study. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 857–865. ACM, 2008.

[518] Kolya Malkin, Caleb Robinson, Le Hou, Rachel Soobitsky, Jacob Czawlytko, Dimitris Samaras, Joel Saltz, Lucas Joppa, and Nebojsa Jojic. Label super-resolution networks. 2018.

[519] Lior Bragilevsky and Ivan V Bajić. Deep learning for Amazon satellite image analysis. In *2017 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*, pages 1–5. IEEE, 2017.

[520] Nate G McDowell, Nicholas C Coops, Pieter SA Beck, Jeffrey Q Chambers, Chandana Gangodagamage, Jeffrey A Hicke, Cho-ying Huang, Robert Kennedy, Dan J Krofcheck, Marcy Litvak, et al. Global satellite monitoring of climate-induced vegetation disturbances. *Trends in plant science*, 20(2):114–123, 2015.

[521] Duy Huynh and Nathalie Neptune. Annotation automatique d'images: le cas de la déforestation. In *Actes de la conférence Traitement Automatique de la Langue Naturelle, TALN 2018*, page 101.

[522] Kirk R Klausmeyer and M Rebecca Shaw. Climate change, habitat loss, protected areas and the climate adaptation potential of species in Mediterranean ecosystems worldwide. *PloS one*, 4(7):e6392, 2009.

[523] Xiaohui Feng, María Uriarte, Grizelle González, Sasha Reed, Jill Thompson, Jess K Zimmerman, and Lora Murphy. Improving predictions of tropical forest response to climate change through integration of field studies and ecosystem modeling. *Global change biology*, 24(1):e213–e232, 2018.

[524] Jane K Hart and Kirk Martinez. Environmental sensor networks: A revolution in the earth system science? *Earth-Science Reviews*, 78(3-4):177–191, 2006.

[525] RW Hut, NC van de Giesen, and JS Selker. The TAHMO project: Designing an unconventional weather station. In *EGU General Assembly Conference Abstracts*, volume 14, page 8963, 2012.

[526] G Griffiths, NW Millard, SD McPhail, P Stevenson, JR Perrett, M Peabody, AT Webb, and DT Meldrum. Towards environmental monitoring with the Autosub autonomous underwater vehicle. In *Proceedings of 1998 International Symposium on Underwater Technology*, pages 121–125. IEEE, 1998.

[527] Matthew Dunbabin and Lino Marques. Robots for environmental monitoring: Significant advancements and applications. *IEEE Robotics & Automation Magazine*, 19(1):24–39, 2012.

[528] Ethan W Dereszynski and Thomas G Dietterich. Probabilistic models for anomaly detection in remote sensor data streams. *Preprint arXiv:1206.5250*, 2012.

[529] David J Hill and Barbara S Minsker. Anomaly detection in streaming environmental sensor data: A data-driven modeling approach. *Environmental Modelling & Software*, 25(9):1014–1022, 2010.

[530] Jnaneshwar Das, Frédéric Py, Julio BJ Harvey, John P Ryan, Alyssa Gellene, Rishi Graham, David A Caron, Kanna Rajan, and Gaurav S Sukhatme. Data-driven robotic sampling for marine ecosystem monitoring. *The International Journal of Robotics Research*, 34(12):1435–1452, 2015.

[531] Genevieve Flaspohler, Nicholas Roy, and Yogesh Girdhar. Feature discovery and visualization of robot mission data using convolutional autoencoders and bayesian nonparametric topic models. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1–8. IEEE, 2017.

[532] Jochem Marotzke, Christian Jakob, Sandrine Bony, Paul A Dirmeyer, Paul A O'Gorman, Ed Hawkins, Sarah Perkins-Kirkpatrick, Corinne Le Quere, Sophie Nowicki, Katsia Paulavets, et al. Climate research must sharpen its view. *Nature climate change*, 7(2):89, 2017.

[533] Project Zamba computer vision for wildlife research & conservation. `https://zamba.drivendata.org/`.

[534] Sara Beery, Yang Liu, Dan Morris, Jim Piavis, Ashish Kapoor, Markus Meister, and Pietro Perona. Synthetic examples improve generalization for rare classes. *Preprint arXiv:1904.05916*, 2019.

[535] Mohammad Sadegh Norouzzadeh, Anh Nguyen, Margaret Kosmala, Alexandra Swanson, Meredith S Palmer, Craig Packer, and Jeff Clune. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences*, 115(25):E5716–E5725, 2018.

[536] Jan C van Gemert, Camiel R Verschoor, Pascal Mettes, Kitso Epema, Lian Pin Koh, and Serge Wich. Nature conservation drones for automatic localization and counting of animals. In *European Conference on Computer Vision*, pages 255–270. Springer, 2014.

[537] Dana M Ghioca-Robrecht, Carol A Johnston, and Mirela G Tulbure. Assessing the use of multiseason quickbird imagery for mapping invasive species in a lake erie coastal marsh. *Wetlands*, 28(4):1028–1039, 2008.

[538] Robin Faillettaz, Marc Picheral, Jessica Y Luo, Cédric Guigand, Robert K Cowen, and Jean-Olivier Irisson. Imperfect automatic image classification successfully describes plankton distribution patterns. *Methods in Oceanography*, 15:60–77, 2016.

[539] Grace Young, Vassileios Balntas, and Victor Prisacariu. Convolutional neural networks predict fish abundance from underlying coral reef texture. *MarXiv. August*, 31, 2018.

[540] Brian L Sullivan, Christopher L Wood, Marshall J Iliff, Rick E Bonney, Daniel Fink, and Steve Kelling. ebird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, 142(10):2282–2292, 2009.

[541] PlantSnap. Homepage. `https://www.plantsnap.com/`.

[542] Simone Branchini, Francesco Pensa, Patrizia Neri, Bianca Maria Tonucci, Lisa Mattielli, Anna Collavo, Maria Elena Sillingardi, Corrado Piccinetti, Francesco Zaccanti, and Stefano Goffredo. Using a citizen science program to monitor coral reef biodiversity through space and time. *Biodiversity and conservation*, 24(2):319–336, 2015.

[543] Sreejith Menon, Tanya Berger-Wolf, Emre Kiciman, Lucas Joppa, Charles V Stewart, Jason Parham, Jonathan Crall, Jason Holmberg, and Jonathan Van Oast. Animal population estimation using flickr images. 2016.

[544] Jeffrey F Kelly, Kyle G Horton, Phillip M Stepanian, Kirsten M de Beurs, Todd Fagin, Eli S Bridge, and Phillip B Chilson. Novel measures of continental-scale avian migration phenology related to proximate environmental cues. *Ecosphere*, 7(9), 2016.

[545] Grant Van Horn, Steve Branson, Ryan Farrell, Scott Haber, Jessie Barry, Panos Ipeirotis, Pietro Perona, and Serge Belongie. Building a bird recognition app and large scale dataset with citizen scientists: The fine print in fine-grained dataset collection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 595–604, 2015.

[546] Eric Ralls. Systems and methods for electronically identifying plant species, November 8 2018. US Patent App. 15/973,660.

[547] Grant Van Horn and Pietro Perona. The devil is in the tails: Fine-grained classification in the wild. *Preprint arXiv:1709.01450*, 2017.

[548] Yexiang Xue, Ian Davies, Daniel Fink, Christopher Wood, and Carla P Gomes. Avicaching: A two stage game for bias reduction in citizen science. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pages 776–785. International Foundation for Autonomous Agents and Multiagent Systems, 2016.

[549] Di Chen and Carla P Gomes. Bias reduction via end-to-end shift learning: Application to citizen science. *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence*, 2019.

[550] Pushpendra Rana and Daniel C Miller. Machine learning to analyze the social-ecological impacts of natural resource policy: insights from community forest management in the indian himalaya. *Environmental Research Letters*, 2018.

[551] Heidi J Albers, Kim Meyer Hall, Katherine D Lee, Majid Alkaee Taleghan, and Thomas G Dietterich. The role of restoration and key ecological invasion mechanisms in optimal spatial-dynamic management of invasive species. *Ecological Economics*, 151:44–54, 2018.

[552] Andreas Lydakis, Jenica M Allen, Marek Petrik, and Tim Szewczyk. Computing robust strategies for managing invasive plants.

[553] Rajendra K Pachauri. *Climate Change 2014 Synthesis Report*. 2014.

