Data, Knowledge Practices, and Naturecultural Worlds: Vehicle Emissions in the Anthropocene

Lindsay Poirier
Smith College, lpoirier@smith.edu

CHAPTER 14

Data, Knowledge Practices, and Naturecultural Worlds: Vehicle Emissions in the Anthropocene

Observation

Lindsay Poirier

Standing at the heart of the US capital city amidst a sea of March for Science protesters on 22 April 2017, the rain had soaked through my jacket. For a protest with over 100,000 people in attendance, it was an oddly anti-social (though not dispassionate) event; looking towards the speaker stage, I could see little more than rows of soaked hoods and iPhones emerging above multi-coloured umbrellas, trying to snap photos of scientific superstars like Dr Michael Mann and Bill Nye. I was attending the march with my cousin Robert—a geologist—and his family. We both care deeply about the preservation of climate data and evidence-based decision making. Robert studies climate variability and sea-level change over thousands of years by examining rock sediments and fossil corals from deep in the Earth. As an anthropologist of data infrastructure and culture, prior to the event I had been getting involved to the extent that I could with the Environmental Data Governance Initiative (EDGI)—a group of researchers, practitioners, and activists convening to plan and execute guerrilla archiving efforts to safeguard environmental data from deletion by the Trump administration.

As a diverse array of speakers shared thoughts on the importance of advancing and advocating on behalf of science, it became clear that the stakes for generating and disseminating robust data about environmental health were

L. Poirier (✉)
Smith College, Northampton, MA, USA
e-mail: lpoirier@smith.edu

M. H. Bruun et al. (eds.), *The Palgrave Handbook of the Anthropology of Technology*, https://doi.org/10.1007/978-981-16-7084-8_14

high. I got emotional as Dr Mona Hanna-Attisha—the paediatrician who exposed heightened blood lead levels in children after they had been poisoned by lead-contaminated drinking water in Flint, MI, in 2015—described how ‘science spoke truth to power’. I was starstruck as Dr Mann detailed his work devising the hockey stick graph in the 1990s. I *also* felt unsettled by rhetoric that data alone should drive environmental regulation. It was not only the reappropriation of derogatory memes such as were seen in protest signs exclaiming ‘Grab ’em by the data’;[1] it was also the counterpoising of data against partisanship, of empiricism against bias. Encountering protest signs stating, ‘We want scientific data, not alternative facts’, I did not disagree but grimaced, a signal of how my own relationship with environmental governance data can be best characterised as ambivalent. Quantitative data about our anthropogenic world are indispensable; they are also at least partially mediated by cultural forces that prioritise profit and technological progress over environmental health equity.

Over a decade after Chris Anderson (2008) claimed that ‘the data deluge makes the scientific method obsolete’, there has been bubbling scepticism in many governance communities over the hype of big data in knowledge production and decision making.[2] In the environmental health domain—a domain that has been ‘informating’ (Fortun 2004) since the 1980s—widespread expert recognition of how uncertainty and estimation figure centrally in the measurement of natural worlds predates Anderson’s claims. In environmental health, researchers and policy makers time and again confront how anthropogenic history is being rewritten through new data and revised models, and how (much like the ecological systems constantly transforming under human feet, and around and through human bodies) the knowledge we have about our anthropogenic worlds is also constantly transforming in response to scientific advancements, political turnover, and industrial pressures. In constructing technologies for making sense of something as spectacular and incomprehensible as the Anthropocene, researchers and policy makers in environmental health have been forced to consider how to enforce environmental regulation when they cannot solely rely on a ‘trust in numbers’ (Porter 1996), and when they have to manage what they cannot (at least comprehensively) measure.

In this chapter, I detail various techno-cultural assemblages from which data collected to model and measure anthropogenic worlds emerge, arguing that data-based technologies both represent and co-produce the Anthropocene. Drawing on a case study of how vehicle emissions are measured and regulated in the US, I examine how US environmental health researchers and regulators grapple with the meaning of evidence and the basis for regulatory decisions as they confront the limits of automated data-collecting and modelling technologies. Finally, I meditate on the role of data-based technologies in mediating the environments we inhabit and the knowledge through which we perceive them.

