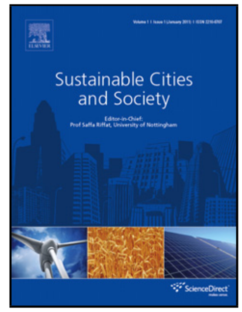
Journal Pre-proof



Machine learning for geographically differentiated climate change mitigation in urban areas

Nikola Milojevic-Dupont, Felix Creutzig

PII: S2210-6707(20)30742-3
DOI: https://doi.org/10.1016/j.scs.2020.102526
Reference: SCS 102526

To appear in: *Sustainable Cities and Society*

Received Date: 16 June 2020
Revised Date: 25 September 2020
Accepted Date: 26 September 2020

Please cite this article as: { doi: https://doi.org/

Machine learning for geographically differentiated climate change mitigation in urban areas

Authors: Nikola Milojevic-Dupont[1,2] and Felix Creutzig[1,2*]

Affiliations:

[1]Mercator Research Institute on Global Commons and Climate Change, Torgauer Straße 12-15, EUREF Campus #19, 10829 Berlin, Germany

[2]Technische Universität Berlin, Straße des 17. Juni 135, 10623 Berlin, Germany

*creutzig@mcc-berlin.net

Highlights:

- We review the ML literature used for climate change mitigation in cities
- We find a rapidly growing literature base and discern algorithmic approaches
- Existing researches focuses on increasing efficiency
- Machine learning can be effectively utilized for low-carbon urban planning

Abstract:

Artificial intelligence and machine learning are transforming scientific disciplines, but their full potential for climate change mitigation remains elusive. Here, we conduct a systematic review of applied machine learning studies that are of relevance for climate change mitigation, focusing on spatial data and specifically on the fields of remote sensing, urban transportation, and buildings. The relevant body of literature spans twenty years and is growing exponentially. We show that the emergence of big data and machine learning methods enables climate solution research to overcome generic recommendations and provide policy solutions at urban, street, building and household scale, adapted to specific contexts, but scalable to global mitigation potentials. We

suggest a meta-algorithmic architecture and framework for using machine learning to optimize urban planning for accelerating, improving and transforming urban infrastructure provision.

Keywords: Machine Learning; cities; climate change mitigation; Urban governance

Main Text:

Policy makers, mayors, administrations and individuals seek to advance climate solutions and demand tailored approaches that match their political, economic and climatic conditions (Hermwille et al., 2017; Reckien et al., 2018; Shan et al., 2018). A main source of reference are the assessment report of the Intergovernmental Panel on Climate Change (IPCC) that provides comprehensive overviews on technology assessments, sectoral approaches, integrated scenarios, and policy studies (IPCC, 2014, 2018). Modelers emphasize multiple socially optimal decarbonization pathways consistent both with global average temperature stabilization targets and stylized societal or environmental constraints. As a result, broad recommendations about technological change, fossil fuel phase-out, and national policies emerge and serve as a reference for governments.

Yet, detailed and contextualized policy options that reflect the idiosyncrasies of places and cultures are scarce and often insufficient. Policy makers are left disoriented on which measures are adequate and impactful, and how everyday decisions related to infrastructure investments or urban planning can be modified to adjust to a low-carbon future. In research, important disagreements remain when quantifying mitigation potential, especially for energy end-uses, demand and services, and for human settlements that are differentiated with across developments, geographies, and spatial structure (Creutzig, Fernandez, et al., 2016; Wilson et al., 2012). For example, integrated assessment models and bottom-up sectoral studies differ on the future emission reduction from energy efficiency measures in the transport (Creutzig, 2016) and buildings sectors (Lucon et al., 2014). Global assessments insufficiently reflect variations in local specificities e.g. infrastructural, economic, climatic, social or political contexts (Creutzig et al., 2019); and place-specific studies investigating the building, street, or urban scale remain idiosyncratic in context and differ in their boundaries of analysis, thus rendering their contribution to global climate change mitigation a distant goal, but also making comparison between local-scale studies difficult (Lamb et al., 2018). The IPCC's AR5 report knowledge gaps on urban climate action (IPCC, 2014): there is too little understanding of the magnitude of the emissions reductions from altering urban form, and emissions savings from integrated infrastructure and land use planning. New analyses are required both on understanding relationships (Silva et al., 2017) and simulating future pathways with and without interventions (Silva et al., 2018).

Therefore, researchers increasingly call for systematic, methodically well-grounded research to *upscale* place-specific climate solutions while respecting local variation and context (Creutzig & Kammen, 2009; Acuto et al., 2018; Creutzig et al., 2019). Two new developments have high potential to help address this desired transformation.

The first component of this transformation is the emergence of big data (Ford et al., 2016) at high spatial resolution and individual heterogeneity. Big data is often produced from sensors or user data; it includes satellite and aerial imagery, volunteered geographic information (Haklay, 2010), such as OpenStreetMap, geo-localized devices data, such as transaction data (Di Clemente et al., 2018) or social media data (Ilieva & McPhearson, 2018), and surveys or governmental data, such as cadastral data. These rich sources of information unfold great potential for analyses of improved granularity of climate solutions (Creutzig et al., 2019; Ilieva & McPhearson, 2018).

The second component is machine learning (ML). Breakthroughs in computer science theory, algorithmic research and computational power provide the instruments to extract meaningful information from massive data. ML methods, based on learning theory (Vapnik, 1999), permit to generalize pattern recognition on unseen data outside of the observed sample (Vapnik, 1999). Since the mid-2000s, deep multi-layer architectures – so-called deep learning (DL) methods (LeCun et al., 2015) – increased ML performance by learning high-level representations of the data. See Box 1 for a general introduction on machine learning.

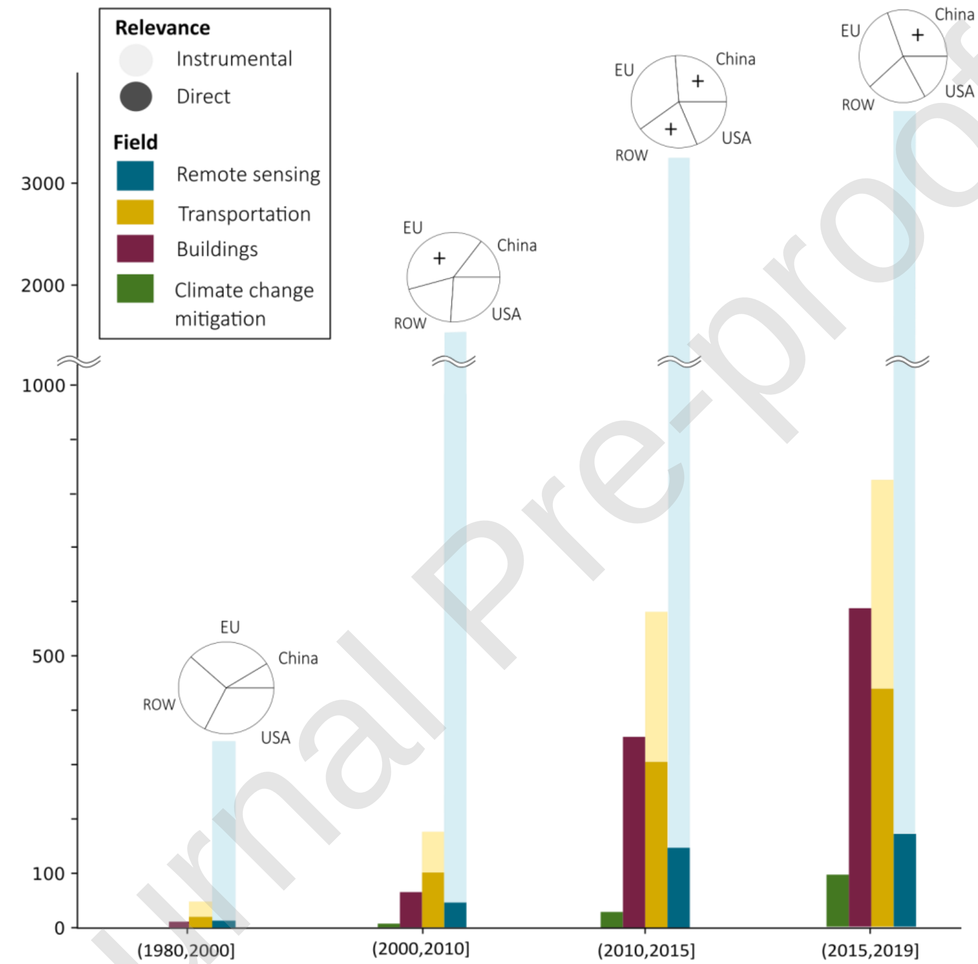
Both dimensions appear also in the concept of smart cities, an umbrella term for (digital) technologies that aim to make cities more efficient, and possibly more sustainable (B. N. Silva et al., 2018; Yigitcanlar et al., 2019) – even though researchers and politicians must remain careful not to conflate smartness with sustainability (Noy & Givoni, 2018). For the purpose of this article, we are mostly concerned with the application and potential of AI algorithms for climate change mitigation, a more specific topic than what is generally understood when referring to smart cities.

The use of ML in climate change mitigation research remains nascent. (Rolnick et al., 2019) provide a broad review on ML applications for tackling climate change, and find relevant applications spanning many domains. Most successful applications in climate change include Earth system analysis (Reichstein et al., 2019), such as modelling multi-scale atmospheric processes (Rasp et al., 2018), and modelling climate impacts at high resolution by making use of big data from satellites, weather stations, radars, and other sources to specify the consequences of hurricanes and deforestation on ecosystems, or of drought on crop yields (*Atlas AI*, n.d.; Fletcher et al., 2019; McDowell et al., 2015). However, ML is not yet a common tool in core climate change research communities. For example, the contribution of the working group III on mitigation in the IPCC fifth assessment report (AR5) (IPCC, 2014) does not mention “machine learning” or “neural network”. This may be rooted in the dominance of the scenario literature within the IPCC reports, along with literature-based assessments of costs and potentials.

A confluence of several factors is likely to make ML more prevalent in mitigation research in the next years. First, increasingly rich sources of big data enable detailed studies in end-use sectors, including buildings and transportation. Second, the spatial dimension within cities and human settlements is gaining traction, since its explicit consideration in the IPCC’s AR5 (Seto et al., 2014). The spatial solutions considered correspond to urban end-use sectors: building and transportation accounted for 23% of the global energy-related CO_2 emissions and for 31% and 28% of global final energy use in 2014 (IPCC, 2018). Third, demand-side climate solutions, including behavioral nudges and digitalization, are introduced as explicit consideration in the IPCC’s AR6 (Creutzig et al., 2018), as such solutions might deliver short-term reductions in greenhouse gas (GHG) emissions (Grubler et al., 2018). Demand-side solutions require fine-grained analyses as they depend on local factors that vary widely across geographies (Grubler et al., 2018). Forth, big corporate players like Google are getting increasingly interested in applying their ML expertise on climate-change relevant issues (Google, 2019). It is hence important to overview the nascent literature and insights at the interface of machine learning and climate change mitigation and provide guidance on applications with relevance to GHG emission abatement.

Here, we chose the spatial dimension as an entry point and ask for the potential of machine learning to upscale geographically differentiated solutions in urban areas, and provide tailored urban planning solutions for decarbonizing cities worldwide. We suggest that machine learning methods could be central for spatially explicit and scalable climate solutions. For this, (i) we systematically review four domains of relevance for spatially explicit climate change solutions in cities – dedicated climate change mitigation studies, remote sensing, building and transport –, and (ii) we propose and detail a nested architecture of ML algorithms that combines the methods and insights from all these domains. As a result, we point to the potential of ML methods that could be systematically harvested by scientists to populate high-resolution urban planning models for climate change mitigation, which could help plan low-carbon cities with high contextual accuracy.

Fig. 1 | Growth in applications of machine learning in research on climate solutions. The literature using AI for climate mitigation explicitly (green) is almost inexistent before 2010, but relevant studies can be found in the other fields since the 1980's. The literature exhibits overall an exponential growth over the two last decades. Records are aggregated by time slices and divided by field and relevance to mitigation. Directly relevant records (dark colored) address greenhouse gas emissions or energy use, while instrumental records (light colored) do not make the link explicitly but offer important intermediary material (see supplementary methods). Pie charts show the share of publications per regions; "+" indicated regions with growing share of records between periods. (EU: European Union; ROW: Rest of the world.)