[554] Mark Pelling. *Adaptation to climate change: from resilience to transformation*. Routledge, 2010.

[555] Linda Shi, Eric Chu, Isabelle Anguelovski, Alexander Aylett, Jessica Debats, Kian Goh, Todd Schenk, Karen C Seto, David Dodman, Debra Roberts, et al. Roadmap towards justice in urban climate adaptation research. *Nature Climate Change*, 6(2):131, 2016.

[556] Amrita Gupta, Caleb Robinson, and Bistra Dilkina. Infrastructure resilience for climate adaptation. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, page 28. ACM, 2018.

[557] Vanessa Frias-Martinez, Cristina Soguero, and Enrique Frias-Martinez. Estimation of urban commuting patterns using cellphone network data. In *Proceedings of the ACM SIGKDD international workshop on urban computing*, pages 9–16. ACM, 2012.

[558] Vipin Jain, Ashlesh Sharma, and Lakshminarayanan Subramanian. Road traffic congestion in the developing world. In *Proceedings of the 2nd ACM Symposium on Computing for Development*, page 11. ACM, 2012.

[559] David Pastor-Escuredo, Alfredo Morales-Guzmán, Yolanda Torres-Fernández, Jean-Martin Bauer, Amit Wadhwa, Carlos Castro-Correa, Liudmyla Romanoff, Jong Gun Lee, Alex Rutherford, Vanessa Frias-Martinez, et al. Flooding through the lens of mobile phone activity. In *IEEE Global Humanitarian Technology Conference (GHTC 2014)*, pages 279–286. IEEE, 2014.

[560] Ami Wiesel, Avinatan Hassidim, Gal Elidan, Guy Shalev, Mor Schlesinger, Oleg Zlydenko, Ran El-Yaniv, Sella Nevo, Yossi Matias, Yotam Gigi, et al. Ml for flood forecasting at scale. 2018.

[561] Barak Oshri, Annie Hu, Peter Adelson, Xiao Chen, Pascaline Dupas, Jeremy Weinstein, Marshall Burke, David Lobell, and Stefano Ermon. Infrastructure quality assessment in africa using satellite imagery and deep learning. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 616–625. ACM, 2018.

[562] Fitore Muharemi, Doina Logofătu, and Florin Leon. Machine learning approaches for anomaly detection of water quality on a real-world data set. *Journal of Information and Telecommunication*, pages 1–14, 2019.

[563] Roshanak Nateghi. Multi-dimensional infrastructure resilience modeling: An application to hurricane-prone electric power distribution systems. *IEEE Access*, 6:13478–13489, 2018.

[564] Mathaios Panteli and Pierluigi Mancarella. The grid: Stronger bigger smarter?: Presenting a conceptual framework of power system resilience. *IEEE Power Energy Mag*, 13(3):58–66, 2015.

[565] Xi Fang, Satyajayant Misra, Guoliang Xue, and Dejun Yang. Smart grid—the new and improved power grid: A survey. *IEEE communications surveys & tutorials*, 14(4):944–980, 2012.

[566] Sarah Fletcher, Megan Lickley, and Kenneth Strzepek. Learning about climate change uncertainty enables flexible water infrastructure planning. *Nature communications*, 10(1):1782, 2019.

[567] I Delpla, A-V Jung, E Baures, M Clement, and O Thomas. Impacts of climate change on surface water quality in relation to drinking water production. *Environment international*, 35(8):1225–1233, 2009.

[568] The water, peace and security partnership. *Institute for Water Education website*, 2019.

[569] Julianne D Quinn, Patrick M Reed, and Klaus Keller. Direct policy search for robust multi-objective management of deeply uncertain socio-ecological tipping points. *Environmental Modelling & Software*, 92:125–141, 2017.

[570] Matteo Giuliani, Andrea Castelletti, Francesca Pianosi, Emanuele Mason, and Patrick M Reed. Curses, tradeoffs, and scalable management: Advancing evolutionary multiobjective direct policy search to improve water reservoir operations. *Journal of Water Resources Planning and Management*, 142(2):04015050, 2015.

[571] Chaopeng Shen. A trans-disciplinary review of deep learning research for water resources scientists. *Preprint arXiv:1712.02162*, 2017.

[572] Satyam Srivastava, Saikrishna Vaddadi, Pankaj Kumar, and Shashikant Sadistap. Design and development of reverse osmosis (RO) plant status monitoring system for early fault prediction and predictive maintenance. *Applied Water Science*, 8(6):159, 2018.

[573] Otilia Elena Dragomir, Rafael Gouriveau, Florin Dragomir, Eugenia Minca, and Noureddine Zerhouni. Review of prognostic problem in condition-based maintenance. In *2009 European Control Conference (ECC)*, pages 1587–1592. IEEE, 2009.

[574] Zubair A Baig. On the use of pattern matching for rapid anomaly detection in smart grid infrastructures. In *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 214–219. IEEE, 2011.

[575] Djellel Eddine Difallah, Philippe Cudre-Mauroux, and Sean A McKenna. Scalable anomaly detection for smart city infrastructure networks. *IEEE Internet Computing*, 17(6):39–47, 2013.

[576] JR Porter, L Xie, AJ Challinor, K Cochrane, MM Howden, DB Lobell, and MI Travasso. Food security and food production systems. *Climate Change 2014: Impacts, Adaptation, Vulnerability.*, pages 485–533, 2014.

[577] Aiguo Dai. Drought under global warming: a review. *Wiley Interdisciplinary Reviews: Climate Change*, 2(1):45–65, 2011.

[578] Adeline Decuyper, Alex Rutherford, Amit Wadhwa, Jean-Martin Bauer, Gautier Krings, Thoralf Gutierrez, Vincent D Blondel, and Miguel A Luengo-Oroz. Estimating food consumption and poverty indices with mobile phone data. *Preprint arXiv:1412.2595*, 2014.

[579] 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.

[580] Jaewoo Kim, Meeyoung Cha, and Jong Gun Lee. Nowcasting commodity prices using social media. *PeerJ Computer Science*, 3:e126, 2017.

[581] S. Chakraborty and A. C. Newton. Climate change, plant diseases and food security: an overview. *Plant Pathology*, 60(1):2–14, jan 2011.

[582] Anna X Wang, Caelin Tran, Nikhil Desai, David Lobell, and Stefano Ermon. Deep transfer learning for crop yield prediction with remote sensing data. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, page 50. ACM, 2018.

[583] C Tebaldi and DB Lobell. Towards probabilistic projections of climate change impacts on global crop yields. *Geophysical Research Letters*, 35(8), 2008.

[584] Cynthia Rosenzweig, Joshua Elliott, Delphine Deryng, Alex C Ruane, Christoph Müller, Almut Arneth, Kenneth J Boote, Christian Folberth, Michael Glotter, Nikolay Khabarov, et al. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proceedings of the National Academy of Sciences*, 111(9):3268–3273, 2014.

[585] Venkata Shashank Konduri, Jitendra Kumar, Forrest Hoffman, Udit Bhatia, Tarik Gouthier, and Auroop Ganguly. Physics-guided data science for food security and climate. KDD Feed Workshop 2019.

[586] Michela Paganini, Luke de Oliveira, and Benjamin Nachman. Accelerating science with generative adversarial networks: an application to 3D particle showers in multilayer calorimeters. *Physical review letters*, 120(4):042003, 2018.

[587] Max Welling. Are ML and statistics complementary? In *IMS-ISBA Meeting on 'Data Science in the Next 50 Years*, 2015.

[588] Marie-Ève Rancourt, Jean-François Cordeau, Gilbert Laporte, and Ben Watkins. Tactical network planning for food aid distribution in Kenya. *Computers & Operations Research*, 56:68–83, 2015.

[589] Gautam Prasad, Upendra Reddy Vuyyuru, and Mithun Das Gupta. Agriculture commodity arrival prediction using remote sensing data: Insights and beyond. KDD Feed Workshop 2019. `https://drive.google.com/file/d/1BQ5QH036yifiza8TOKt_8FbimYyQ0SYA/view`.

[590] DrivenData. Mapping agricultural supply chains from source to shelf. `http://drivendata.co/case-studies/mapping-agricultural-supply-chains-from-source-to-shelf/`.

[591] Ernest Mwebaze, Washington Okori, and John Alexander Quinn. Causal structure learning for famine prediction. In *2010 AAAI Spring Symposium Series*, 2010.