Data Infrastructure and Knowledge Practices in the Anthropocene

Scholarship emerging at the intersection of science and technology studies (STS) and information studies (IS) has demonstrated that data infrastructure and data modelling (and the expertise and advocacy that emerge around both) are important nodes within the anthropogenic assemblages that shape history (Edwards 2017). The large-scale data infrastructures and models that enable us to visualise and make sense of the Anthropocene emerge from a series of localised practices of defining, classifying, and counting, wherein recognition, belonging, and uncertainty are continually being renegotiated. For instance, Bowker (1998) articulates how devising the neat boxes into which observations get classified involves negotiating the messiness of natural experience, navigating power struggles, and temporarily stabilising perpetually evolving worlds. Similarly, Martin and Lynch (2009) argue that counting, while seemingly trivial, does involve not only numerical operations but also discernments of what counts, calling for categorical judgements of identity and difference. Indeed, producing (ac)counts of natural observations is an embodied practice, demanding attunement to sensory experiences and eliciting emotions that style measurements and inscriptions (Calvillo 2018; Garnett 2016; Lorimer 2008).

Scholars in STS and IS have also characterised how practices of naming, classifying, and structuring data have become sites of collaboration, conflict, and politics. Standards for describing and storing data—often designed to network data across disciplinary, geographic, and temporal borders—emerge and transform in the face of capitalist, regulatory, and activist pressures (Bowker and Star 1999; Lampland and Star 2008; Ottinger 2010; Timmermans and Epstein 2010). This research has shown how, as data migrate across time and space, representations of anthropogenic worlds evolve alongside iterations in data semiotics (Bowker 2005). For example, Waterton (2002) shows how, as classification systems concerned with vegetation and natural habitats mutate in the face of controversy and instability, they come to reflect the dynamism and fluidity of the cultural systems within which they operate more than the contexts of their production.

Cultural practices of classifying and counting shape how identities form through data and how problems become both discernible and governable. Citing Hacking (2006), Kitchin and Lauriault (2014) argue that practices of counting and classifying are both contentious and consequential, ‘making up people’ and at least temporarily stabilising certain social and natural orders. Asdal (2008) demonstrates how nature-wholes are enacted—rendered real—through political methods of quantification and accounting designed to produce governable spaces. Kirksey (2015) demonstrates how species come into being through their entangled intra-actions with taxonomists and their technologies of classification: a dance of recognition, differentiation, and stabilisation on which many organisms depend to avert extinction. Similarly, Hepler-Smith (2019, p. 552) shows how toxic chemicals are identifiable in US regulatory structures through information practices that encode them on a molecule-by-molecule basis, a result of a 'molecular bureaucracy' in which 'law, administration, and politics meet empirical measurement and the material world'. Contemporary environmental problems receive public attention and enter debates through community engagement in environmental sense-making and the technologies they leverage. For example, Fortun et al. (2016) document how public pollution problems emerge as critical data designers couple skill in data visualisation with a hermeneutic sensibility to read the social, cultural, and political conditions that have eclipsed those problems. Other scholars have shown how research is left 'undone' because expert data systems—designed in ignorance of certain socio-cultural histories—preclude it (Frickel et al. 2010; Frickel and Vincent 2007; Nafus 2018).

Since different communities produce and consume data in different ways and with a diversity of ascribed meanings, it can be difficult to integrate data produced in different settings. Edwards et al. (2011) summarise this set of issues as 'data friction': the abrasive contact of the differing technologies, standards, and worldviews that represent and consume data. Scholars in this field have gone on to characterise how scientists address data friction in a cultural practice that involves attempts to cleanse data of their cultural influences. For instance, drawing on research in biology laboratories, Leonelli (2010) describes how, in order to facilitate the re-interpretation of data in new settings, data managers have had to learn to package data for travel, a practice that involves attempting to strip from data the personality and nuance of the contexts in which they were produced, meanwhile documenting their provenance so that others may re-contextualise the data for their own purposes. Since the contexts of data production and dissemination are often amorphous, power-laden, and unequal, Lampland and Star (2008) argue that translating data through various means of establishing common ground is always a political practice, one privileging certain semiotic orders over others.