Few ML research targets climate change mitigation, but many provide relevant substance

In the following, we systematically map the relevant literature, based on the archive available in the online search engine Web of Science. For each publication, we extract its specific topic and the ML method used. Four search queries were designed to detect relevant literature on 1) ML and climate change mitigation; 2) ML and remote sensing, 3) ML and buildings; 4) ML and urban transportation. Queries were iteratively defined based on expressions present in the articles. The protocol follows the ROSES reporting standards for systematic maps (Haddaway & Macura, 2018; James et al., 2016), and includes manual labelling of several thousand of articles and non-negative matrix factorization (Lee & Seung, 1999) for topic discovery.

We find few studies relying on ML methods to directly address climate change mitigation, but more than 10 times more for sector-specific topics with direct or instrumental relevance to mitigation strategies (see Fig 1 and supplementary methods). Directly relevant records address GHG emissions or energy use. Instrumental records do not make the link explicitly but offer important intermediary material, for example, data on the building stock that in turn can help develop mitigation strategies in the sector. Overall, we identify 121 publications for climate change mitigation, 1,120 for buildings, 1,705 for transportation and 8,824 for remote sensing (see Fig 1).

Only 121 studies apply ML methods for researching explicitly climate change mitigation. We defined "climate change mitigation" using a conservative search string, which aimed to retrieve studies that specifically use language and goals of core climate change mitigation communities (see details in supplementary methods). This explains why we find a much smaller set of studies compared to (Rolnick et al., 2019). Supervised learning methods account for about 80% of all ML applications in mitigation studies (Fig. 2), relying on random forest, support vector learning or shallow artificial neural networks to generalize insights from data to other temporal or geographical domains.

	Deep Learning	Shallow neural networks	Decision trees	Support Vector Learning	Others Clas./Reg.	Recurrent neural networks	Reinforcement learning*	Clustering	Dimensionality reduction	ML not defined
Climate change mitigation *n=121*										
Agriculture	0	1	5	0	0	0	0	1	1	0
Soil	0	2	10	2	0	0	0	2	1	2
Cities	0	8	1	3	1	0	0	0	1	1
Forest and Biomass	0	6	21	4	1	0	1	1	4	4
Energy	0	8	2	4	0	0	1	3	2	1
Human behavior	0	0	1	0	0	0	0	1	3	1
Macro-scale	0	6	0	3	1	0	0	4	0	2
Remote sensing *n=8824*										
Air	35	228	146	114	40	5	6	51	94	21
Biomass	118	1069	1778	1076	202	17	39	233	696	82
Carbon	2	103	171	62	16	2	5	18	60	16
Earth surface	16	388	304	231	35	2	5	54	204	23
Impervious/Built-up	120	395	410	471	126	7	17	120	195	52
Water	37	797	483	403	62	6	23	116	354	51
Others	394	707	38	784	362	28	36	321	730	92
Buildings *n=1120*										
Building sub-system	10	189	23	38	13	3	47	17	37	42
Whole building	10	247	33	60	27	7	7	12	26	9
Environmental factors	1	43	5	12	1	0	1	3	3	2
Medium scale	2	70	25	17	8	1	8	15	10	8
Large scale	2	57	18	12	3	1	3	9	3	6
Human factor	1	35	10	17	8	3	15	11	10	15
Urban transportation *n=1705*										
Mobility patterns	8	28	27	13	13	7	30	45	20	17
Road traffic optimization	45	228	35	62	40	27	86	42	25	46
Public transport opt.	8	77	21	30	11	0	13	21	17	9
Travel behavior	1	35	30	9	3	1	21	4	5	7
Low carbon mobility	2	17	5	1	2	2	13	12	4	7
Planning	0	28	9	5	1	0	2	16	13	3
Others	60	203	61	102	48	7	41	57	71	50

Classes of machine learning methods

Fig. 2 | Summary of machine learning methods reviewed. Remote sensing, and to lesser degree, spatial studies in mobility and buildings, rely on ML methods, while climate change mitigation studies only scarcely build on ML methods. Supervised learning tasks (columns 1 to 6) are the most frequent applications in all fields. The information was extracted from the publicly available metadata of the records; 'Machine Learning not defined' is reported when there is no specific method available from the metadata. When several groups of methods are used in a record (e.g. dimensionality reduction and supervised learning), the record is counted in both categories. The gradient of the colors on the heat-map reflects the number of articles. *Graphical models were added to this category.

Recent progress in remote sensing is grounded in ML and DL methods, which are now widely used in this field (Zhu et al., 2017). ML has been a common tool in image processing for years, which has likely encouraged remote sensing researchers to use the same techniques. Remote sensing inputs are high-dimensional pictures with a large number of pixels and potentially many spectral bands. Finding relevant features is therefore challenging, requiring methods that can learn complex patterns, which may also explain why deep learning was used most often in remote sensing studies compared to the other fields reviewed (see Fig 2).

In the 1120 studies in the building sector, ML-based analysis, mostly at individual building scale, investigates efficient energy use, an objective in support of climate change mitigation. For example, ML identifies contextual determinants of energy uses, e.g. appliance use from electricity load (Kelly & Knottenbelt, 2015) or faults in mechanical systems (Lei et al., 2016). These results have helped analysts assess solutions that are tailored to highly specific contexts, and have the potential to inform how solutions can be spread across building stock and users.

Out of 1700 publications on transportation, we find about half (48%) directly relevant, although with varying mitigation potential. The efficiency of road networks is the option most investigated, primarily with the goal of reducing congestion; but only few studies investigated energy use explicitly (3% each). The other half (52%) of the relevant records can play an instrumental role for mitigation studies, by providing background data on travel determinants, flows, modes and infrastructures, which are all relevant to low-carbon cities. Transportation has the highest number of successful reinforcement learning applications (Fig 2), such as (Wen et al., 2017).

From our screening of the identified literature, we extracted four structural components of ML studies in support of climate mitigation solutions in cities, described in Fig 3A and summarized with relevant examples in Table 2: (1) mapping the infrastructure, (2) improving energy efficiency, (3) identifying behavioral patterns, and (4) infrastructure planning. In the following, for each component we show selected examples of how ML catalyzes geographically explicit climate solutions (Fig. 3B, see Table 1 for a summary of examples). ML is used for general tasks across application areas, for example forecasts; we summarized those tasks in Box 2.

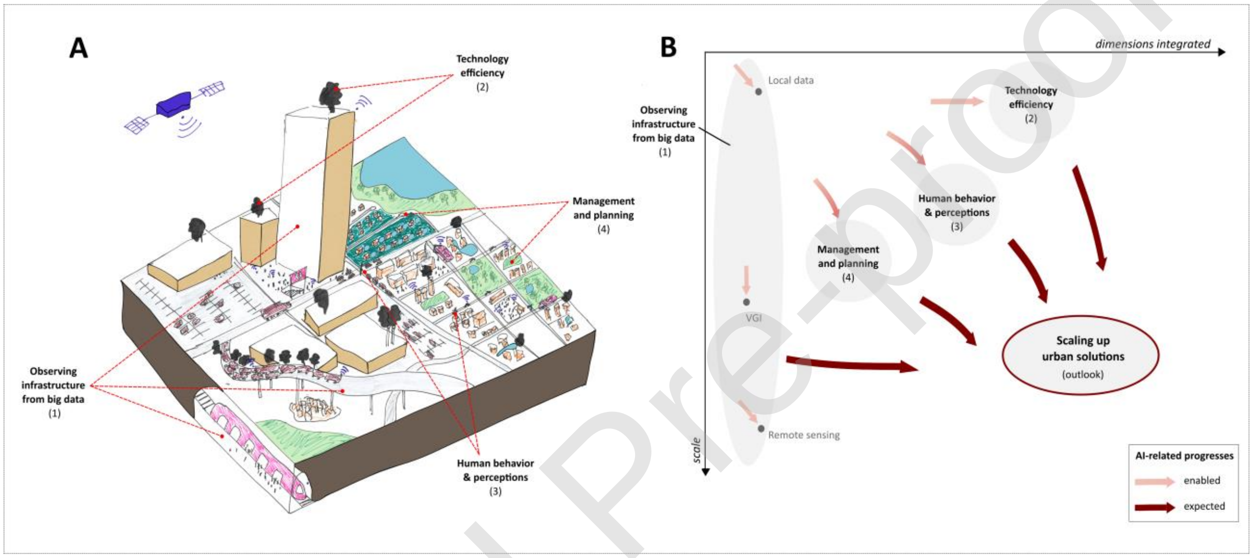


Fig. 3 | Towards scaling up urban solutions with machine learning and big data. (**A**) Main components of ML research on urban spaces relevant to climate change mitigation. Data sensed in the physical world and processed by ML enables modeling and predicting cities' infrastructures and activities, assess at fine-grain their energy use and emission patterns, and model different future pathways towards low-carbon societies. (**B**) Integrating ML studies on cities, has the potential to provide high-resolution low-carbon urban planning worldwide. Integrating approaches realizes a scaling up of urban solutions. (VGI: Voluntary Geographical Information)

(1) Mapping infrastructure

Infrastructure is the physical basis of human societies and preformats energy use and GHG emissions (F. Creutzig, Fernandez, et al., 2016). ML extracts from unstructured information, like remote sensing, knowledge about geometrics and semantics of infrastructure, helps enrich data-scarce environments, e.g., in developing countries, and offer ways to integrate spatial models with climate information.

From remote sensing to infrastructure models. Remote sensing, and specifically satellite imaging, enables to acquire data remotely on land-use patterns, land-use change, and other spatial metrics. Operating on these data, ML algorithms characterize human settlements by identifying geometric, e.g. footprint (Microsoft, 2018), height (Biljecki et al., 2017), and location (Esch et al., 2017), as well as semantic information, e.g. buildings usage (Sturrock et al., 2018; Wurm et al., 2016) and physical properties (Blaha et al., 2016; Tusting et al., 2019). As output, maps can display built-up areas worldwide at a few meters resolution (Esch et al., 2017), or at the country level, precise building footprints (Microsoft, 2018). The purpose of building usage, for example, commercial vs. residential or new vs. historical buildings can be classified from open geographical data (Sturrock et al., 2018) or survey data (Tusting et al., 2019) with supervised learning. A new class of studies aim for complete semantic 3D reconstruction, e.g. from aerial imagery with class labels for different parts of buildings (Blaha et al., 2016).

Generalizing to data-scarce environments. Where there is little ground data, as it is often the case in developing countries, ML can discover pivotal proxies from remote sensing. Such proxies go beyond describing infrastructure and can also identify socio-economic characteristics of a region. For this purpose, deep learning methods compress information from the spatial settings of cities into abstract features that are relevant for predicting metrics such as poverty (Jean et al., 2016) or demographics (Naik et al., 2016). For example, (Jean et al., 2016) used nighttime light intensity and corresponding daytime imagery to extract spatial features that revealed predictive of consumption expenditure and asset wealth.

Integration of spatial big data and climate semantics. Remote sensing contributed most directly to mitigation research in identifying deforestation and similar land use dynamics. Several pioneering publications in the 2000's used regression trees to assess the deforestation in tropical or humid forests (R. S. DeFries et al., 2002), to retrieve the associated carbon emission with spatially explicit resolution (Baccini et al., 2012) or to demonstrate that urban growth drives deforestation more than, counterintuitively, rural population growth (Ruth S. DeFries et al., 2010).

Remote sensing applications also support the identification of spatial patterns of CO_2 emissions and deployment of mitigation technologies. For example, Bayesian networks characterize the relationships between remote sensing data and urban CO_2 concentrations (J. Tao et al., 2014). DL applications identify installed photovoltaic systems but also the most suitable roofs and regions for photovoltaics deployment with high geographical precision (Yu et al., 2018). And to assess which settlements are most-suited for district heating system, clustering methods determine relevant building types by fusing data sources (Geiß et al., 2011).

(2) Improving energy efficiency

Individual infrastructural components are critical for the efficiency of the urban system (Gershenfeld et al., 2010) and have been subject to more precise modelling at small scale. ML helps optimize buildings by making sense of the complex electricity data, as well as transportation modes, for instance by analyzing their trajectories. A key challenge for ML applications is to extend the relevance of idiosyncratic results from single cases to larger populations.

Vehicles efficiency. In road transport, emissions and other externalities have been assessed in only a limited but growing set of applications (3%). Emissions have been in certain cases be metered and analyzed at finer scale than before: ML approaches range from downscaling national transportation emissions to the street level (Alam et al., 2018), to estimating of vehicle emission from smartphone GPS traces (Lehmann & Gross, 2017), and analyzing of the emissions associated with the current German fleet development (Krause et al., 2016).