[592] Arun Agrawal and Nicolas Perrin. Climate adaptation, local institutions and rural livelihoods. *Adapting to climate change: thresholds, values, governance*, pages 350–367, 2009.

[593] Daivi Rodima-Taylor. Social innovation and climate adaptation: Local collective action in diversifying Tanzania. *Applied Geography*, 33:128–134, 2012.

[594] Solomon Assefa. Hello Tractor pilot agriculture digital wallet based on AI and blockchain.

[595] UN Global Pulse. Landscaping study: Digital signals & access to finance in Kenya, Sep 2013.

[596] Vanessa Frias-Martinez, Victor Soto, Jesus Virseda, and Enrique Frias-Martinez. Computing cost-effective census maps from cell phone traces. In *Workshop on pervasive urban applications*, 2012.

[597] Vukosi Marivate and Nyalleng Moorosi. Employment relations: a data driven analysis of job markets using online job boards and online professional networks. In *Proceedings of the International Conference on Web Intelligence*, pages 1110–1113. ACM, 2017.

[598] Kirk Bansak, Jeremy Ferwerda, Jens Hainmueller, Andrea Dillon, Dominik Hangartner, Duncan Lawrence, and Jeremy Weinstein. Improving refugee integration through data-driven algorithmic assignment. *Science*, 359(6373):325–329, 2018.

[599] UN Global Pulse. Improving professional training in Indonesia with gaming data, 2017.

[600] Emilio Zagheni, Ingmar Weber, Krishna Gummadi, et al. Leveraging Facebook's advertising platform to monitor stocks of migrants. *Population and Development Review*, 43(4):721–734, 2017.

[601] Sibren Isaacman, Vanessa Frias-Martinez, Lingzi Hong, and Enrique Frias-Martinez. Climate change induced migrations from a cell phone perspective. *NetMob*, page 46, 2017.

[602] Joshua E Blumenstock. Inferring patterns of internal migration from mobile phone call records: evidence from Rwanda. *Information Technology for Development*, 18(2):107–125, 2012.

[603] John A Quinn, Marguerite M Nyhan, Celia Navarro, Davide Coluccia, Lars Bromley, and Miguel Luengo-Oroz. Humanitarian applications of machine learning with remote-sensing data: review and case study in refugee settlement mapping. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2128):20170363, 2018.

[604] Katherine Hoffmann Pham, Jeremy Boy, and Miguel Luengo-Oroz. Data fusion to describe and quantify search and rescue operations in the Mediterranean sea. In *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, pages 514–523. IEEE, 2018.

[605] Vincenzo Lomonaco, Angelo Trotta, Marta Ziosi, Juan De Dios Yáñez Ávila, and Natalia Díaz-Rodríguez. Intelligent drone swarm for search and rescue operations at sea. *Preprint arXiv:1811.05291*, 2018.

[606] UN Global Pulse. Social media and forced displacement: Big data analytics & machine learning, Sep 2017.

[607] Andy Haines, R Sari Kovats, Diarmid Campbell-Lendrum, and Carlos Corvalán. Climate change and human health: impacts, vulnerability and public health. *Public health*, 120(7):585–596, 2006.

[608] MC Sarofim, Shubhayu Saha, MD Hawkins, DM Mills, Jeremy J Hess, Radley M Horton, Patrick L Kinney, Joel D Schwartz, and Alexis St Juliana. The impacts of climate change on human health in the United States: a scientific assessment. *The Impacts of Climate Change on Human Health in the United States: A Scientific Assessment*, 2016.

[609] Joel Schwartz, Jonathan M Samet, and Jonathan A Patz. Hospital admissions for heart disease: the effects of temperature and humidity. *Epidemiology*, 15(6):755–761, 2004.

[610] Francesca Dominici, Roger D Peng, Michelle L Bell, Luu Pham, Aidan McDermott, Scott L Zeger, and Jonathan M Samet. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *Jama*, 295(10):1127–1134, 2006.

[611] Muin J Khoury, Tram Kim Lam, John PA Ioannidis, Patricia Hartge, Margaret R Spitz, Julie E Buring, Stephen J Chanock, Robert T Croyle, Katrina A Goddard, Geoffrey S Ginsburg, et al. Transforming epidemiology for 21st century medicine and public health. *Cancer Epidemiology and Prevention Biomarkers*, 22(4):508–516, 2013.

[612] Marcel Salathe, Linus Bengtsson, Todd J Bodnar, Devon D Brewer, John S Brownstein, Caroline Buckee, Ellsworth M Campbell, Ciro Cattuto, Shashank Khandelwal, Patricia L Mabry, et al. Digital epidemiology. *PLoS computational biology*, 8(7):e1002616, 2012.

[613] Nicholas Clinton and Peng Gong. MODIS detected surface urban heat islands and sinks: Global locations and controls. *Remote Sensing of Environment*, 134:294–304, 2013.

[614] Hung Chak Ho, Anders Knudby, Paul Sirovyak, Yongming Xu, Matus Hodul, and Sarah B Henderson. Mapping maximum urban air temperature on hot summer days. *Remote Sensing of Environment*, 154:38–45, 2014.

[615] Jackson Voelkel, Vivek Shandas, and Brendon Haggerty. Peer reviewed: Developing high-resolution descriptions of urban heat islands: A public health imperative. *Preventing chronic disease*, 13, 2016.

[616] Sidrah Hafeez, Man Sing Wong, Hung Chak Ho, Majid Nazeer, Janet Nichol, Sawaid Abbas, Danling Tang, Kwon Ho Lee, and Lilian Pun. Comparison of machine learning algorithms for retrieval of water quality indicators in case-II waters: a case study of Hong Kong. *Remote Sensing*, 11(6):617, 2019.

[617] Nikhil Kumar Koditala and Purnendu Shekar Pandey. Water quality monitoring system using IoT and machine learning. In *2018 International Conference on Research in Intelligent and Computing in Engineering (RICE)*, pages 1–5. IEEE, 2018.

[618] Qian Di, Petros Koutrakis, Christine Choirat, Francesca Dominici, and Joel D Schwartz. Machine learning approach for spatially and temporally resolved PM2.5 exposures in the continental United States. In *ISEE Conference Abstracts*, 2018.

[619] Jie Chen, Kees de Hoogh, Maciek Strak, Jules Kerckhoffs, Roel Vermeulen, Bert Brunekreef, and Gerard Hoek. OP III–4 exposure assessment models for NO2 and PM2.5 in the elapse study: a comparison of supervised linear regression and machine learning approaches, 2018.

[620] Nick Watts, W Neil Adger, Sonja Ayeb-Karlsson, Yuqi Bai, Peter Byass, Diarmid Campbell-Lendrum, Tim Colbourn, Peter Cox, Michael Davies, Michael Depledge, et al. The Lancet Countdown: tracking progress on health and climate change. *The Lancet*, 389(10074):1151–1164, 2017.

[621] Alex Pentland, David Lazer, Devon Brewer, and Tracy Heibeck. Using reality mining to improve public health and medicine. *Stud Health Technol Inform*, 149:93–102, 2009.

[622] Sherri Rose. Mortality risk score prediction in an elderly population using machine learning. *American journal of epidemiology*, 177(5):443–452, 2013.

[623] Patrick Meier. Human computation for disaster response. In *Handbook of human computation*, pages 95–104. Springer, 2013.

[624] Carlos Castillo. *Big crisis data: Social media in disasters and time-critical situations*. Cambridge University Press, 2016.

[625] William A Yasnoff, Patrick W O Carroll, Denise Koo, Robert W Linkins, and Edwin M Kilbourne. Public health informatics: improving and transforming public health in the information age. *Journal of Public Health Management and Practice*, 6(6):67–75, 2000.

[626] Fahad Pervaiz, Mansoor Pervaiz, Nabeel Abdur Rehman, and Umar Saif. FluBreaks: early epidemic detection from Google flu trends. *Journal of medical Internet research*, 14(5):e125, 2012.

[627] Vasileios Lampos, Tijl De Bie, and Nello Cristianini. Flu detector-tracking epidemics on Twitter. In *Joint European conference on machine learning and knowledge discovery in databases*, pages 599–602. Springer, 2010.

[628] Michael A Johansson, Nicholas G Reich, Aditi Hota, John S Brownstein, and Mauricio Santillana. Evaluating the performance of infectious disease forecasts: A comparison of climate-driven and seasonal dengue forecasts for Mexico. *Scientific reports*, 6:33707, 2016.