As data move through complex and distributed socio-technical assemblages, frictional data practices call attention to their context-dependence and areas where they are incomplete or uncertain. Studies of the history and practices of data modelling have examined the ways in which such technologies mediate how knowledge is legitimised in the face of uncertainty. Oreskes (2000) argues that global data models have emerged to represent natural systems in instances when scientists lack complete access to the phenomena they are studying. Building upon this work, Knox (2018) ethnographically demonstrates how models serve as 'baseline data' against which messy and inconsistent observational data can be compared, enabling local-level administrative decision making in the face of missing data and other observational limits. However, Edwards (1999) shows that models themselves are also unstable, controversial, and constantly evolving, often in response to data frictions and inconsistencies with locally derived observational data. The movement of data across borders, scales, classification systems, and models troubles the local/global data binary. While Loukissas (2019) argues that 'all data are local'—that is all data are situated in a particular time and place—scholarship on practices and politics of data modelling demonstrates that data are also always more than local, the products of a 'cultural heterogenisation' of people, technologies, capital, media, and ideologies that propagate data flows, integrations, and disjuncture (Appadurai 1990). Through this scholarship we can see that, resonating with anthropologists' arguments that globalisation is not seamless and totalising (Ferguson 2006), global data infrastructuring is not moving us towards a mono-cultural data world. Data modelling and integration can divide the sciences and the representations they produce just as much as they bring them together.

Settings where scientists and policy makers grapple with the ambiguity and uncertainty woven through data practices and environmental sense-making are prominent sites for assessing shifting cultures of science and environmental regulation. Work in STS has documented how uncertainty can cripple scientific authority in policy making (Jasanoff 1987). For instance, Murphy (2006) has noted that, in the 1980s, the purposeful promotion of studies that furthered scientific uncertainty became a tool for anti-regulation at the US Environmental Protection Agency (EPA). However, other work in STS has shown that some scientific communities have responded to the limits of data, the complexity of environmental problems, and the extent of the unknown with 'humility and ambition'. Fortun and Fortun (2005), for instance, have described the culture of toxicology as shifting towards one that privileges experimentalism, wherein research is not necessarily designed to confirm what is already known but to generate new knowledge. In such communities, uncertainty is not seen as debilitating, and the knowledge produced through data systems and applied science is not the only knowledge useful in advancing regulation.

In summary, scholarship examining knowledge practices for characterising the natural world has demonstrated that technologies designed to measure and model the impact of human (and more-than-human) activity on earth systems are profoundly animated by the very human (and more-than-human) activities they attempt to measure. While dominant metaphors equate data with natural resources to be controlled or extracted, formulated in claims such as 'data is the new oil', or that we can be 'flooded with data' (Puschmann and Burgess 2014), scholars in critical data studies (e.g. see Gitelman 2013) often echo Bowker's (2000) claim that the term 'raw data is an oxymoron'. Suggestions that data could emerge from or return to a 'pure' or uncooked state mirror the blundering calls to return nature to its pure state. Data (and the worlds they inhabit) are always naturecultural (Haraway 2003; Subramaniam 2014). Anthropological attention to the materialities and mutability of data-producing technologies, along with the cultures and politics that shape them, can help to unpack how expertise operates, how knowledge is legitimised, and the way both have styled our experience of the Anthropocene.

Technologies for Counting and Estimating Vehicle Emissions in the US

What does it mean for environmental policies to be enacted 'based on' scientific data that are at times contested, always context-dependent, and modelled to measure things to which researchers do not have direct access? I explore this question in the following case study as I archaeologically trace select lineages of annual vehicle emissions estimates in the US to the moment when cars are first counted on federal highways. Rather than providing a holistic picture of how vehicle emissions estimates come into being, I ethnographically describe the data-collecting technologies involved in specific moments of their production in order to characterise environmental air quality regulation as a technologically mediated knowledge practice. In looking 'under the hood' at the configuration of a subset of technologies for measuring vehicle emissions, I elaborate on the diverse techno-cultural assemblages that animate systems of anthropogenic knowledge production and demonstrate the inextricable ties between the Anthropocene and the tools developed to understand it. Following Peter-Paul Verbeek's (2016) scholarship on 'technological mediation', I examine how technologies of data production mediate relationships between humans and the natural world, style everyday environments, and shape perception around what constitutes an empirical foundation for scientific claims.

Techno-Cultural Mediations of Emissions Standards

The US Clean Air Act, first signed into law in 1963 and updated several times since, was the first US policy to legislate air pollution control at the federal level.[3] A significant fortification of the federal government's role in air pollution control came with the 1970 amendments, which authorised the newly formed EPA to set National Ambient Air Quality Standards (NAAQS), and required each state to submit a periodical State Implementation Plan to the EPA outlining the policies and programmes they would enact to attain or maintain the standards.[4] NAAQS have been the subject of contentious debate and continuous evolution since the 1970s, with activists pressuring the EPA to strengthen regulation, corporations suing the EPA over the standards' stringency, and successive administrations revisiting the standards' review process, loosening or tightening the role of EPA staff in recommending policy options. Debates around the technical feasibility of implementing the standards have always been at the forefront of controversy. As a result, standards have emerged from a discursive space where technology, both available and speculative, tends to be positioned as a privileged signifier, in turn provoking changes in technological landscapes.