Reducing inefficacies in driving offers marginal emissions reduction potential, with opportunities for granular options well-fitted to the application of ML. Trading-off travel time and emission reduction for eco-routing entails combining a large number of combinations in a highly interactive way, which is possible with support vector machines (Zeng et al., 2017). Inefficient driving behavior can be learned at the individual level: DL can reduce energy use of a trip by reasoning about observed driving behavior and related driver characteristics (de Penning et al., 2014), or discover regions where drivers repetitively drive inefficiently (Corcoba Magana & Munoz-Organero, 2015).

Single building optimization. A central tenet of the literature applying ML to buildings is to attain thermal comfort while minimizing energy use. Single-building models predict energy use patterns, relying on building-scale big data such as smart metering and internet of things (Shaikh et al., 2014), and by making sense of time series of electricity demand in a multiplicity of contexts (from single-household passive building to large hospital) and under the influence of various impacting factors (from weather to defections in ventilation systems) (Harish & Kumar, 2016). When there are no smart meters, energy disaggregation techniques enable to augment the available level of granularity by understanding individual appliance loads from the total electricity consumption signal (Kelly & Knottenbelt, 2015). In turn, ML applications can help reduce energy losses in specific appliances and systems in the building, in particular heating, ventilation, and air conditioning (HVAC) systems (Afroz et al., 2018). Examples include correcting malfunctions and unnecessary operations in energy intensive systems (Lei et al., 2016), or better controlling systems e.g. water heating (Kazmi et al., 2018). However, an important limitation of these applications is that relying on a large amount of precise data confines most of the studies to one or a few buildings (~75% of the applications).

Integrating high precision modeling. Upscaling the spatial relevance of current high-resolution applications in building energy use modelling also requires generalization from data-rich to data-scarce contexts. At various scales, from building to regions, the literature exhibits examples with supervised learning algorithms trained on high-quality data (Khayatian et al., 2017; Kontokosta & Tull, 2017). Methodologies for cross-building transfer learning include training neural networks on merged data from similar buildings with different distributions and different seasonal profiles (Ribeiro et al., 2018), and using reinforcement learning and deep neural networks conjointly to develop models that generalize from commercial to residential buildings, or from gas- to power-heated buildings (Mocanu et al., 2016).

(3) Identifying behavioral patterns

Dwellers' choices ultimately determine activity levels and resulting emissions (F. Creutzig et al., 2018). Behavioral data and models are focal to assess the propensity of actors to change towards more sustainable societies. Integrating human behaviors in city models can help identify dynamical feedbacks: for example, infrastructure provision (such as bike lanes) can foster changes in mobility choices. ML offer insights in sensed data and recorded traces of users of digital services, and supports the analysis of surveys and experimental data.

Understanding cities dynamics. ML is used to preprocess data from urban sensors and to later identify behavioral patterns of mobility and building users. For example, computer vision autonomously classifies and detects objects on images and videos (Boukerche et al., 2017); and thus computer vision techniques monitor traffic, automatically count vehicles and even derive socioeconomic information relying on data from street and road imagery (traffic cameras and Google Street View). For example, (Gebru et al., 2017) applied a convolutional neural network to street view images to classify motor vehicles encountered in particular neighborhoods, by this predicted income and voting patterns in neighborhoods across 200 US cities.

By making sense of mobility patterns, researchers provide valuable information on the structure of cities for planning (6% of transport applications, see an overview (Zhao et al., 2016)). Mobility pattern mining investigates locations where dwellers tend to go often, what are the trip purposes, as well as which are urban activities or functions, e.g. using so-called points of interest (POIs). Examples include retrieving transportation mode from GPS (X. Zhu et al., 2016) or trip chains from smart cards (G. Han & Sohn, 2016). Beyond mobility, call details records can be used to predict directly domestic energy use (Bogomolov et al., 2016).

Engaging with human behaviors. At the individual level, behavioral research investigates how psychological aspects – e.g. engagement (Jones et al., 2017) or acceptance of novelty (Carr-Cornish et al., 2011) – are relevant to better target interventions. For example, predicting environmental attitudes with ML revealed that the scale of future consequences better predicts attitudes towards climate than the factors traditionally considered, e.g. income and education (Beiser-McGrath & Huber, 2018)**.**

Understanding the human dimension of building energy use (∼10%), such as occupant behavior or investments, is instrumental to reducing absolute consumption levels. One type of studies investigates how to monitor occupancy and usage (∼6%), mostly with the goals of more accurately predicting energy usage or controlling smart adaptable devices. Another type of studies scrutinizes occupant behaviors (∼2%) and engages directly with occupants by using heuristic models of behavioral triggers and resistances to more energy-efficient lifestyles. An example of the second category uses interpretable “trees”, where branches identify homogeneous sub-groups of consumers with respect to socio-economic attributes, and predictive of their likely reaction to a specific type of energy efficiency communication (Albert & Maasoumy, 2016). In another example, hierarchical clustering of people's mental models about their level of domestic energy consumption consistently revealed categorizations emphasizing the importance of the location of appliances in the house (Gabe-Thomas et al., 2016).

Understanding mobility choices is key to device low-carbon policies in urban transportation. Estimating the determinants of decisions, ML methods perform better than standard discrete choice models, for instance in modelling car ownership (Paredes et al., 2017). DL (Wang & Zhao, 2018) or semi-supervised Bayesian learning (Yang et al., 2018) directly substitute for the traditional logit and probit functions in the model formulation. Coupling neural network with an agent-based model, an innovative study explored the role of emotions in agents' decision to buy an electric vehicle and simulated the effect of public policies on decisions, e.g., by introducing exclusive lanes for electric vehicles (Wolf et al., 2015).

(4) Planning and management

Providing and managing infrastructures frames current and future behaviors. Planning and operating low-carbon infrastructure can be supported by detailed maps, models of energy efficiency, and models of human behavior.

Infrastructure management. The tension between the need of case-specific studies and scalable solutions for deep decarbonization – particularly exacerbated in the research on buildings – could be partly alleviated by ML applications that provide new tools to identify and transfer individual solutions between different contexts (Ma & Cheng, 2017; Mocanu et al., 2016; Zhang et al., 2018).

ML methods can accelerate spatially explicit assessments of the energy and retrofit potential to entire building stocks, by imputing building level data where none is available (Zhang et al., 2018). Going beyond micro-optimization with large-scale and precise assessments is crucial to leverage the full potential of ML for mitigating the emissions from the building sector. About 7**%** of the literature base applies ML algorithms to identify energy use patterns in large building stocks. For example, in New York City, building characteristics and fuel type for over 15,000 buildings was used to predict the energy use of the city's building stock, which is about one million buildings (Kontokosta & Tull, 2017). The same research group could also identify in these data which buildings drive aggregate energy use and where energy conservation schemes were most effective (Papadopoulos et al., 2018). A further step is the production of retrofit typologies at regional or national scale. For example, ML approaches can compress building characteristics data into indicators that support national building retrofit policies (Khayatian et al., 2017).

Intelligent transportation systems make traffic patterns more efficient, but may be compromised by rebound effects. However, coupled with public policy, intelligent transportation systems could deliver substantial system-wide efficiency gains and resulting contributions to climate change mitigation. For traffic control – with ramp metering, traffic lights or signal control – reinforcement learning can reduce energy use by optimizing circulation (Yau et al., 2017).

Modify infrastructures. A handful of promising studies (1%) have targeted urban planning, investigating the links between urban form at different scales and sustainability metrics (Ding et al., 2018; T. Tao et al., 2020; Wu et al., 2019). For example, one application at the city-level found a strongly non-linear relationship between distance to city center and driving distance in Oslo – offering novel evidence relevant for the mitigation potential of denser cities (Ding et al., 2018). At the neighborhood level, another study could show the impact of walking distance to a station on transit ridership (T. Tao et al., 2020).

ML can also enable carbon intensity improvements in land transport by supporting the development of low-carbon modes. For example, one class of ML applications optimizes the charging infrastructure of electric vehicles (Rigas et al., 2015). ML was also applied to estimate the power demand of electric vehicles (Longo et al., 2017). For choosing the location of new docking-station-free shared bikes, ML can predict the bike trips supply and demand (Xu et al., 2018) and reinforcement learning can help rebalance the fleet within the city (Wen et al., 2017).

Transfers of knowledge across geographies. By "decoding deep similarities" as phrased by P. Ball (Ball, 2017), comparative research enables local policy makers to learn about structural components of urban transitions into low-carbon futures, identified from large scale comparisons, and how they could apply or not to the specifics of their cities. Many papers investigate the spatial transferability of their approach (Biljecki et al., 2017; Jean et al., 2016; Ma & Cheng, 2017; Mocanu et al., 2016; Tang et al., 2018). For example, some ML-based travel demand models can transfer well to other cities with similar characteristics (Tang et al., 2018).

Typologies using supervised learning allow sorting contexts according to the drivers of energy use and GHG emissions, and also offer further inputs for projection in new geographies. For example, Ma and Cheng identified socio-economic features from the leading US counties in the adoption of green building standards, which they used to predict green building deployment in Chinese provinces (Ma & Cheng, 2017). Another example incorporated city-specific context into worldwide analysis by creating a typology of hundreds of cities' energy use, consisting of clusters of cities, each characterized by thresholding combinations of energy-use drivers, such as income, population density, fuel prices, and local climate variables (Creutzig et al., 2015).

The existing research shows many relevant strands for contextualized and geographically differentiated climate change mitigation. But it is unclear how the individual components act together, and how the existing expertise can be organized to a community-wide effort in providing agile, efficient, and scalable solutions for low-carbon cities. In the following, we discuss an algorithmic architecture that aims to fill this gap.

An algorithmic architecture for scalable low-carbon urban planning

Machine learning can foster a new class of spatially precise climate mitigation solutions in urban planning that can be scaled to data-scarce environments, if certain conditions are fulfilled. Building on our systematic analysis of the literature, we detail how to tailor workflows that are intended to support decision making at the municipal level and to foster systematic knowledge sharing between cities. Our focus is on urban planning, but most points are also pertinent for other mitigation studies with explicit spatial resolution.

We first discuss limitations and shortcomings we identified in the existing literature base. Second, we articulate an architecture of "Machine learning for low-carbon Urban Planning" (ML-UP) that systematically utilizes machine learning algorithms at different stages of data processing. Third, we outline the potential of the proposed modelling architecture to enhance current global urban climate mitigation scenarios, and to provide geographically differentiated mitigation strategies across human settlements.

Optimizing existing infrastructure with ML is insufficient. We see three main limitations for applying the surveyed literature for geographically differentiated climate change mitigation: (i) a predominant focus on behavioral models and business applications that also increase social risks of surveillance; (ii) a large dominance of utilizing ML for efficient use of existing infrastructure; (iii) a resulting lack of public policy analysis.

Many ML applications with relevance for reducing energy demand or GHG emissions intend to change user behavior (Albert & Maasoumy, 2016; Bertone et al., 2018; Gabe-Thomas et al., 2016; Wolf et al., 2015). The role of behavioral models for creating mitigation-relevant social knowledge is promising but they also bear important concerns about privacy and freedom. Behavioral models take advantage of user data to capture patterns on specific people or groups of people lifestyles and choices. Even if individual data, such as mobility patterns, can be anonymized, guaranteeing user privacy in behavioral models is very challenging (de Montjoye et al., 2013). Current developments hold value for research on urban functions but remain highly sensitive in the context of surveillance states (Couldry & Mejias, 2018). ML applications for behavioral models are nonetheless likely to expand due to multiple industries' interests, and opportunities from geo-located app data, which are widely collected and cheaper to access.

Overall, the reviewed literature from end-use sectors mostly investigates efficiency improvements, provided by measures such as smart metering and traffic flow optimization. 95% of the literature on buildings and more than 60% of the literature on transport covers such efficiency measures. Efficiency improvements counterfactually and effectively reduce energy demand, but this effect is impacted by rebound effects (Azevedo, 2014). Reducing road congestion can produce new demand and lead to more additional traffic, partially negating the effect of demand reduction (Gossart, 2015). Similarly, smart appliances can generate energy use from new end-uses. Both sorts of rebound effects are typically in the order of 10–20% (Chen et al., 2018; Hymel et al., 2010); hence they reduce the effectiveness of technological efficiency improvements without disqualifying it.