[629] David Lazer, Ryan Kennedy, Gary King, and Alessandro Vespignani. The parable of Google Flu: traps in big data analysis. *Science*, 343(6176):1203–1205, 2014.

[630] Sudhakar V Nuti, Brian Wayda, Isuru Ranasinghe, Sisi Wang, Rachel P Dreyer, Serene I Chen, and Karthik Murugiah. The use of google trends in health care research: a systematic review. *PloS one*, 9(10):e109583, 2014.

[631] Charles C Onu, Innocent Udeogu, Eyenimi Ndiomu, Urbain Kengni, Doina Precup, Guilherme M Sant'Anna, Edward Alikor, and Peace Opara. Ubenwa: Cry-based diagnosis of birth asphyxia. *Preprint arXiv:1711.06405*, 2017.

[632] John A Quinn, Alfred Andama, Ian Munabi, and Fred N Kiwanuka. Automated blood smear analysis for mobile malaria diagnosis. *Mobile Point-of-Care Monitors and Diagnostic Device Design*, 31:115, 2014.

[633] Joel Robertson and Del J DeHart. An agile and accessible adaptation of Bayesian inference to medical diagnostics for rural health extension workers. In *2010 AAAI Spring Symposium Series*, 2010.

[634] Emma Brunskill and Neal Lesh. Routing for rural health: optimizing community health worker visit schedules. In *2010 AAAI Spring Symposium Series*, 2010.

[635] Jigar Doshi, Saikat Basu, and Guan Pang. From satellite imagery to disaster insights. *Preprint arXiv:1812.07033*, 2018.

[636] Favyen Bastani, Songtao He, Sofiane Abbar, Mohammad Alizadeh, Hari Balakrishnan, Sanjay Chawla, and Sam Madden. Machine-assisted map editing. In *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 23–32. ACM, 2018.

[637] Stefan Voigt, Thomas Kemper, Torsten Riedlinger, Ralph Kiefl, Klaas Scholte, and Harald Mehl. Satellite image analysis for disaster and crisis-management support. *IEEE transactions on geoscience and remote sensing*, 45(6):1520–1528, 2007.

[638] Ritwik Gupta, Bryce Goodman, Nirav Patel, Ricky Hosfelt, Sandra Sajeev, Eric Heim, Jigar Doshi, Keane Lucas, Howie Choset, and Matthew Gaston. Creating xBD: A dataset for assessing building damage from satellite imagery. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 10–17, 2019.

[639] Alexandra Olteanu, Carlos Castillo, Fernando Diaz, and Sarah Vieweg. CrisisLex: A lexicon for collecting and filtering microblogged communications in crises. In *Eighth International AAAI Conference on Weblogs and Social Media*, 2014.

[640] Muhammad Imran, Carlos Castillo, Fernando Diaz, and Sarah Vieweg. Processing social media messages in mass emergency: A survey. *ACM Computing Surveys (CSUR)*, 47(4):67, 2015.

[641] David W Keith. Geoengineering the climate: History and prospect. *Annual review of energy and the environment*, 25(1):245–284, 2000.

[642] John G Shepherd. *Geoengineering the climate: science, governance and uncertainty*. Royal Society, 2009.

[643] Peter J Irvine, Ben Kravitz, Mark G Lawrence, and Helene Muri. An overview of the Earth system science of solar geoengineering. *Wiley Interdisciplinary Reviews: Climate Change*, 7(6):815–833, 2016.

[644] David Keith and Peter Irvine. The science and technology of solar geoengineering: A compact summary. *Governance of the Deployment of Solar Geoengineering*, page 1, 2018.

[645] Andy Parker and Peter J Irvine. The risk of termination shock from solar geoengineering. *Earth's Future*, 6(3):456–467, 2018.

[646] Peter Irvine, Kerry Emanuel, Jie He, Larry W Horowitz, Gabriel Vecchi, and David Keith. Halving warming with idealized solar geoengineering moderates key climate hazards. *Nature Climate Change*, page 1, 2019.

[647] Andy Jones, Jim Haywood, and Olivier Boucher. Climate impacts of geoengineering marine stratocumulus clouds. *Journal of Geophysical Research: Atmospheres*, 114(D10), 2009.

[648] Trude Storelvmo, WR Boos, and N Herger. Cirrus cloud seeding: a climate engineering mechanism with reduced side effects? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2031):20140116, 2014.

[649] Philip J Rasch, Simone Tilmes, Richard P Turco, Alan Robock, Luke Oman, Chih-Chieh Chen, Georgiy L Stenchikov, and Rolando R Garcia. An overview of geoengineering of climate using stratospheric sulphate aerosols. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1882):4007–4037, 2008.

[650] Hashem Akbari, H Damon Matthews, and Donny Seto. The long-term effect of increasing the albedo of urban areas. *Environmental Research Letters*, 7(2):024004, 2012.

[651] Roger Angel. Feasibility of cooling the Earth with a cloud of small spacecraft near the inner Lagrange point (L1). *Proceedings of the National Academy of Sciences*, 103(46):17184–17189, 2006.

[652] Justin McClellan, David W Keith, and Jay Apt. Cost analysis of stratospheric albedo modification delivery systems. *Environmental Research Letters*, 7(3):034019, 2012.

[653] Jordan P Smith, John A Dykema, and David W Keith. Production of sulfates onboard an aircraft: implications for the cost and feasibility of stratospheric solar geoengineering. *Earth and Space Science*, 5(4):150–162, 2018.

[654] Alan Robock, Douglas G MacMartin, Riley Duren, and Matthew W Christensen. Studying geoengineering with natural and anthropogenic analogs. *Climatic Change*, 121(3):445–458, 2013.

[655] Sebastian D Eastham, Debra K Weisenstein, David W Keith, and Steven RH Barrett. Quantifying the impact of sulfate geoengineering on mortality from air quality and UV-B exposure. *Atmospheric environment*, 187:424–434, 2018.

[656] Jonathan Proctor, Solomon Hsiang, Jennifer Burney, Marshall Burke, and Wolfram Schlenker. Estimating global agricultural effects of geoengineering using volcanic eruptions. *Nature*, 560(7719):480, 2018.

[657] Simon Gruber, Ulrich Blahak, Florian Haenel, Christoph Kottmeier, Thomas Leisner, Harel Muskatel, Trude Storelvmo, and Bernhard Vogel. A process study on thinning of arctic winter cirrus clouds with high-resolution icon-art simulations. *Journal of Geophysical Research: Atmospheres*, 0(ja), 2019.

[658] JA Dykema, DW Keith, and FN Keutsch. Improved aerosol radiative properties as a foundation for solar geoengineering risk assessment. *Geophysical Research Letters*, 43(14):7758–7766, 2016.

[659] Christopher G Fletcher, Ben Kravitz, and Bakr Badawy. Quantifying uncertainty from aerosol and atmospheric parameters and their impact on climate sensitivity. *Atmospheric Chemistry and Physics*, 18(23):17529–17543, 2018.

[660] Douglas G MacMartin and Ben Kravitz. The engineering of climate engineering. *Annual Review of Control, Robotics, and Autonomous Systems*, (0), 2018.

[661] Signe Moe, Anne Marthine Rustad, and Kristian G Hanssen. Machine learning in control systems: An overview of the state of the art. In *International Conference on Innovative Techniques and Applications of Artificial Intelligence*, pages 250–265. Springer, 2018.

[662] Ross Boczar, Nikolai Matni, and Benjamin Recht. Finite-data performance guarantees for the output-feedback control of an unknown system. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 2994–2999. IEEE, 2018.

[663] Brandon Amos, Ivan Jimenez, Jacob Sacks, Byron Boots, and J Zico Kolter. Differentiable MPC for end-to-end planning and control. In *Advances in Neural Information Processing Systems*, pages 8289–8300, 2018.

[664] Christian Schroeder de Witt and Thomas Hornigold. Stratospheric aerosol injection as a deep reinforcement learning problem. *Preprint arXiv:1905.07366*, 2019.

[665] Qian Di, Itai Kloog, Petros Koutrakis, Alexei Lyapustin, Yujie Wang, and Joel Schwartz. Assessing PM2.5 exposures with high spatiotemporal resolution across the continental United States. *Environmental science & technology*, 50(9):4712–4721, 2016.