In contrast, public policy aiming to utilize ML for climate change mitigation or energy conversation is only scarcely investigated in the literature. Only 1.5% of the transport and building literature investigates policies (26 and 18 studies respectively). These include ex-post evaluation of fuel standard policies relying on statistical learning techniques (Huseynov & Palma, 2018), and the application of Bayesian networks to investigate how different financing mechanisms would affect the willingness to retrofit (Bertone et al., 2018). For the uptake of improved technologies, policies are crucial to signal and incentivize shifts. Public policies could foster the use of ML towards more effective climate change mitigation, and constrain rebound effects of efficiency improvements – through infrastructure provision, price incentives or standards.

Leveraging the mitigation potential of urban planning. Urban spatial configurations harbor robust and long-term mitigation potential (Creutzig et al., 2018). For example, urban sprawl induces carbon-intensive transport and building use; but if connectivity is high, land-use is mixed, and structures are compact, the spatial structures are conducive of low-carbon transport systems and short-distance travel (Seto et al., 2014). Urban planning with its focus on building structures and mobility patterns is central to providing low-carbon infrastructures, but until now has only been considered in a small segment of the machine learning literature pertinent to climate change mitigation.

Focusing on urban planning presents multiple advantages. First, it targets absolute demand reduction, not only efficiency improvements. For example, for transportation, planning can reduce the distances traveled (Ewing & Cervero, 2017); for buildings, planning can help develop neighborhoods that synergically re-utilize resources (Petit-Boix & Leipold, 2018). Second, infrastructures can endogenously shape people's preferences (F. Creutzig, Fernandez, et al., 2016). People are more likely to shift to less emitting transport modes, such as bikes, if bike lanes are abundant, convenient, and safe. Third, it is important to prevent lock-ins in poorly designed infrastructure: an infrastructure will be kept for decades, and bad planning alone could lead to the exhaustion of the 1.5C carbon budget by 2050 (F. Creutzig, Agoston, et al., 2016). Fourth, urban planning requires to a much lesser extend behavioral models and micro-sensors, and hence less or even zero privacy-compromising big data.

There are two main challenges towards providing low-carbon urban planning scenarios at high spatial and contextual resolution: (1) generating high-resolution energy or emission data to create digital models of cities' energy use and emission patterns; and (2) subsequently identifying concrete urban planning solutions for climate change mitigation.

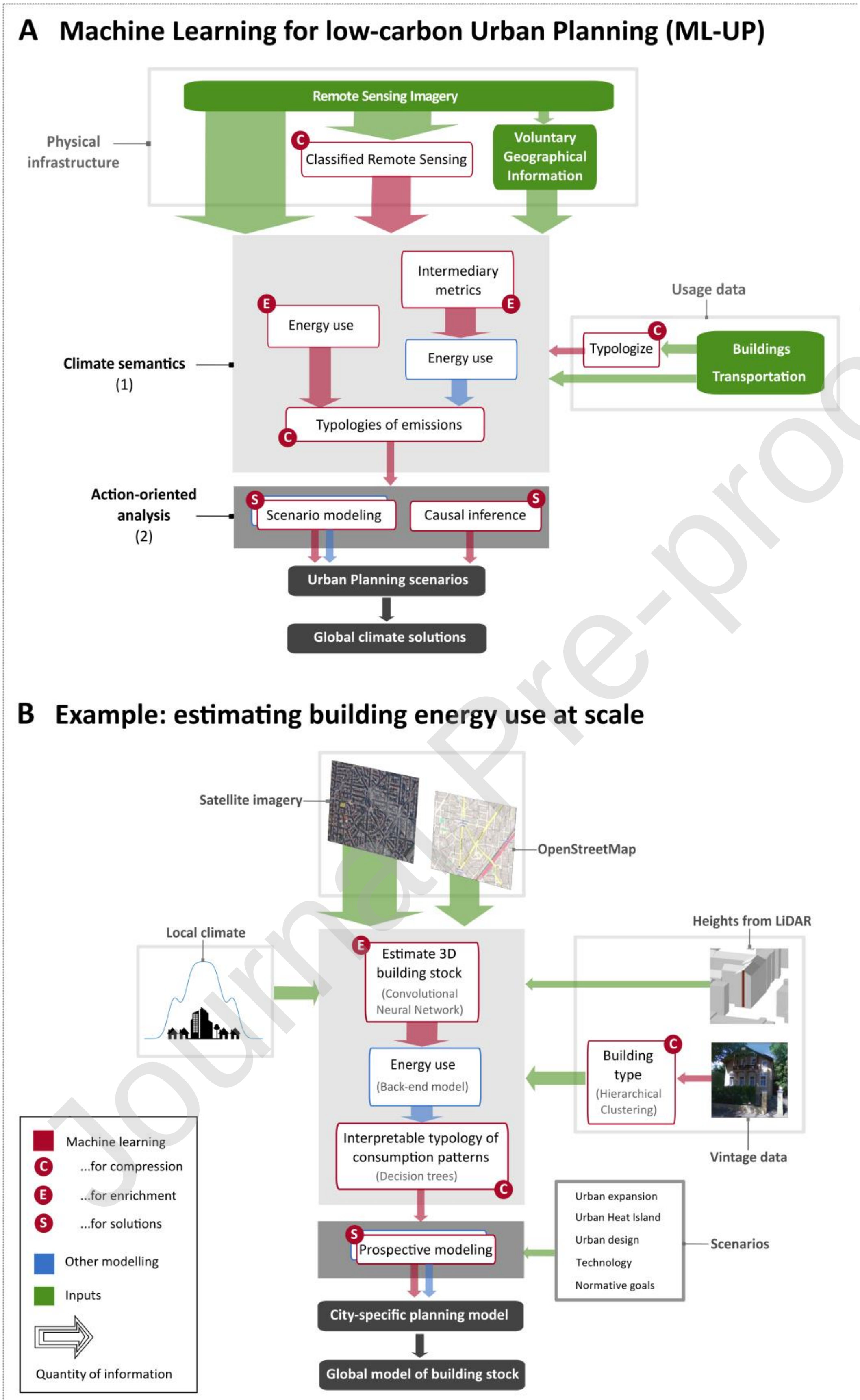

A Machine Learning for low-carbon Urban Planning (ML-UP)
Remote Sensing Imagery
Physical infrastructure
Classified Remote Sensing
Voluntary Geographical Information
Intermediary metrics
Energy use
Energy use
Usage data
Typologize
Buildings
Transportation
Climate semantics (1)
Typologies of emissions
Action-oriented analysis (2)
Scenario modeling
Causal inference
Urban Planning scenarios
Global climate solutions
B Example: estimating building energy use at scale
Satellite imagery
OpenStreetMap
Local climate
Heights from LiDAR
Estimate 3D building stock (Convolutional Neural Network)
Energy use (Back-end model)
Building type (Hierarchical Clustering)
Vintage data
Interpretable typology of consumption patterns (Decision trees)
Machine learning
...for compression
...for enrichment
...for solutions
Other modelling
Inputs
Quantity of information
Urban expansion
Urban Heat Island
Urban design
Technology
Normative goals
Scenarios
Prospective modeling
City-specific planning model
Global model of building stock

Fig. 4 | An architecture of machine learning for low-carbon urban planning (ML-UP). (**A**) ML-UP is an information flow from big data to semantically relevant data for climate change mitigation-oriented urban planning. The data can be processed by a succession of different phases including ML and other media. (**B**) An example workflow for estimating energy use of individual buildings at large scale. Spatial data available at large scale are trained with precisely metered building data.

The proposed infrastructure of "Machine Learning for low-carbon Urban Planning" (ML-UP) addresses these two classes of challenges, and is summarized in Fig 4A and exemplified in Fig 4B for a concrete example on buildings. The architecture is structured around sequentially stacked ML applications. ML methods are further grouped into three usages: information compression (from high dimensional data to useful features), information enrichment (by predicting future or inaccessible information) and solution-oriented information processing – see red circles in Fig 4A.

(1) Generating climate semantics. As first key challenge, generating mitigation-relevant digital models of cities requires spatialized climate data, i.e. information on urban GHG emissions, or more specifically on the energy use of each building of a given building stock. The main bottleneck for modelling spatially at such scales is the requirement of consistent gridded data – while existing data on cities are selective, biased, and inconsistent (F. Creutzig et al., 2019).

The starting point of ML-UP is to process the available spatial data on cities (Fig 4A). Remote sensing data is compressed by unsupervised or supervised learning to classify earth observations into semantic and geometric information on land use classes. Additionally, volunteered geographic information, such as OpenStreetMap, provides primary data that is spatially explicit (Fig 4A, physical infrastructure). Many geo-localized digital traces can be also exploited. For example, natural language processing tools can identify geo-localized urban activities from publicly available text obtained from, e.g. Twitter (Rahim Taleqani et al., 2019), where location information can be extracted and associated to activities of relevance.

The central task consists in generating climate change mitigation semantics, such as energy use or GHG emissions (central box in Fig 4A), by integrating spatialized information on energy end-uses with the processed information on the spatial structure of cities. Climate semantics can be generated using different routes: directly or through intermediary metrics.

The direct route is to train urban form data with spatialized energy use in supervised learning algorithms (Fig 4A central box left). It has already been demonstrated that climate mitigation semantics can be derived from the characteristics of urban activities and form – e.g. density, land use mix, connectivity and accessibility (Creutzig, Fernandez, et al., 2016; Ewing & Cervero, 2017; Silva et al., 2018; Silva et al., 2017). A spatially-explicit analysis of the energy demand for buildings and transportation using ML shows that urban form explains about 80% of the variation in energy use in the city of Porto (Silva et al., 2017). Mobility patterns, transport networks, attraction poles, or population densities are all related to the energy used for transportation. Similarly, features on buildings – such as their own geometries or their configuration within neighborhoods, for example whether building are contiguous or not – also contain relevant information about their energy use and emission patterns. The extent of the potential of this direct route remains so far speculative and depends on the availability of training data as well transferability performances.

An alternative route is to predict intermediate metrics and use back-end modelling (Fig 4A central box right). Intermediary metrics extracted from spatial data – either abstract encodings or real quantities e.g. the volume of a building or the distance travelled by a vehicle – can be predictors for energy use and GHG emissions (Liu et al., 2018; Robinson et al., 2017). This route utilizes several ML methods sequentially for different tasks (Kaack et al., 2019; Liu et al., 2018). ML and other modelling can also be combined (Abdulkareem, 2019). Simple back-of-the-envelope models relying on identified mechanical relations, regression techniques (Jean et al., 2016; Liu et al., 2018), as well as fully-fledged systems dynamics or agent-based models (Toole et al., 2015), use the intermediary metrics as inputs. The example on Fig. 4B follows the sequential route. Urban form data is trained on ground truth building heights with a convolutional neural network to estimate a 3D building stock as an intermediary metric. Main types of buildings in the region are clustered in an interpretable fashion, that reveals the key parameters influencing energy use across clusters relying on regression trees (Baiocchi et al., 2015; Creutzig et al., 2015). Both sets of information are fed into a simple energy model together with climatic data, to find the minimal energy use for heating and cooling the entire building stock. These distinct phases can favor more interpretability – and in certain cases, more scalability – between feature extraction and prediction.

Overall, information on energy use and emissions from buildings and the transportation sectors could be generated at large scale and at high resolution, while respecting heterogeneities across scales. Another phase of compression via typologies can help relate energy use patterns to categories of emitters (F. Creutzig et al., 2015), and grasp better in which conditions certain GHG emissions patterns occur.

(2) Actionable solutions. The second key challenge is to make the knowledge about the energy use and emission patterns of cities actionable for policy makers. All cities are different and develop with geography, culture, demographics, and economy along path-dependent trajectories (Arthur, 1988). Such context requires accounting dynamically for changes in relevant interacting dimensions within urban environments, including economic mechanisms, urban form, local climate, and social norms or ties. Hence, a second block of ML-UP focuses on action-oriented prospective predictions, making use of scenario techniques and advances in causal inference research (Fig 4A). The energy performance of urban planning scenarios can be assessed with supervised learning. A study shows that higher density and constructing around key public transport stations were yielding the most energy consumption reductions (M. Silva et al., 2018). To further evaluate policies, causal inference and explanatory models (Bertone et al., 2018) complement the ML-UP architecture. For example, the city-based difference-in-difference methodology (Blake et al., 2015) might be helpful to evaluate the causal effect of providing cycling infrastructure on modal shift. As another example, lasso regression used on electricity consumption data from schools in California finds a causal effect of energy efficiency interventions that falls short of the expected savings (Burlig et al., 2017).