[666] A Crane-Droesch, B Kravitz, and JT Abatzoglou. Using deep learning to model potential impacts of geoengineering via solar radiation management on US agriculture. In *AGU Fall Meeting Abstracts*, 2018.

[667] Marshall Burke, Solomon M Hsiang, and Edward Miguel. Global non-linear effect of temperature on economic production. *Nature*, 527(7577):235, 2015.

[668] Noah S Diffenbaugh and Marshall Burke. Global warming has increased global economic inequality. *Proceedings of the National Academy of Sciences*, 116(20):9808–9813, 2019.

[669] David L Kelly and Charles D Kolstad. Integrated assessment models for climate change control. *International yearbook of environmental and resource economics*, 2000:171–197, 1999.

[670] John Weyant. Some contributions of integrated assessment models of global climate change. *Review of Environmental Economics and Policy*, 11(1):115–137, 2017.

[671] Albert C Lin. Does geoengineering present a moral hazard. *Ecology LQ*, 40:673, 2013.

[672] Christopher J Preston. Ethics and geoengineering: reviewing the moral issues raised by solar radiation management and carbon dioxide removal. *Wiley Interdisciplinary Reviews: Climate Change*, 4(1):23–37, 2013.

[673] David W Keith. Toward a responsible solar geoengineering research program. *Issues in Science and Technology*, 33(3):71–77, 2017.

[674] Douglas G MacMartin, Ben Kravitz, and Philip J Rasch. On solar geoengineering and climate uncertainty. *Geophysical Research Letters*, 42(17):7156–7161, 2015.

[675] K. Williamson, A. Satre-Meloy, K. Velasco, and K. Green. Climate Change Needs Behavior Change: Making the case for behavioral solutions to reduce global warming. Technical report, Center for Behavior and the Environment, 2018.

[676] Alberto Mucci. The supermarket of the future knows exactly what you're eating. `https://www.vice.com/en_us/article/4xbppn/the-supermarket-of-the-future-knows-exactly-what-youre-eating`, 2016.

[677] Karen Ehrhardt-Martinez, Kat A Donnelly, Skip Laitner, et al. Advanced metering initiatives and residential feedback programs: a meta-review for household electricity-saving opportunities. American Council for an Energy-Efficient Economy Washington, DC, 2010.

[678] Adrian Albert and Mehdi Maasoumy. Predictive segmentation of energy consumers. *Applied energy*, 177:435–448, 2016.

[679] Hunt Allcott. Social norms and energy conservation. *Journal of public Economics*, 95(9-10):1082–1095, 2011.

[680] Hunt Allcott and Todd Rogers. The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review*, 104(10):3003–37, 2014.

[681] Christopher M Jones and Daniel M Kammen. Quantifying carbon footprint reduction opportunities for US households and communities. *Environmental science & technology*, 45(9):4088–4095, 2011.

[682] Christopher Jones and Daniel M Kammen. Spatial distribution of US household carbon footprints reveals suburbanization undermines greenhouse gas benefits of urban population density. *Environmental science & technology*, 48(2):895–902, 2014.

[683] K Carrie Armel, Abhay Gupta, Gireesh Shrimali, and Adrian Albert. Is disaggregation the holy grail of energy efficiency? the case of electricity. *Energy Policy*, 52:213–234, 2013.

[684] Vasughi Sundramoorthy, Grahame Cooper, Nigel Linge, and Qi Liu. Domesticating energy-monitoring systems: Challenges and design concerns. *IEEE pervasive Computing*, 10(1):20–27, 2011.

[685] David MacKay. *Sustainable Energy-without the hot air*. UIT Cambridge, 2008.

[686] David Klenert, Linus Mattauch, Emmanuel Combet, Ottmar Edenhofer, Cameron Hepburn, Ryan Rafaty, and Nicholas Stern. Making carbon pricing work for citizens. *Nature Climate Change*, 8(8):669–677, 2018.

[687] Olivier Corradi. Estimating the marginal carbon intensity of electricity with machine learning. `https://medium.com/electricitymap/using-machine-learning-to-estimate-the-hourly-marginal-carbon-intensity-of-electricity-49eade43b421`, 2018.

[688] Goran Strbac. Demand side management: Benefits and challenges. *Energy policy*, 36(12):4419–4426, 2008.

[689] FC Schweppe, B Daryanian, and RD Tabors. Algorithms for a spot price responding residential load controller. *IEEE Transactions on Power Systems*, 4(2):507–516, 1989.

[690] Elena Mocanu, Decebal Constantin Mocanu, Phuong H Nguyen, Antonio Liotta, Michael E Webber, Madeleine Gibescu, and Johannes G Slootweg. On-line building energy optimization using deep reinforcement learning. *IEEE Transactions on Smart Grid*, 2018.

[691] T Remani, EA Jasmin, and TP Imthias Ahamed. Residential load scheduling with renewable generation in the smart grid: A reinforcement learning approach. *IEEE Systems Journal*, (99):1–12, 2018.

[692] Liam F Beiser-McGrath and Robert A Huber. Assessing the relative importance of psychological and demographic factors for predicting climate and environmental attitudes. *Climatic change*, 149(3-4):335–347, 2018.

[693] Simone Carr-Cornish, Peta Ashworth, John Gardner, and Stephen J Fraser. Exploring the orientations which characterise the likely public acceptance of low emission energy technologies. *Climatic change*, 107(3-4):549–565, 2011.

[694] Cristóbal De La Maza, Alex Davis, Cleotilde Gonzalez, and Inês Azevedo. A graph-based model to discover preference structure from choice data. In *40th Annual Meeting of the Cognitive Science Society (CogSci 2018)*, pages 25–28, 2018.

[695] Elizabeth Gabe-Thomas, Ian Walker, Bas Verplanken, and Gavin Shaddick. Householders' mental models of domestic energy consumption: using a sort-and-cluster method to identify shared concepts of appliance similarity. *PloS one*, 11(7):e0158949, 2016.

[696] Shan-lin Yang, Chao Shen, et al. A review of electric load classification in smart grid environment. *Renewable and Sustainable Energy Reviews*, 24:103–110, 2013.

[697] Cristóbal de la Maza Guzmán. Willingness to pay to avoid environmental impacts of electricity generation. Technical report, Latin American and Caribbean Environmental Economics Program, 2013.

[698] Jiansong Zhang and Nora M. El-Gohary. Automated information transformation for automated regulatory compliance checking in construction. *Journal of Computing in Civil Engineering*, 29(4):B4015001, 2015.

[699] Wanda Bell, Lewis Ahron Kaufman, William Joseph Krajewski, John J McGillicuddy, Paul Aloysius Scanlon Jr, Abhijit Dey, Sharon Ameet Fanse, Giridhar Holenarsipur Nagaraj, Shyamli Rai, Sunitha Sundaramurthy, et al. Systems and methods for automated data privacy compliance, November 29 2016. US Patent 9,507,960.

[700] Charlotte Jones, Donald W Hine, and Anthony DG Marks. The future is now: reducing psychological distance to increase public engagement with climate change. *Risk Analysis*, 37(2):331–341, 2017.

[701] Liam F Beiser-McGrath and Thomas Bernauer. Commitment failures are unlikely to undermine public support for the paris agreement. *Nature climate change*, 9(3):248, 2019.

[702] Victor Schmidt, Alexandra Luccioni, S Karthik Mukkavilli, Narmada Balasooriya, Kris Sankaran, Jennifer Chayes, and Yoshua Bengio. Visualizing the consequences of climate change using cycle-consistent adversarial networks. *Preprint arXiv:1905.03709*, 2019.

[703] Ioannis C Konstantakopoulos, Andrew R Barkan, Shiying He, Tanya Veeravalli, Huihan Liu, and Costas Spanos. A deep learning and gamification approach to improving human-building interaction and energy efficiency in smart infrastructure. *Applied Energy*, 237:810–821, 2019.

[704] Marc Gunther. The power of peer pressure in combatting climate change. `https://www.greenbiz.com/blog/2010/01/19/power-peer-pressure-combatting-climate-change`, 2010.

[705] Dominic Dudley. Renewable energy will be consistently cheaper than fossil fuels by 2020, report claims. `https://www.forbes.com/sites/dominicdudley/2018/01/13/renewable-energy-cost-effective-fossil-fuels-2020/#63a450834ff2`, 2018.

[706] Scott De Marchi and Scott E Page. Agent-based models. *Annual Review of political science*, 17:1–20, 2014.

[707] Joshua M. Epstein. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton University Press, stu - student edition edition, 2006.

[708] Varun Rai and Scott A. Robinson. Agent-based modeling of energy technology adoption: Empirical integration of social, behavioral, economic, and environmental factors. *Environmental Modelling & Software*, 70:163 – 177, 2015.

[709] Leonore Haelg, Marius Waelchli, and Tobias S Schmidt. Supporting energy technology deployment while avoiding unintended technological lock-in: a policy design perspective. *Environmental Research Letters*, 13(10):104011, oct 2018.

[710] Tao Zhang and William J Nuttall. An agent-based simulation of smart metering technology adoption. *International Journal of Agent Technologies and Systems (IJATS)*, 4(1):17–38, 2012.

[711] Mehdi Noori and Omer Tatari. Development of an agent-based model for regional market penetration projections of electric vehicles in the united states. *Energy*, 96:215 – 230, 2016.

[712] Haifeng Zhang, Yevgeniy Vorobeychik, Joshua Letchford, and Kiran Lakkaraju. Data-driven agent-based modeling, with application to rooftop solar adoption. *Autonomous Agents and Multi-Agent Systems*, 30(6):1023–1049, Nov 2016.

[713] Varun Rai, D Cale Reeves, and Robert Margolis. Overcoming barriers and uncertainties in the adoption of residential solar pv. *Renewable Energy*, 89:498–505, 2016.

[714] Bryan Bollinger and Kenneth Gillingham. Peer effects in the diffusion of solar photovoltaic panels. *Marketing Science*, 31(6):900–912, 2012.

[715] Maja Schlüter, Alessandro Tavoni, and Simon Levin. Robustness of norm-driven cooperation in the commons. *Proceedings of the Royal Society B: Biological Sciences*, 283(1822):20152431, 2016.

[716] Sylvie Geisendorf. Evolutionary climate-change modelling: A multi-agent climate-economic model. *Computational Economics*, 52(3):921–951, Oct 2018.

[717] Jule Thober, Nina Schwarz, and Kathleen Hermans. Agent-based modeling of environment-migration linkages: a review. *Ecology and society*, 23(2), 2018.

[718] Haifeng Zhang and Yevgeniy Vorobeychik. Empirically grounded agent-based models of innovation diffusion: a critical review. *Artificial Intelligence Review*, 52(1):707–741, Jun 2019.

[719] Chathika Gunaratne, Ivan Garibay, and Nguyen Dang. Evolutionary model discovery of causal factors behind the socio-agricultural behavior of the ancestral pueblo, 2018.

[720] Christian Hilbe, Štěpán Šimsa, Krishnendu Chatterjee, and Martin A. Nowak. Evolution of cooperation in stochastic games. *Nature*, 559(7713):246–249, 2018.

[721] Liviu Panait and Sean Luke. Cooperative multi-agent learning: The state of the art. *Autonomous Agents and Multi-Agent Systems*, 11(3):387–434, November 2005.

[722] Hyun-Rok Lee and Taesik Lee. Improved cooperative multi-agent reinforcement learning algorithm augmented by mixing demonstrations from centralized policy. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '19, pages 1089–1098, Richland, SC, 2019. International Foundation for Autonomous Agents and Multiagent Systems.

[723] Natasha Jaques, Angeliki Lazaridou, Edward Hughes, Çaglar Gülçehre, Pedro A. Ortega, DJ Strouse, Joel Z. Leibo, and Nando de Freitas. Intrinsic social motivation via causal influence in multi-agent RL. *CoRR*, abs/1810.08647, 2018.

[724] David Martimort and Wilfried Sand-Zantman. A mechanism design approach to climate agreements. 2011.

[725] Ariel D. Procaccia. Cake cutting: Not just child's play. *Commun. ACM*, 56(7):78–87, July 2013.

[726] Thomas Sterner, Edward B. Barbier, Ian Bateman, Inge van den Bijgaart, Anne-Sophie Crépin, Ottmar Edenhofer, Carolyn Fischer, Wolfgang Habla, John Hassler, Olof Johansson-Stenman, Andreas Lange, Stephen Polasky, Johan Rockström, Henrik G. Smith, Will Steffen, Gernot Wagner, James E. Wilen, Francisco Alpízar, Christian Azar, Donna Carless, Carlos Chávez, Jessica Coria, Gustav Engström, Sverker C. Jagers, Gunnar Köhlin, Åsa Löfgren, Håkan Pleijel, and Amanda Robinson. Policy design for the anthropocene. *Nature Sustainability*, 2(1):14–21, 2019.

[727] M. Granger Morgan. *Theory and Practice in Policy Analysis: Including Applications in Science and Technology*. Cambridge University Press, 2017.

[728] D. S. Patton, C. V. annd Sawicki and J. Clark. *Basic methods of policy analysis and planning*. 2015.

[729] Giuseppe A Veltri and Dimitrinka Atanasova. Climate change on twitter: Content, media ecology and information sharing behaviour. *Public Understanding of Science*, 26(6):721–737, 2017.

[730] Hywel TP Williams, James R McMurray, Tim Kurz, and F Hugo Lambert. Network analysis reveals open forums and echo chambers in social media discussions of climate change. *Global environmental change*, 32:126–138, 2015.

[731] Andrei P Kirilenko and Svetlana O Stepchenkova. Public microblogging on climate change: One year of twitter worldwide. *Global environmental change*, 26:171–182, 2014.

[732] John Weyant. Some Contributions of Integrated Assessment Models of Global Climate Change. *Review of Environmental Economics and Policy*, 11(1):115–137, 03 2017.

[733] Richard H. Moss, Jae A. Edmonds, Kathy A. Hibbard, Martin R. Manning, Steven K. Rose, Detlef P. van Vuuren, Timothy R. Carter, Seita Emori, Mikiko Kainuma, Tom Kram, Gerald A. Meehl, John F. B. Mitchell, Nebojsa Nakicenovic, Keywan Riahi, Steven J. Smith, Ronald J. Stouffer, Allison M. Thomson, John P. Weyant, and Thomas J. Wilbanks. The next generation of scenarios for climate change research and assessment. *Nature*, 463(7282):747–756, 2010.

[734] Jan Philipp Dietrich, Alexander Popp, and Hermann Lotze-Campen. Reducing the loss of information and gaining accuracy with clustering methods in a global land-use model. *Ecological modelling*, 263:233–243, 2013.

[735] Christian Folberth, Artem Baklanov, Juraj Balkovič, Rastislav Skalskỳ, Nikolay Khabarov, and Michael Obersteiner. Spatio-temporal downscaling of gridded crop model yield estimates based on machine learning. *Agricultural and forest meteorology*, 264:1–15, 2019.

[736] Lianfa Li. Geographically weighted machine learning and downscaling for high-resolution spatiotemporal estimations of wind speed. *Remote Sensing*, 11(11):1378, 2019.

[737] Wan Li, Li Ni, Zhao-Liang Li, Si-Bo Duan, and Hua Wu. Evaluation of machine learning algorithms in spatial downscaling of modis land surface temperature. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12:2299–2307, 2019.

[738] Marc Jaxa-Rozen and Jan Kwakkel. Tree-based ensemble methods for sensitivity analysis of environmental models: A performance comparison with sobol and morris techniques. *Environmental Modelling & Software*, 107:245 – 266, 2018.

[739] Simon Scheidegger and Ilias Bilionis. Machine learning for high-dimensional dynamic stochastic economies. *Journal of Computational Science*, 33:68–82, 2019.

[740] Victor Duarte. Machine learning for continuous-time economics. 2018.

[741] Shunsuke Mori, Toyoaki Washida, Atsushi Kurosawa, and Toshihiko Masui. Assessment of mitigation strategies as tools for risk management under future uncertainties: a multi-model approach. *Sustainability Science*, 13(2):329–349, Mar 2018.

[742] S.D. Pohekar and M. Ramachandran. Application of multi-criteria decision making to sustainable energy planning—a review. *Renewable and Sustainable Energy Reviews*, 8(4):365 – 381, 2004.

[743] Alessandro Mattiussi, Michele Rosano, and Patrizia Simeoni. A decision support system for sustainable energy supply combining multi-objective and multi-attribute analysis: An Australian case study. *Decision Support Systems*, 57:150 – 159, 2014.

[744] Multi-objective optimization for sustainable development of the power sector: An economic, environmental, and social analysis of Iran. *Energy*, 161:493 – 507, 2018.