There are several barriers to ML-UP. First, aggregating heterogeneous sources of data bears considerably search costs, as well as complex (ML-based and other) harmonization methods, like data matching (Bordes et al., 2014). The quality of training data could be insufficient in certain cases, making it impossible even for advanced generalization methods to fill the gaps completely. Second, the data underlying behavioral models are inconsistent with privacy concerns and undesired social control if not carefully governed. Third, interpretability is repetitively raised as a central issue in machine learning (Montavon et al., 2018; Reichstein et al., 2019) and consistency issues could appear within phases of ML-UP. Fourth, not everything is quantifiable and much of the local context in term of culture, jurisdictions, etc. are difficult to fit in this framework and require other types of analyses. However, Google's Environmental

Insights Explorer and recent advances in Earth System Modelling (Rasp et al., 2018; Reichstein et al., 2019) demonstrate that modular and sequential architectures similar to ML-UP successfully integrate machine learning and physical modelling for spatial and dynamical problems.

Governance implications. The outlined architecture of ML for low-carbon urban planning – if deployed at scale – could have implications for the structuring of the solution space for climate change mitigation. It also stipulates more agile and rapid deployment of effective solution strategies with the potential to generalize insights from data-rich to data-poor settings, and thus help policy identification for more human settlements than currently possible.

First, results from high spatial resolution models, if applied to a large number of cities at the global scale, would restructure the solution space for climate change mitigation. Going beyond the highly aggregated state-of-the-art global models or back-of-the-envelope calculations is a necessity to provide realistic global estimates of the mitigation potential from cities. Spatially explicit modeling can improve generic global models by enabling a better accounting of local heterogeneity.

Generating models of urban climate mitigation solutions at high spatial resolution would transform global environmental assessments, such as those of the IPCC. Instead of providing long-term scenarios with abstract policy suggestions, place-specific solution strategies could then be compared and evaluated. Up to now, IPCC reports emphasized technologies that were possible to operationalize in energy system models. Contextualized demand-side solutions, such as a myriad of behavior and infrastructure options relevant for energy savings in buildings, in contrast, were difficult if not impossible to model and to represent in high-level models. With the big data/ML approach, contextualized, place-specific and demand-side options could be evaluated at the global scale, while respecting the local context. As a result, the understanding of the solution space might experience a considerable shift away from energy supply-side technologies, to urban planning and contextualized solutions.

Second, ML-UP aims to provide urban policy makers actionable information for implementing municipal climate action. Mayors of the biggest and richest cities advance climate action, but medium and small-sized cities, where the largest part of the world's urban population lives, mostly lack data-driven insights and policy commitment (Lamb et al., 2019; Nagendra et al., 2018). With data and learning across municipal jurisdictions, ML-UP aims to empower policy makers also of smaller and medium-sized cities to advance data science supported strategies. For example, standardized and comparable information of the ecological footprint of physical assets could support public investors' choices by giving them a better visibility on the impacts of their portfolio.

Typologies and synthesis of cases studies together help cities learn about climate solutions (Lamb et al., 2019). Developing taxonomies of cities may enable groups of similar cities to draw from the same pool of solutions, or learn from early pioneers in climate policy (Lamb et al., 2019). Existing qualitative typologies, like the Atkins Future Proofing Cities report, which linked for more than a hundred cities their structural features and detailed policies options, could become systematized and quantitatively accurate with ML. A growing research strain identified in this review uses ML to generate quantitative typologies at the city (Creutzig et al., 2015; Han et al., 2018), district (Baiocchi et al., 2015) and street scales (Louf & Barthelemy, 2014) that both capture universalities (e.g. fuel prices influence the sprawl of cities (Creutzig et al., 2015) or similar street patterns can identified across continents (Louf & Barthelemy, 2014)), and consider local specificities (e.g. by using as input local socio-economic variables (Creutzig et al., 2015) or urban form data (Louf & Barthelemy, 2014)). Nonetheless, quantitative typologies do not fully explore the underlying political and social conditions, dimensions that only can be brought through evidence syntheses of case studies as a complementary strategy (Lamb et al., 2019).

Third, ML-UP would have the highest value in low-income countries with small resources for bottom-up policy modelling. Cities in low-income countries with rapid population growth and urbanization are a priority for mitigation strategies (IPCC, 2014; Nagendra et al., 2018)**,** while the literature has been biased towards the global North and disproportionally focused on megacities (Lamb et al., 2019). Low-income countries tend to have little data available, which is often due to weaker governance and statistics collection capacities. This situation can hinder the use of ML, but it also make particularly relevant for predicting lacking information. In this context, those ML techniques that are most capable of taking advantage of small amount of available training data are particularly relevant and require further applications. Few-shots learning (Fei-Fei et al., 2006), meta-learning (Mishra et al., 2018), transfer learning (Jean et al., 2016) can be determinant in overcoming current limitations. ML could help formulate policies but their implementation depend local actors, with strong institutions and political will.

Conclusion

Artificial intelligence and machine learning hold considerable but still underutilized potential for geographically differentiated and contextualized design of measures that reduce GHG emissions. Distinct research contributions include the spatially explicit mapping of human settlements, the design of energy-efficient infrastructure use, the understanding of behavioral patterns, and the design of low-carbon urban infrastructures. We argue that the planning of low-carbon urban infrastructures carries high potential but currently receives scarce attention in the literature. We suggest an algorithmic architecture, ML-UP, that is designed to orient the intersection of machine learning research, urban planning and studies of climate change mitigation towards a common research framework. Climate change mitigation at relevant scale will only be achieved on conjunction with well-designed public policy.

The authors declare that they there is no conflict of interests.

References:

Abdulkareem, S. A. (2019). *Enhancing agent-based models with artificial intelligence for complex decision making* [PhD thesis]. University of Twente.

Acuto, M., Parnell, S., & Seto, K. C. (2018). Building a global urban science. *Nature Sustainability*, *1*(1), 2–4. https://doi.org/10.1038/s41893-017-0013-9

Afroz, Z., Shafiullah, G., Urmee, T., & Higgins, G. (2018). Modeling techniques used in building HVAC control systems: A review. *Renewable and Sustainable Energy Reviews*, *83*, 64–84. https://doi.org/10.1016/j.rser.2017.10.044

Alam, Md. S., Duffy, P., Hyde, B., & McNabola, A. (2018). Downscaling national road transport emission to street level: A case study in Dublin, Ireland. *Journal of Cleaner Production*, *183*, 797–809. https://doi.org/10.1016/j.jclepro.2018.02.206

Albert, A., & Maasoumy, M. (2016). Predictive segmentation of energy consumers. *Applied Energy*, *177*, 435–448. https://doi.org/10.1016/j.apenergy.2016.05.128

Arthur, W. B. (1988). Urban systems and historical path-dependence. In *Cities and Their Vital Systems: Infrastructure Past, Present, and Future*. https://www.popline.org/node/362909

Athey, S. (2017). Beyond prediction: Using big data for policy problems. *Science*, *355*(6324), 483–485. https://doi.org/10.1126/science.aal4321

Atlas AI. (n.d.). https://www.atlasai.co/

Azevedo, I. M. L. (2014). Consumer End-Use Energy Efficiency and Rebound Effects. *Annual Review of Environment and Resources*, *39*(1), 393–418. https://doi.org/10.1146/annurev-environ-021913-153558

Baccini, A., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P. S. A., Dubayah, R., Friedl, M. A., Samanta, S., & Houghton, R. A. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change*, *2*(3), 182–185. https://doi.org/10.1038/nclimate1354

Baiocchi, G., Creutzig, F., Minx, J., & Pichler, P.-P. (2015). A spatial typology of human settlements and their CO2 emissions in England. *Global Environmental Change*, *34*, 13–21. https://doi.org/10.1016/j.gloenvcha.2015.06.001

Ball, P. (2017). Complexity: Decoding deep similarities. *Nature*, *545*, 154. https://doi.org/10.1038/545154a

Beiser-McGrath, L. F., & Bernauer, T. (2019). Commitment failures are unlikely to undermine public support for the Paris agreement. *Nature Climate Change*, *9*(3), 248–252. https://doi.org/10.1038/s41558-019-0414-z

Beiser-McGrath, L. F., & Huber, R. A. (2018). Assessing the relative importance of psychological and demographic factors for predicting climate and environmental attitudes. *Climatic Change*, *149*(3–4), 335–347. https://doi.org/10.1007/s10584-018-2260-9

Bertone, E., Sahin, O., Stewart, R. A., Zou, P. X. W., Alam, M., Hampson, K., & Blair, E. (2018). Role of financial mechanisms for accelerating the rate of water and energy efficiency retrofits in Australian public buildings: Hybrid Bayesian Network and System Dynamics modelling approach. *Applied Energy*, *210*, 409–419. https://doi.org/10.1016/j.apenergy.2017.08.054

Biljecki, F., Ledoux, H., & Stoter, J. (2017). Generating 3D city models without elevation data. *Computers, Environment and Urban Systems*, *64*, 1–18. https://doi.org/10.1016/j.compenvurbsys.2017.01.001

Blaha, M., Vogel, C., Richard, A., Wegner, J. D., Pock, T., & Schindler, K. (2016). Large-Scale Semantic 3D Reconstruction: An Adaptive Multi-resolution Model for Multi-class Volumetric Labeling. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3176–3184. https://doi.org/10.1109/CVPR.2016.346

Blake, T., Nosko, C., & Tadelis, S. (2015). Consumer Heterogeneity and Paid Search Effectiveness: A Large-Scale Field Experiment: Paid Search Effectiveness. *Econometrica*, *83*(1), 155–174. https://doi.org/10.3982/ECTA12423

Bogomolov, A., Lepri, B., Larcher, R., Antonelli, F., Pianesi, F., & Pentland, A. (2016). Energy consumption prediction using people dynamics derived from cellular network data. *EPJ Data Science*, *5*(1), 13. https://doi.org/10.1140/epjds/s13688-016-0075-3

Bordes, A., Glorot, X., Weston, J., & Bengio, Y. (2014). A semantic matching energy function for learning with multi-relational data. *Machine Learning*, *94*(2), 233–259. https://doi.org/10.1007/s10994-013-5363-6

Boukerche, A., Siddiqui, A. J., & Mammeri, A. (2017). Automated Vehicle Detection and Classification: Models, Methods, and Techniques. *ACM Computing Surveys*, *50*(5), 1–39. https://doi.org/10.1145/3107614

Burlig, F., Knittel, C., Rapson, D., Reguant, M., & Wolfram, C. (2017). *Machine Learning from Schools about Energy Efficiency* (No. w23908). National Bureau of Economic Research. https://doi.org/10.3386/w23908

Callaghan, M. W., Minx, J. C., & Forster, P. M. (2020). A topography of climate change research. *Nature Climate Change*, *10*(2), 118–123. https://doi.org/10.1038/s41558-019-0684-5

Carr-Cornish, S., Ashworth, P., Gardner, J., & Fraser, S. J. (2011). Exploring the orientations which characterise the likely public acceptance of low emission energy technologies. *Climatic Change*, *107*(3–4), 549–565. https://doi.org/10.1007/s10584-010-9964-9

Chen, K., Li, Z., Lu, T.-P., Rau, P.-L. P., & Huang, D. (2018). Influence of Rebound Effect on Energy Saving in Smart Homes. In P.-L. P. Rau (Ed.), *Cross-Cultural Design. Applications in Cultural Heritage, Creativity and Social Development* (pp. 266–274). Springer International Publishing.