[745] Qinru Shi, Jonathan M. Gomes-Selman, Roosevelt García-Villacorta, Suresh Sethi, Alexander S. Flecker, and Carla P. Gomes. Efficiently optimizing for dendritic connectivity on tree-structured networks in a multi-objective framework. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, COMPASS '18, pages 26:1–26:8, New York, NY, USA, 2018. ACM.

[746] Jee-Hoon Han, Yu-Chan Ahn, and In-Beum Lee. A multi-objective optimization model for sustainable electricity generation and CO2 mitigation (EGCM) infrastructure design considering economic profit and financial risk. *Applied Energy*, 95:186 – 195, 2012.

[747] H Hassine, Maher Barkallah, and A Bellacicco. Multi objective optimization for sustainable manufacturing, application in turning. *International Journal of Simulation Modelling*, 14:98–109, 03 2015.

[748] A. Chaabane, A. Ramudhin, and M. Paquet. Design of sustainable supply chains under the emission trading scheme. *International Journal of Production Economics*, 135(1):37 – 49, 2012. Advances in Optimization and Design of Supply Chains.

[749] Milena Lakicevic, Zorica Srdjevic, Bojan Srdjevic, and Miodrag Zlatic. Decision making in urban forestry by using approval voting and multicriteria approval method (case study: Zvezdarska forest, Belgrade, Serbia). *Urban Forestry & Urban Greening*, 13(1):114 – 120, 2014.

[750] Vivek K. Varma, Ian Ferguson, and Ian Wild. Decision support system for the sustainable forest management. *Forest Ecology and Management*, 128(1):49 – 55, 2000.

[751] Serna-González M. Ponce-Ortega J.M. et al. Gutiérrez-Arriaga, C.G. Multi-objective optimization of steam power plants for sustainable generation of electricity. *Clean Techn Environ Policy*, 15(551), 2013.

[752] Riccardo Minciardi, Massimo Paolucci, Michela Robba, and Roberto Sacile. Multi-objective optimization of solid waste flows: Environmentally sustainable strategies for municipalities. *Waste Management*, 28(11):2202 – 2212, 2008.

[753] Ching-Ho Chen, Wei-Lin Liu, and Chia-Hsing Chen. Development of a multiple objective planning theory and system for sustainable air quality monitoring networks. *Science of The Total Environment*, 354(1):1 – 19, 2006.

[754] Dalia Streimikiene and Tomas Balezentis. Multi-objective ranking of climate change mitigation policies and measures in Lithuania. *Renewable and Sustainable Energy Reviews*, 18:144 – 153, 2013.

[755] Taimoor Akhtar and Christine A Shoemaker. Efficient multi-objective optimization through population-based parallel surrogate search. *arXiv preprint arXiv:1903.02167*, 2019.

[756] Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P Adams, and Nando De Freitas. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE*, 104(1):148–175, 2015.

[757] Jasper Snoek, Hugo Larochelle, and Ryan P. Adams. Practical bayesian optimization of machine learning algorithms. In *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 2*, NIPS'12, pages 2951–2959, USA, 2012. Curran Associates Inc.

[758] David Silver, Aja Huang, Christopher J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the game of go with deep neural networks and tree search. *Nature*, 529:484–503, 2016.

[759] Justin Grimmer and Brandon M. Stewart. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3):267–297, 2013.

[760] Judea Pearl. The seven tools of causal inference, with reflections on machine learning. *Commun. ACM*, 62(3):54–60, February 2019.

[761] Susan Athey and Guido W Imbens. Machine learning methods that economists should know about. *Annual Review of Economics*, 11, 2019.

[762] Miguel A. Hernán, John Hsu, and Brian Healy. A second chance to get causal inference right: A classification of data science tasks. *CHANCE*, 32(1):42–49, 2019.

[763] Noemi Kreif and Karla DiazOrdaz. Machine learning in policy evaluation: new tools for causal inference, 2019.

[764] Susan Athey. Beyond prediction: Using big data for policy problems. *Science*, 355(6324):483–485, 2017.

[765] Isabel Hovdahl. On the use of machine learning for causal inference in climate economics. 2019.

[766] Jianing Zhao, Daniel M. Runfola, and Peter Kemper. Quantifying heterogeneous causal treatment effects in world bank development finance projects. In Yasemin Altun, Kamalika Das, Taneli Mielikäinen, Donato Malerba, Jerzy Stefanowski, Jesse Read, Marinka Žitnik, Michelangelo Ceci, and Sašo Džeroski, editors, *Machine Learning and Knowledge Discovery in Databases*, pages 204–215, Cham, 2017. Springer International Publishing.

[767] Joseph E Stiglitz, Nicholas Stern, Maosheng Duan, Ottmar Edenhofer, Gaël Giraud, Geoffrey M Heal, Emilio Lèbre la Rovere, Adele Morris, Elisabeth Moyer, Mari Pangestu, et al. Report of the high-level commission on carbon prices. 2017.

[768] Nicholas Stern. The economics of climate change. *American Economic Review*, 98(2):1–37, 2008.

[769] A Denny Ellerman, Frank J Convery, and Christian De Perthuis. *Pricing carbon: the European Union emissions trading scheme*. Cambridge University Press, 2010.

[770] Hamed Ghoddusi, Germán G Creamer, and Nima Rafizadeh. Machine learning in energy economics and finance: A review. *Energy Economics*, 81:709–727, 2019.

[771] Bangzhu Zhu, Dong Han, Ping Wang, Zhanchi Wu, Tao Zhang, and Yi-Ming Wei. Forecasting carbon price using empirical mode decomposition and evolutionary least squares support vector regression. *Applied energy*, 191:521–530, 2017.

[772] Wei Sun and Chongchong Zhang. Analysis and forecasting of the carbon price using multi—resolution singular value decomposition and extreme learning machine optimized by adaptive whale optimization algorithm. *Applied energy*, 231:1354–1371, 2018.

[773] Bangzhu Zhu, Shunxin Ye, Ping Wang, Kaijian He, Tao Zhang, and Yi-Ming Wei. A novel multiscale nonlinear ensemble leaning paradigm for carbon price forecasting. *Energy Economics*, 70:143–157, 2018.

[774] Sun Wei, Zhang Chongchong, and Sun Cuiping. Carbon pricing prediction based on wavelet transform and K-ELM optimized by bat optimization algorithm in China ETS: the case of Shanghai and Hubei carbon markets. *Carbon Management*, 9(6):605–617, 2018.

[775] Bangzhu Zhu, Ping Wang, Julien Chevallier, and Yiming Wei. Carbon price analysis using empirical mode decomposition. *Computational Economics*, 45(2):195–206, 2015.

[776] Michael Rothschild. A two-armed bandit theory of market pricing. *Journal of Economic Theory*, 9(2):185 – 202, 1974.

[777] Rajkumar Ragupathi and Tapas Das. A stochastic game approach for modeling wholesale energy bidding in deregulated power markets. *Power Systems, IEEE Transactions on*, 19:849 – 856, 06 2004.

[778] Vishnuteja Nanduri and Tapas Das. A reinforcement learning model to assess market power under auction-based energy pricing. *Power Systems, IEEE Transactions on*, 22:85 – 95, 03 2007.

[779] Guochang Fang, Lixin Tian, Min Fu, Mei Sun, Ruijin Du, and Menghe Liu. Investigating carbon tax pilot in YRD urban agglomerations – analysis of a novel ESER system with carbon tax constraints and its application. *Applied energy*, 194:635–647, 2017.

[780] Guochang Fang, Lixin Tian, Menghe Liu, Min Fu, and Mei Sun. How to optimize the development of carbon trading in China – enlightenment from evolution rules of the EU carbon price. *Applied energy*, 211:1039–1049, 2018.

[781] Prashant Nagapurkar and Joseph D Smith. Techno-economic optimization and social costs assessment of microgrid-conventional grid integration using genetic algorithm and artificial neural networks: A case study for two US cities. *Journal of Cleaner Production*, 229:552–569, 2019.

[782] Nicolò Barbieri. Fuel prices and the invention crowding out effect: Releasing the automotive industry from its dependence on fossil fuel. *Technological Forecasting and Social Change*, 111:222–234, 2016.

[783] Xiping Zheng, Qiang Guo, Zenglu Li, and Ting Zhang. Optimal choice of enterprise's production strategy under constraints of carbon quota. *International Journal of Computational Intelligence Systems*, 11(1):1268–1277, 2018.

[784] Qunli Wu and Hongjie Zhang. Research on optimization allocation scheme of initial carbon emission quota from the perspective of welfare effect. *Energies*, 12(11):2118, 2019.

[785] Ramon Granell, Colin J Axon, and David CH Wallom. Predicting winning and losing businesses when changing electricity tariffs. *Applied energy*, 133:298–307, 2014.

[786] Paulo Picchetti. Hedonic residential property price estimation using geospatial data: a machine-learning approach. *Instituto Brasileiro de Economia*, 04 2017.

[787] Byeonghwa Park and Jae Bae. Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data. *Expert Systems with Applications*, 42, 04 2015.

[788] Timothy Oladunni and Sharad Sharma. Hedonic housing theory – a machine learning investigation. 12 2016.

[789] Mark A Delucchi, James J Murphy, and Donald R McCubbin. The health and visibility cost of air pollution: a comparison of estimation methods. *Journal of Environmental Management*, 64(2):139 – 152, 2002.

[790] Charles D Kerchner and William S Keeton. California's regulatory forest carbon market: Viability for northeast landowners. *Forest Policy and Economics*, 50:70–81, 2015.

[791] UNESCO. *Not just hot air: putting climate change education into practice*. United Nations Educational, Scientific and Cultural Organization, 2015.

[792] Douglas H Fisher, Zimei Bian, and Selina Chen. Incorporating sustainability into computing education. *IEEE Intelligent Systems*, 31(5):93–96, 2016.

[793] Heather Randell and Clark Gray. Climate variability and educational attainment: Evidence from rural ethiopia. *Global environmental change*, 41:111–123, 2016.

[794] Heather Randell and Clark Gray. Climate change and educational attainment in the global tropics. *Proceedings of the National Academy of Sciences*, 116(18):8840–8845, 2019.

[795] Devendra Singh Chaplot, Christopher MacLellan, Ruslan Salakhutdinov, and Kenneth Koedinger. Learning cognitive models using neural networks. In *International Conference on Artificial Intelligence in Education*, pages 43–56. Springer, 2018.

[796] Benjamin Clement, Didier Roy, Pierre-Yves Oudeyer, and Manuel Lopes. Multi-armed bandits for intelligent tutoring systems. *Journal of Educational Data Mining*, 7(2):20–48, 2013.

[797] Xinya Du, Junru Shao, and Claire Cardie. Learning to ask: Neural question generation for reading comprehension. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, 2017.

[798] Ana Iglesias, Paloma Martínez, Ricardo Aler, and Fernando Fernández. Learning teaching strategies in an adaptive and intelligent educational system through reinforcement learning. *Applied Intelligence*, 31(1):89–106, 2009.

[799] Kenneth R Koedinger, Emma Brunskill, Ryan SJd Baker, Elizabeth A McLaughlin, and John Stamper. New potentials for data-driven intelligent tutoring system development and optimization. *AI Magazine*, 34(3):27–41, 2013.

[800] Ulrich Gnewuch, Stefan Morana, Carl Heckmann, and Alexander Maedche. Designing conversational agents for energy feedback. In *International Conference on Design Science Research in Information Systems and Technology*, pages 18–33. Springer, 2018.

[801] Cristóbal Romero, Sebastián Ventura, Pedro G Espejo, and César Hervás. Data mining algorithms to classify students. In *Educational data mining 2008*, 2008.

[802] Vanessa Svihla and Marcia C Linn. A design-based approach to fostering understanding of global climate change. *International Journal of Science Education*, 34(5):651–676, 2012.

[803] Roger Nkambou, Riichiro Mizoguchi, and Jacqueline Bourdeau. *Advances in intelligent tutoring systems*, volume 308. Springer Science & Business Media, 2010.

[804] Srećko Joksimović, Oleksandra Poquet, Vitomir Kovanović, Nia Dowell, Caitlin Mills, Dragan Gašević, Shane Dawson, Arthur C Graesser, and Christopher Brooks. How do we model learning at scale? A systematic review of research on MOOCs. *Review of Educational Research*, 88(1):43–86, 2018.

[805] Niels Pinkwart. Another 25 years of AIED? challenges and opportunities for intelligent educational technologies of the future. *International journal of artificial intelligence in education*, 26(2):771–783, 2016.

[806] Ido Roll, Daniel M Russell, and Dragan Gašević. Learning at scale. *International Journal of Artificial Intelligence in Education*, 28(4):471–477, 2018.

[807] Chris Dede. Immersive interfaces for engagement and learning. *Science*, 323(5910):66–69, 2009.

[808] Maya Cakmak and Andrea L Thomaz. Eliciting good teaching from humans for machine learners. *Artificial Intelligence*, 217:198–215, 2014.

[809] Benjamin D Nye. Intelligent tutoring systems by and for the developing world: A review of trends and approaches for educational technology in a global context. *International Journal of Artificial Intelligence in Education*, 25(2):177–203, 2015.

[810] Eugene C Cordero, Anne Marie Todd, and Diana Abellera. Climate change education and the ecological footprint. *Bulletin of the American Meteorological Society*, 89(6):865–872, 2008.

[811] Allison Anderson. Climate change education for mitigation and adaptation. *Journal of Education for Sustainable Development*, 6(2):191–206, 2012.

[812] Jeannette Angel, Alicia LaValle, Deepti Mathew Iype, Stephen Sheppard, and Aleksandra Dulic. Future delta 2.0 an experiential learning context for a serious game about local climate change. In *SIGGRAPH Asia 2015 Symposium on Education*, page 12. ACM, 2015.

[813] Simon Dietz, Alex Bowen, Charlie Dixon, and Philip Gradwell. 'climate value at risk' of global financial assets. *Nature Climate Change*, 6(7):676, 2016.

[814] Jean Boissinot, Doryane Huber, and Gildas Lame. Finance and climate: The transition to a low-carbon and climate-resilient economy from a financial sector perspective. *OECD Journal: Financial Market Trends*, 2016.

[815] Stefano Battiston, Antoine Mandel, Irene Monasterolo, Franziska Schütze, and Gabriele Visentin. A climate stress-test of the financial system. *Nature Climate Change*, 7(4):283, 2017.

[816] Emanuele Campiglio, Yannis Dafermos, Pierre Monnin, Josh Ryan-Collins, Guido Schotten, and Misa Tanaka. Climate change challenges for central banks and financial regulators. *Nature Climate Change*, 8(6):462, 2018.

[817] Luc Eyraud, Benedict Clements, and Abdoul Wane. Green investment: Trends and determinants. *Energy Policy*, 60:852–865, 2013.

[818] Ivan Diaz-Rainey, Becky Robertson, and Charlie Wilson. Stranded research? leading finance journals are silent on climate change. *Climatic Change*, 143(1-2):243–260, 2017.

[819] Gianfranco Gianfrate. Designing carbon-neutral investment portfolios. In *Designing a Sustainable Financial System*, pages 151–171. Springer, 2018.

[820] Ariel Bergmann, Nick Hanley, and Robert Wright. Valuing the attributes of renewable energy investments. *Energy policy*, 34(9):1004–1014, 2006.

[821] Elizabeth Stanny and Kirsten Ely. Corporate environmental disclosures about the effects of climate change. *Corporate Social Responsibility and Environmental Management*, 15(6):338–348, 2008.

[822] Robert F Engle, Stefano Giglio, Heebum Lee, Bryan T Kelly, and Johannes Stroebel. Hedging climate change news. *Available at SSRN 3317570*, 2019.

[823] Mats Andersson, Patrick Bolton, and Frédéric Samama. Hedging climate risk. *Financial Analysts Journal*, 72(3):13–32, 2016.

[824] William A Pizer. Choosing price or quantity controls for greenhouse gases. In Wallace E Oates, editor, *The RFF Reader in Environmental and Resource Policy*, pages 225–234. Resources for the Future, 2006.

[825] Bangzhu Zhu and Julien Chevallier. Carbon price forecasting with a hybrid arima and least squares support vector machines methodology. In *Pricing and Forecasting Carbon Markets*, pages 87–107. Springer, 2017.

[826] Jianguo Zhou, Xuechao Yu, and Xiaolei Yuan. Predicting the carbon price sequence in the shenzhen emissions exchange using a multiscale ensemble forecasting model based on ensemble empirical mode decomposition. *Energies*, 11(7):1907, 2018.