Corcoba Magana, V., & Munoz-Organero, M. (2015). Discovering Regions Where Users Drive Inefficiently on Regular Journeys. *IEEE Transactions on Intelligent Transportation Systems*, *16*(1), 221–234. https://doi.org/10.1109/TITS.2014.2328517

Couldry, N., & Mejias, U. A. (2018). Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject. *Television & New Media*, 1527476418796632. https://doi.org/10.1177/1527476418796632

Creutzig, F. (2016). Evolving Narratives of Low-Carbon Futures in Transportation. *Transport Reviews*, *36*(3), 341–360. https://doi.org/10.1080/01441647.2015.1079277

Creutzig, F., Agoston, P., Minx, J. C., Canadell, J. G., Andrew, R. M., Quéré, C. L., Peters, G. P., Sharifi, A., Yamagata, Y., & Dhakal, S. (2016). Urban infrastructure choices structure climate solutions. *Nature Climate Change*, *6*(12), 1054–1056. https://doi.org/10.1038/nclimate3169

Creutzig, F., Baiocchi, G., Bierkandt, R., Pichler, P.-P., & Seto, K. C. (2015). Global typology of urban energy use and potentials for an urbanization mitigation wedge. *Proceedings of the National Academy of Sciences*, *112*(20), 6283–6288. https://doi.org/10.1073/pnas.1315545112

Creutzig, F., Fernandez, B., Haberl, H., Khosla, R., Mulugetta, Y., & Seto, K. C. (2016). Beyond Technology: Demand-Side Solutions for Climate Change Mitigation. *Annual Review of Environment and Resources*, *41*(1), 173–198. https://doi.org/10.1146/annurev-environ-110615-085428

Creutzig, F., Lohrey, S., Bai, X., Baklanov, A., Dawson, R., Dhakal, S., Lamb, W. F., McPhearson, T., Minx, J., Munoz, E., & Walsh, B. (2019). Upscaling urban data science for global climate solutions. *Global Sustainability*, *2*. https://doi.org/10.1017/sus.2018.16

Creutzig, F., Roy, J., Lamb, W. F., Azevedo, I. M. L., Bruine, W., Dalkmann, H., Edelenbosch, O. Y., Geels, F. W., Grubler, A., Hepburn, C., Hertwich, E. G., Khosla, R., Mattauch, L., Minx, J. C., Ramakrishnan, A., Rao, N. D., Steinberger, J. K., Tavoni, M., Ürge-Vorsatz, D., & Weber, E. U. (2018). Towards demand-side solutions for mitigating climate change. *Nature Climate Change*, *8*, 4.

Creutzig, F. S., & Kammen, D. M. (2009). The Post-Copenhagen Roadmap Towards Sustainability: Differentiated Geographic Approaches, Integrated Over Goals. *Innovations: Technology, Governance, Globalization*, *4*(4), 301–321. https://doi.org/10.1162/itgg.2009.4.4.301

de Montjoye, Y.-A., Hidalgo, C. A., Verleysen, M., & Blondel, V. D. (2013). Unique in the Crowd: The privacy bounds of human mobility. *Scientific Reports*, *3*(1), 1376. https://doi.org/10.1038/srep01376

de Penning, L., d'Avila Garcez, A. S., Lamb, L. C., Stuiver, A., & Meyer, J.-J. Ch. (2014). Applying Neural-Symbolic Cognitive Agents in Intelligent Transport Systems to reduce CO2 emissions. *2014 International Joint Conference on Neural Networks (IJCNN)*, 55–62. https://doi.org/10.1109/IJCNN.2014.6889788

DeFries, R. S., Houghton, R. A., Hansen, M. C., Field, C. B., Skole, D., & Townshend, J. (2002). 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. https://doi.org/10.1073/pnas.182560099

DeFries, Ruth S., Rudel, T., Uriarte, M., & Hansen, M. (2010). Deforestation driven by urban population growth and agricultural trade in the twenty-first century. *Nature Geoscience*, *3*(3), 178–181. https://doi.org/10.1038/ngeo756

Di Clemente, R., Luengo-Oroz, M., Travizano, M., Xu, S., Vaitla, B., & González, M. C. (2018). Sequences of purchases in credit card data reveal lifestyles in urban populations. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-05690-8

Ding, C., Cao, X. (Jason), & Næss, P. (2018). 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. https://doi.org/10.1016/j.tra.2018.02.009

Esch, T., Heldens, W., Hirner, A., Keil, M., Marconcini, M., Roth, A., Zeidler, J., Dech, S., & Strano, E. (2017). Breaking new ground in mapping human settlements from space – The Global Urban Footprint. *ISPRS Journal of Photogrammetry and Remote Sensing*, *134*, 30–42. https://doi.org/10.1016/j.isprsjprs.2017.10.012

Ewing, R., & Cervero, R. (2017). "Does Compact Development Make People Drive Less?" The Answer Is Yes. *Journal of the American Planning Association*, *83*(1), 19–25. https://doi.org/10.1080/01944363.2016.1245112

Fei-Fei, L., Fergus, R., & Perona, P. (2006). One-shot learning of object categories. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *28*(4), 594–611. https://doi.org/10.1109/TPAMI.2006.79

Fletcher, S., Lickley, M., & Strzepek, K. (2019). Learning about climate change uncertainty enables flexible water infrastructure planning. *Nature Communications*, *10*(1), 1782. https://doi.org/10.1038/s41467-019-09677-x

Ford, J. D., Tilleard, S. E., Berrang-Ford, L., Araos, M., Biesbroek, R., Lesnikowski, A. C., MacDonald, G. K., Hsu, A., Chen, C., & Bizikova, L. (2016). Opinion: Big data has big potential for applications to climate change adaptation. *Proceedings of the National Academy of Sciences*, *113*(39), 10729–10732. https://doi.org/10.1073/pnas.1614023113

Gabe-Thomas, E., Walker, I., Verplanken, B., & Shaddick, G. (2016). 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. https://doi.org/10.1371/journal.pone.0158949

Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E. L., & Fei-Fei, L. (2017). Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States. *Proceedings of the National Academy of Sciences*, *114*(50), 13108–13113. https://doi.org/10.1073/pnas.1700035114

Geiß, C., Taubenböck, H., Wurm, M., Esch, T., Nast, M., Schillings, C., & Blaschke, T. (2011). Remote Sensing-Based Characterization of Settlement Structures for Assessing Local Potential of District Heat. *Remote Sensing*, *3*(7), 1447–1471. https://doi.org/10.3390/rs3071447

Gershenfeld, N., Samouhos, S., & Nordman, B. (2010). Intelligent Infrastructure for Energy Efficiency. *Science*, *327*(5969), 1086–1088. https://doi.org/10.1126/science.1174082

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. The MIT Press.

Google. (2019). *Environmental Insights Explorer*. https://insights.sustainability.google/

Gossart, C. (2015). Rebound Effects and ICT: A Review of the Literature. In L. M. Hilty & B. Aebischer (Eds.), *ICT Innovations for Sustainability* (pp. 435–448). Springer International Publishing.

Grubler, A., Wilson, C., Bento, N., Boza-Kiss, B., Krey, V., McCollum, D. L., Rao, N. D., Riahi, K., Rogelj, J., Stercke, S. D., Cullen, J., Frank, S., Fricko, O., Guo, F., Gidden, M., Havlík, P., Huppmann, D., Kiesewetter, G., Rafaj, P., … Valin, H. (2018). A low energy demand scenario for meeting the 1.5 °C target and sustainable development goals without negative emission technologies. *Nature Energy*, *3*(6), 515. https://doi.org/10.1038/s41560-018-0172-6

Haddaway, N. R., & Macura, B. (2018). The role of reporting standards in producing robust literature reviews. *Nature Climate Change*, *8*(6), 444. https://doi.org/10.1038/s41558-018-0180-3

Haklay, M. (2010). How Good is Volunteered Geographical Information? A Comparative Study of OpenStreetMap and Ordnance Survey Datasets. *Environment and Planning B: Planning and Design*, *37*(4), 682–703. https://doi.org/10.1068/b35097

Han, G., & Sohn, K. (2016). Activity imputation for trip-chains elicited from smart-card data using a continuous hidden Markov model. *Transportation Research Part B: Methodological*, *83*, 121–135. https://doi.org/10.1016/j.trb.2015.11.015

Han, Y., Oke, J., Hua, S., Zhou, J., Azevedo, C. L., Zegras, C., Ferreira, J., & Ben-Akiva, M. (2018). *Global Urban Typology Discovery with a Latent Class Choice Model*. Transportation Research Board 97th Annual MeetingTransportation Research Board. https://trid.trb.org/view/1497385

Harish, V. S. K. V., & Kumar, A. (2016). A review on modeling and simulation of building energy systems. *Renewable and Sustainable Energy Reviews*, *56*, 1272–1292. https://doi.org/10.1016/j.rser.2015.12.040

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd edition). Springer.

Hermwille, L., Obergassel, W., Ott, H. E., & Beuermann, C. (2017). UNFCCC before and after Paris – what's necessary for an effective climate regime? *Climate Policy*, *17*(2), 150–170. https://doi.org/10.1080/14693062.2015.1115231

Huseynov, S., & Palma, M. A. (2018). Does California's Low Carbon Fuel Standards reduce carbon dioxide emissions? *PLOS ONE*, *13*(9), e0203167. https://doi.org/10.1371/journal.pone.0203167

Hymel, K. M., Small, K. A., & Dender, K. V. (2010). Induced demand and rebound effects in road transport. *Transportation Research Part B: Methodological*, *44*(10), 1220–1241. https://doi.org/10.1016/j.trb.2010.02.007

Ilieva, R. T., & McPhearson, T. (2018). Social-media data for urban sustainability. *Nature Sustainability*, *1*(10), 553–565. https://doi.org/10.1038/s41893-018-0153-6

IPCC. (2014). *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.

IPCC. (2018). *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.)]*.

James, K. L., Randall, N. P., & Haddaway, N. R. (2016). A methodology for systematic mapping in environmental sciences. *Environmental Evidence*, *5*(1). https://doi.org/10.1186/s13750-016-0059-6

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, *353*(6301), 790–794. https://doi.org/10.1126/science.aaf7894

Jones, C., Hine, D. W., & Marks, A. D. G. (2017). The Future is Now: Reducing Psychological Distance to Increase Public Engagement with Climate Change. *Risk Analysis*, *37*(2), 331–341. https://doi.org/10.1111/risa.12601

Kaack, L. H., Chen, G. H., & Morgan, M. G. (2019). Truck Traffic Monitoring with Satellite Images. *Proceedings of the 2Nd ACM SIGCAS Conference on Computing and Sustainable Societies*, 155–164. https://doi.org/10.1145/3314344.3332480

Kazmi, H., Mehmood, F., Lodeweyckx, S., & Driesen, J. (2018). Gigawatt-hour scale savings on a budget of zero: Deep reinforcement learning based optimal control of hot water systems. *Energy*, *144*, 159–168. https://doi.org/10.1016/j.energy.2017.12.019

Kelly, J., & Knottenbelt, W. (2015). Neural NILM: Deep Neural Networks Applied to Energy Disaggregation. *Proceedings of the 2Nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, 55–64. https://doi.org/10.1145/2821650.2821672

Khayatian, F., Sarto, L., & Dall’O’, G. (2017). Building energy retrofit index for policy making and decision support at regional and national scales. *Applied Energy*, *206*, 1062–1075. https://doi.org/10.1016/j.apenergy.2017.08.237

Kontokosta, C. E., & Tull, C. (2017). A data-driven predictive model of city-scale energy use in buildings. *Applied Energy*, *197*, 303–317. https://doi.org/10.1016/j.apenergy.2017.04.005

Krause, J., Small, M. J., Haas, A., & Jaeger, C. C. (2016). An expert-based bayesian assessment of 2030 German new vehicle CO2 emissions and related costs. *Transport Policy*, *52*, 197–208. https://doi.org/10.1016/j.tranpol.2016.08.005

Lamb, W. F., Callaghan, M. W., Creutzig, F., Khosla, R., & Minx, J. C. (2018). The literature landscape on 1.5 °C climate change and cities. *Current Opinion in Environmental Sustainability*, *30*, 26–34. https://doi.org/10.1016/j.cosust.2018.02.008

Lamb, W. F., Creutzig, F., Callaghan, M. W., & Minx, J. C. (2019). Learning about urban climate solutions from case studies. *Nature Climate Change*, *9*(4), 279. https://doi.org/10.1038/s41558-019-0440-x

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, *401*(6755), 788–791. https://doi.org/10.1038/44565

Lehmann, A., & Gross, A. (2017). Towards Vehicle Emission Estimation from Smartphone Sensors. *2017 18th IEEE International Conference on Mobile Data Management (MDM)*, 154–163. https://doi.org/10.1109/MDM.2017.29

Lei, Y., Jia, F., Lin, J., Xing, S., & Ding, S. X. (2016). An Intelligent Fault Diagnosis Method Using Unsupervised Feature Learning Towards Mechanical Big Data. *IEEE Transactions on Industrial Electronics*, *63*(5), 3137–3147. https://doi.org/10.1109/TIE.2016.2519325

Liu, Z., Li, Z., Wu, K., & Li, M. (2018). Urban Traffic Prediction from Mobility Data Using Deep Learning. *IEEE Network*, *32*(4), 40–46. https://doi.org/10.1109/MNET.2018.1700411

Longo, M., Foiadelli, F., Franzo, S., Frattini, F., & Latilla, V. M. (2017). Towards the development of residential smart districts: The role of EVs. *2017 IEEE International Conference on Environment and Electrical Engineering and 2017 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I CPS Europe)*, 1–6. https://doi.org/10.1109/EEEIC.2017.7977552

Louf, R., & Barthelemy, M. (2014). A typology of street patterns. *Journal of The Royal Society Interface*, *11*(101), 20140924–20140924. https://doi.org/10.1098/rsif.2014.0924

Lucon, O., Ürge-Vorsatz, D., Zain Ahmed, A., Bertoldi, P., Cabeza, L. F., Eyre, N., Gadgil, A., Harvey, L. D. D., Jiang, Y., Liphoto, S., Mirasgedis, S., Murakami, S., Parikh, J., Pyke, C., & Vilariño, M. V. (2014). 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.*

Ma, J., & Cheng, J. C. P. (2017). Identification of the numerical patterns behind the leading counties in the U.S. local green building markets using data mining. *Journal of Cleaner Production*, *151*, 406–418. https://doi.org/10.1016/j.jclepro.2017.03.083

McDowell, N. G., Coops, N. C., Beck, P. S. A., Chambers, J. Q., Gangodagamage, C., Hicke, J. A., Huang, C., Kennedy, R., Krofcheck, D. J., Litvak, M., Meddens, A. J. H., Muss, J., Negrón-Juarez, R., Peng, C., Schwantes, A. M., Swenson, J. J., Vernon, L. J., Williams, A. P., Xu, C., … Allen, C. D. (2015). Global satellite monitoring of climate-induced vegetation disturbances. *Trends in Plant Science*, *20*(2), 114–123. https://doi.org/10.1016/j.tplants.2014.10.008

Microsoft. (2018). *US Building Footprints*. https://github.com/Microsoft/USBuildingFootprints

Mishra, N., Rohaninejad, M., Chen, X., & Abbeel, P. (2018). *A SIMPLE NEURAL ATTENTIVE META-LEARNER*. 17.

Mocanu, E., Nguyen, P. H., Kling, W. L., & Gibescu, M. (2016). Unsupervised energy prediction in a Smart Grid context using reinforcement cross-building transfer learning. *Energy and Buildings*, *116*, 646–655. https://doi.org/10.1016/j.enbuild.2016.01.030

Monajem, S., & Ekram Nosratian, F. (2015). 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. https://doi.org/10.1016/j.trd.2015.07.009

Montavon, G. (Ed.). (2012). *Neural networks: Tricks of the trade* (2. ed). Springer.

Montavon, G., Samek, W., & Müller, K.-R. (2018). Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, *73*, 1–15. https://doi.org/10.1016/j.dsp.2017.10.011

Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Definitions, methods, and applications in interpretable machine learning. *Proceedings of the National Academy of Sciences*, *116*(44), 22071–22080. https://doi.org/10.1073/pnas.1900654116

Nagendra, H., Bai, X., Brondizio, E. S., & Lwasa, S. (2018). The urban south and the predicament of global sustainability. *Nature Sustainability*, *1*(7), 341–349. https://doi.org/10.1038/s41893-018-0101-5

Naik, N., Raskar, R., & Hidalgo, C. A. (2016). Cities Are Physical Too: Using Computer Vision to Measure the Quality and Impact of Urban Appearance. *American Economic Review*, *106*(5), 128–132. https://doi.org/10.1257/aer.p20161030

Noy, K., & Givoni, M. (2018). Is 'smart mobility'sustainable? Examining the views and beliefs of transport's technological entrepreneurs. *Sustainability*, *10*(2), 422.

Nutkiewicz, A., Yang, Z., & Jain, R. K. (2018). 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. https://doi.org/10.1016/j.apenergy.2018.05.023

Papadopoulos, S., Bonczak, B., & Kontokosta, C. E. (2018). Pattern recognition in building energy performance over time using energy benchmarking data. *Applied Energy*, *221*, 576–586. https://doi.org/10.1016/j.apenergy.2018.03.079

Paredes, M., Hemberg, E., O'Reilly, U., & Zegras, C. (2017). Machine learning or discrete choice models for car ownership demand estimation and prediction? *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, 780–785. https://doi.org/10.1109/MTITS.2017.8005618

Petit-Boix, A., & Leipold, S. (2018). Circular economy in cities: Reviewing how environmental research aligns with local practices. *Journal of Cleaner Production*, *195*, 1270–1281. https://doi.org/10.1016/j.jclepro.2018.05.281

Rahim Taleqani, A., Hough, J., & Nygard, K. E. (2019). Public Opinion on Dockless Bike Sharing: A Machine Learning Approach. *Transportation Research Record*, 0361198119838982. https://doi.org/10.1177/0361198119838982

Rasp, S., Pritchard, M. S., & Gentine, P. (2018). Deep learning to represent subgrid processes in climate models. *Proceedings of the National Academy of Sciences*, *115*(39), 9684–9689. https://doi.org/10.1073/pnas.1810286115

Reckien, D., Salvia, M., Heidrich, O., Church, J. M., Pietrapertosa, F., De Gregorio-Hurtado, S., D'Alonzo, V., Foley, A., Simoes, S. G., Krkoška Lorencová, E., Orru, H., Orru, K., Wejs, A., Flacke, J., Olazabal, M., Geneletti, D., Feliu, E., Vasilie, S., Nador, C., … Dawson, R. (2018). How are cities planning to respond to climate change? Assessment of local climate plans from 885 cities in the EU-28. *Journal of Cleaner Production*, *191*, 207–219. https://doi.org/10.1016/j.jclepro.2018.03.220

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1

Ribeiro, M., Grolinger, K., ElYamany, H. F., Higashino, W. A., & Capretz, M. A. M. (2018). Transfer learning with seasonal and trend adjustment for cross-building energy forecasting. *Energy and Buildings*, *165*, 352–363. https://doi.org/10.1016/j.enbuild.2018.01.034

Rigas, E. S., Ramchurn, S. D., & Bassiliades, N. (2015). Managing Electric Vehicles in the Smart Grid Using Artificial Intelligence: A Survey. *IEEE Transactions on Intelligent Transportation Systems*, *16*(4), 1619–1635. https://doi.org/10.1109/TITS.2014.2376873

Robinson, C., Dilkina, B., Hubbs, J., Zhang, W., Guhathakurta, S., Brown, M. A., & Pendyala, R. M. (2017). Machine learning approaches for estimating commercial building energy consumption. *Applied Energy*, *208*, 889–904. https://doi.org/10.1016/j.apenergy.2017.09.060

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A., Maharaj, T., Sherwin, E. D., Mukkavilli, S. K., Kording, K. P., Gomes, C., Ng, A. Y., Hassabis, D., Platt, J. C., … Bengio, Y. (2019). Tackling Climate Change with Machine Learning. *ArXiv:1906.05433 [Cs, Stat]*. http://arxiv.org/abs/1906.05433

Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2019). Green AI. *ArXiv:1907.10597 [Cs, Stat]*. http://arxiv.org/abs/1907.10597

Seto, K. C., Dhakal, S., Bigio, A., Blanco, H., Delgado, G. C., Dewar, D., Huang, L., Inaba, A., Kansal, A., Lwasa, S., McMahon, J. E., Müller, D. B., Murakami, J., Nagendra, H., & Ramaswami, A. (2014). Human Settlements, Infrastructure and Spatial Planning. 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.*

Shaikh, P. H., Nor, N. B. M., Nallagownden, P., Elamvazuthi, I., & Ibrahim, T. (2014). A review on optimized control systems for building energy and comfort management of smart sustainable buildings. *Renewable and Sustainable Energy Reviews*, *34*, 409–429. https://doi.org/10.1016/j.rser.2014.03.027

Shan, Y., Guan, D., Hubacek, K., Zheng, B., Davis, S. J., Jia, L., Liu, J., Liu, Z., Fromer, N., Mi, Z., Meng, J., Deng, X., Li, Y., Lin, J., Schroeder, H., Weisz, H., & Schellnhuber, H. J. (2018). City-level climate change mitigation in China. *Science Advances*, *4*(6), eaaq0390. https://doi.org/10.1126/sciadv.aaq0390

Silva, B. N., Khan, M., & Han, K. (2018). Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities. *Sustainable Cities and Society*, *38*, 697–713. https://doi.org/10.1016/j.scs.2018.01.053

Silva, M. C., Horta, I. M., Leal, V., & Oliveira, V. (2017). A spatially-explicit methodological framework based on neural networks to assess the effect of urban form on energy demand. *Applied Energy*, *202*, 386–398. https://doi.org/10.1016/j.apenergy.2017.05.113

Silva, M., Leal, V., Oliveira, V., & Horta, I. M. (2018). A Scenario-Based Approach for Assessing the Energy Performance of Urban Development Pathways. *Sustainable Cities and Society*. https://doi.org/10.1016/j.scs.2018.01.028

Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and Policy Considerations for Deep Learning in NLP. *ArXiv:1906.02243 [Cs]*. http://arxiv.org/abs/1906.02243

Sturrock, H. J. W., Woolheater, K., Bennett, A. F., Andrade-Pacheco, R., & Midekisa, A. (2018). Predicting residential structures from open source remotely enumerated data using machine learning. *PLOS ONE*, *13*(9), e0204399. https://doi.org/10.1371/journal.pone.0204399

Tang, L., Xiong, C., & Zhang, L. (2018). Spatial Transferability of Neural Network Models in Travel Demand Modeling. *Journal of Computing in Civil Engineering*, *32*(3), 04018010. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000752

Tao, J., Zhou, Y., Wu, W., & Yu, L. (2014). Estimating carbon dioxide concentrations in urban areas from satellite imagery using Bayesian network. *2014 The Third International Conference on Agro-Geoinformatics*, 1–7. https://doi.org/10.1109/Agro-Geoinformatics.2014.6910674

Tao, T., Wang, J., & Cao, X. (2020). Exploring the non-linear associations between spatial attributes and walking distance to transit. *Journal of Transport Geography*, *82*, 102560. https://doi.org/10.1016/j.jtrangeo.2019.102560

Toole, J. L., Colak, S., Sturt, B., Alexander, L. P., Evsukoff, A., & González, M. C. (2015). The path most traveled: Travel demand estimation using big data resources. *Transportation Research Part C: Emerging Technologies*, *58*, 162–177. https://doi.org/10.1016/j.trc.2015.04.022

Tusting, L. S., Bisanzio, D., Alabaster, G., Cameron, E., Cibulskis, R., Davies, M., Flaxman, S., Gibson, H. S., Knudsen, J., Mbogo, C., Okumu, F. O., Seidlein, L. von, Weiss, D. J., Lindsay, S. W., Gething, P. W., & Bhatt, S. (2019). Mapping changes in housing in sub-Saharan Africa from 2000 to 2015. *Nature*, *568*(7752), 391. https://doi.org/10.1038/s41586-019-1050-5

Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, *10*(5), 988–999. https://doi.org/10.1109/72.788640

Wang, S., & Zhao, J. (2018). Framing Discrete Choice Model as Deep Neural Network with Utility Interpretation. *ArXiv:1810.10465 [Econ, q-Fin]*. http://arxiv.org/abs/1810.10465

Wen, J., Zhao, J., & Jaillet, P. (2017). Rebalancing shared mobility-on-demand systems: A reinforcement learning approach. *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 220–225. https://doi.org/10.1109/ITSC.2017.8317908

Wilson, C., Grubler, A., Gallagher, K. S., & Nemet, G. F. (2012). Marginalization of end-use technologies in energy innovation for climate protection. *Nature Climate Change*, *2*(11), 780–788. https://doi.org/10.1038/nclimate1576

Wolf, I., Schröder, T., Neumann, J., & de Haan, G. (2015). Changing minds about electric cars: An empirically grounded agent-based modeling approach. *Technological Forecasting and Social Change*, *94*, 269–285. https://doi.org/10.1016/j.techfore.2014.10.010

Wu, X., Tao, T., Cao, J., Fan, Y., & Ramaswami, A. (2019). Examining threshold effects of built environment elements on travel-related carbon-dioxide emissions. *Transportation Research Part D: Transport and Environment*, *75*, 1–12. https://doi.org/10.1016/j.trd.2019.08.018

Wurm, M., Schmitt, A., & Taubenbock, H. (2016). Building Types' Classification Using Shape-Based Features and Linear Discriminant Functions. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(5), 1901–1912. https://doi.org/10.1109/JSTARS.2015.2465131

Xu, C., Ji, J., & Liu, P. (2018). The station-free sharing bike demand forecasting with a deep learning approach and large-scale datasets. *Transportation Research Part C: Emerging Technologies*, *95*, 47–60. https://doi.org/10.1016/j.trc.2018.07.013

Yang, J., Shebalov, S., & Klabjan, D. (2018). Semi-Supervised Learning for Discrete Choice Models. *IEEE Transactions on Intelligent Transportation Systems*, 1–15. https://doi.org/10.1109/TITS.2018.2882847

Yau, K.-L. A., Qadir, J., Khoo, H. L., Ling, M. H., & Komisarczuk, P. (2017). A Survey on Reinforcement Learning Models and Algorithms for Traffic Signal Control. *ACM Computing Surveys*, *50*(3), 1–38. https://doi.org/10.1145/3068287

Yigitcanlar, T., Kamruzzaman, Md., Foth, M., Sabatini-Marques, J., da Costa, E., & Ioppolo, G. (2019). Can cities become smart without being sustainable? A systematic review of the literature. *Sustainable Cities and Society*, *45*, 348–365. https://doi.org/10.1016/j.scs.2018.11.033

Yu, J., Wang, Z., Majumdar, A., & Rajagopal, R. (2018). DeepSolar: A Machine Learning Framework to Efficiently Construct a Solar Deployment Database in the United States. *Joule*, *2*(12), 2605–2617. https://doi.org/10.1016/j.joule.2018.11.021

Zeng, W., Miwa, T., & Morikawa, T. (2017). 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. https://doi.org/10.1016/j.trd.2017.10.001

Zhang, W., Robinson, C., Guhathakurta, S., Garikapati, V. M., Dilkina, B., Brown, M. A., & Pendyala, R. M. (2018). Estimating residential energy consumption in metropolitan areas: A microsimulation approach. *Energy*, *155*, 162–173. https://doi.org/10.1016/j.energy.2018.04.161

Zhao, K., Tarkoma, S., Liu, S., & Vo, H. (2016). Urban human mobility data mining: An overview. *2016 IEEE International Conference on Big Data (Big Data)*, 1911–1920. https://doi.org/10.1109/BigData.2016.7840811

Zhou, S., Luccioni, A., Cosne, G., Bernstein, M. S., & Bengio, Y. (2020). Establishing an evaluation metric to quantify climate change image realism. *Machine Learning: Science and Technology*, *1*(2), 025005. https://doi.org/10.1088/2632-2153/ab7657

Zhu, X., Li, J., Liu, Z., Wang, S., & Yang, F. (2016). Learning Transportation Annotated Mobility Profiles from GPS Data for Context-Aware Mobile Services. *2016 IEEE International Conference on Services Computing (SCC)*, 475–482. https://doi.org/10.1109/SCC.2016.68

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine*, *5*(4), 8–36. https://doi.org/10.1109/MGRS.2017.2762307

Box. 1 | Machine learning in a nutshell.

Numerous families and traditions. High-level learning tasks include classification, regression, or probability density estimation. Main ML families include supervised learning (where the target value is known during training), unsupervised learning (no target value is known), or reinforcement learning (learning occurs through interactions with an environment), and sub-families include, for example for supervised learning, kernel methods or tree-based methods. For an introduction to ML, see (Hastie et al., 2009) and to DL specifically see (Goodfellow et al., 2016).

Simple to complex methods. On the simpler end, linear classifiers draw lines between groups of data points, and can be enough for simple tasks at large scales, e.g. classifying built-up areas on satellite imagery (Esch et al., 2017). In contrast, generative adversarial networks can achieve complex tasks like mimicking style of an image, for example they can generate a realistic image of a (climate-induced) flooded area from a picture taken with a normal weather (Zhou et al., 2020).

Data science life cycle. Machine learning tasks are embedded in the larger life cycle of a data science project, which includes for example data collection, problem formulation, and model maintenance when a model is used with new streams of data over time. There are many resources on generic aspects of the data cycle life cycle, see (Murdoch et al., 2019) for a particular focus on interpretable approaches. For practical tips on model training, refer for example to (Montavon, 2012).

Computational cost of ML. The largest deep neural networks train billions of parameters – for example transformers for natural language processing – and consume extreme amounts of energy (Strubell et al., 2019). However, in fact, most ML methods run within seconds to minutes on small personal computers. It is possible to contain deep neural networks' energy footprint, by explicitly considering the energy impact of training models (Schwartz et al., 2019).

Mining sustainability-relevant metrics. Which solutions for making a city more sustainable can one draw from staring at millions of points moving across a map? Big data often comes as a byproduct of a digital activity, rather than being collected for the purpose of a sustainability-related study. A key application area is to retrieve a signal of interest out of high-dimensional and noisy sources, e.g. from activity (Kelly & Knottenbelt, 2015) and textual data (Callaghan et al., 2020). Another application area is, when clean but numerous covariates are available, to identify groups of similar behaviors (Albert & Maasoumy, 2016) or contexts (F. Creutzig et al., 2015).

Generating data. By comparing the aerial images of a dense mixed-use neighborhood with that of a vast commercial area covered by parking lots, one may be able to identify which urban setting induces more or less car mobility. Such information is not directly present in the pixels of the images, but inferred by regularities known about the context of the picture. Supervised ML can learn to draw such inferences from large samples of paired input-target and such capabilities enable to automate the creation of fine-grained datasets on GHG emissions (Alam et al., 2018), energy uses (Kontokosta & Tull, 2017) and their spatial context.

Forecasting energy demands and supplies. Forward-looking predictions occupy a predominant role in climate research, and help decision making under uncertainty. ML is well suited for short- to mid-term predictions, which can offer concrete opportunities for emission savings. For example, *nowcasting* the production from PVs can enable to reduce unnecessary uses of backup power generation from more carbon-intensive carriers like gas. ML methods can learn patterns in time series (Mocanu et al., 2016), emulate physical models (Nutkiewicz et al., 2018), and be used conjointly in hybrid models.

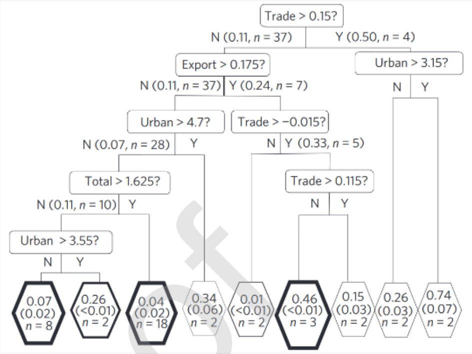
Controlling end-use systems. Various devises, e.g. heat pumps, do not require operate continuously to provide a given comfort, but are rarely adjusted, leading to unnecessary energy consumption. Beyond this simple example, much more complex energy-intensive systems offer opportunities for autonomous monitoring, for example the HVAC system of a ten-floor hospital. Key applications include controlling systems, detecting relevant events like faults (Afroz et al., 2018), and coordinating several processes to optimize the energy used to deliver a service (Kazmi et al., 2018).

Inferring causality. To best allocate limited resources to mitigation efforts, beyond pinpointing where problems are, one needs to ponder how effective possible interventions are likely to be – i.e. inferring their causal effect (Athey, 2017). New data sources offer the opportunity for high-resolution program evaluation, which can enable to target and tailor interventions, including information campaign, regulation, and retrofits. Main applications include estimating heterogeneous treatment effects (Beiser-McGrath & Bernauer, 2019) and counterfactual predictions (Burlig et al., 2017).

Box 2. Main usages of ML for climate change mitigation

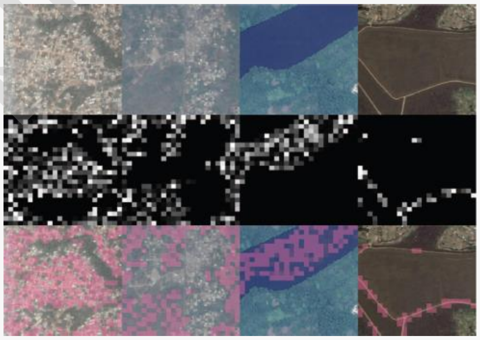
Question	Climate significance	ML algorithm
Climate change mitigation		
Carbon stock estimation (forests (Baccini et al., 2012))	Monitoring; Conservation; Land use policy	All
Mapping GHG emissions (cities (J. Tao et al., 2014))	Monitoring; policy design	SL, RF
Energy system and end uses (determinants (F. Creutzig et al., 2015))	Supporting energy policies, planning and transition pathways analyses	SL, Clu.
Behavioral aspects (attitudes (Beiser-McGrath & Huber, 2018), acceptance (Carr-Cornish et al., 2011), engagement (Jones et al., 2017))	Targeting policy implementation; Designing demand-side measures	DR, RL

Spatially-explicit example



Urbanization linked to deforestation (Ruth S. DeFries et al., 2010)

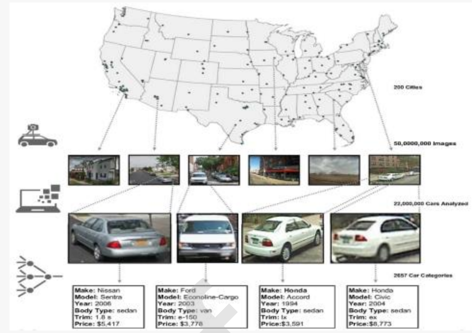
Question	Climate significance	ML algorithm
Remote sensing		
Socio-economic factors inference (poverty (Jean et al., 2016), demographics (Naik et al., 2016))	Monitoring; Planning; Establishing development aid needed	DR, SL
Stocks assessments (buildings (Esch et al., 2017), forest (R. S. DeFries et al., 2002))	Input for urban and energy planning	DL, UL
Geophysical or biochemical quantities (CO_2 concentration (J. Tao et al., 2014))	Monitoring; Parameters for models (climate, agriculture, etc.)	DL
Characterizing infrastructure (building type (Geiß et al., 2011; Sturrock et al., 2018; Wurm et al., 2016))	Input for urban and energy planning	DL



Socio-economic attributes derived from remote sensing (Jean et al., 2016)

Question	Climate significance	ML algorithm
Building energy use		
Contextual factors of energy use (occupancy (Shaikh et al., 2014), behaviors (Albert & Maasoumy, 2016; Gabe-Thomas et al., 2016))	Reducing demand due to consumer behavior	DTs, Clu.
Building energy use estimation (individual (Mocanu et al., 2016; Ribeiro et al., 2018), building stock (Kontokosta & Tull, 2017; Papadopoulos et al., 2018))	Estimating retrofit potential; Efficiency measures	NNs, RF
Typologies (Khayatian et al., 2017) **& solution transfer** (Ma & Cheng, 2017)	Accelerating the spread of best know-hows	Clu., DR, SL
Optimizing heating/cooling devices (HVAC, fault detection, etc. (Afroz et al., 2018; Harish & Kumar, 2016; Shaikh et al., 2014))	Reducing demand by efficiency measures	SL, RL

Parcel-level energy use prediction in NYC (Kontokosta & Tull, 2017)

Urban transportation			
Mobility pattern discovery from sensors or user data (Zhao et al., 2016)	Mitigating emissions from tourism, urban commuting	UL, DL	
Modal shift (discrete choice modelling (Paredes et al., 2017; Wang & Zhao, 2018; Yang et al., 2018))	Reducing carbon intensity	SL, RL	
Electric vehicle deployment (Longo et al., 2017; Rigas et al., 2015; Wolf et al., 2015)	Reducing carbon intensity	SL, RL	*Classifying cars to predict social attributes* (Gebru et al., 2017)
Role of urban form (Ding et al., 2018; Monajem & Ekram Nosratian, 2015)	Reducing transportation demand	SL, UL	

Abbreviations | GHG: GreenHouse Gas(es), **HVAC:** Heating, Ventilation, and Air Conditioning, **GPS:** Global Positioning System, **CDR:** Call Details Record
Methods | SL: Supervised Learning, **UL:** Unsupervised Learning, **DL:** Deep Learning, **RL:** Reinforcement Learning, **DR:** Dimensionality Reduction, **Clu.:** Clustering, **DTs:** Decision Trees, **NNs:** Neural Networks.

Table 2 | Examples of mitigation-relevant applications of machine learning methods for each field reviewed.