Responding to growing environmental concerns about smog, in 1970, Congress' amendments to the Clean Air Act mandated a 90% reduction in vehicle tailpipe emissions (including hydrocarbons, CO, and NOx) for passenger vehicles within five years (Gerard and Lave 2005). At the time, there had been no major improvements to the internal combustion engine in 20 years, and with little incentive for manufacturers to design technologies to reduce emissions, the new standard was considered 'technology-forcing' and designed to provoke innovation (Gerard and Lave 2005). For every car sold that did not meet the standards within the designated timeframe, automakers would face a $10,000 fine, double the average cost of a vehicle at the time (Gerard and Lave 2005). While the reductions were not achieved by the 1975 deadline, this regulatory pressure to innovate did lead to the introduction of the catalytic converter in 1975 and the three-way catalyst in 1981, both of which control tailpipe emissions through a chemical conversion process.

This technology-driving standard did not only spur innovation for emission control technologies but also helped motivate innovation around the chemicals that interfered with them. Until the 1970s, oil refiners had been adding lead to gasoline in order to raise the temperature and pressure at which engine knocking (or a premature ignition) occurs. At the time, lead in gasoline made up approximately 90% of airborne lead pollution, and there was growing concern about the threats the pollutant posed to public health (Stikkers 2002). Further, the combustion by-products of lead in gasoline can 'poison' catalytic converters by coating the metals responsible for converting exhaust chemicals (Stikkers 2002). As it was becoming increasingly clear that catalytic converters would be the primary means of reducing tailpipe emissions in the 1970s, the EPA began to introduce rules demanding the sale of unleaded gasoline. In culmination, the 1990 amendments to the Clean Air Act banned all leaded gasoline by 1996.

The year 1990 marked a historic change for the Clean Air Act for a number of reasons, but perhaps most notably for the way the US Congress further centred technology in standards-setting. When the Clean Air Act was first implemented in 1970, emissions standards for a number of pollutants were to be set based solely on what is requisite to protect public health (Bachmann 2007). However, as it became increasingly clear that scientific uncertainty about exposure risks would continuously immobilise the promulgation of standards, Congress pivoted the Act to require that standards be set based on the best currently available emission control technologies (or technologies available in the foreseeable future) (Flatt 2007; McCubbin 2003). Residual health risks would be assessed eight years after each standard was set. As a result, technical feasibility was increasingly privileged in standards-setting over the normative end goal of protecting health.

Approximately once a decade since the 1990 amendments, the EPA has introduced increasingly stringent standards on vehicle emissions. While in the 1990s emissions standards applied primarily to passenger vehicles, in the 2000s the same standards were extended to medium-duty passenger vehicles such as SUVs and passenger vans, and in the 2010s the standards were further extended to some heavier-duty vehicles such as cargo trucks (box vans). Successive standards also required reductions of sulphur and eventually ethanol in gasoline. Innovations such as hybrid vehicles and clean diesel engines emerged in the wake of these changes. While the EPA has been sued almost every step of the way, courts have reacted favourably to technology-forcing standards. Responding to a petition from automakers to block regulations requiring a 30% reduction in greenhouse gas emissions by 2016, Judge William Sessions III wrote in his ruling, 'History suggests that the ingenuity of the industry, once put in gear, responds admirably to most technological challenges' (Freeman 2007).

Regulators' discursive privileging of technology in connection with emissions standards has provoked shifts in technological, environmental, regulatory, and health landscapes in the US. These technology-driving standards have also, however, mediated the empirical foundation of scientific knowledge production around vehicle emissions. To calculate estimates of annual vehicle emissions, the EPA coordinates a number of scientific studies examining emissions from vehicles *meeting current standards*, along with how they fluctuate with changes in factors such as fuel additives, speed, and outdoor temperature conditions at start-up (US EPA Office of Transportation and Air Quality 2015). The EPA then estimates how many vehicles in the US meet these standards by analysing data about car sales and certifications in model years before, during, and after the phase-in of new standards. In other words, vehicle emissions standards, which emerge in a balance of what is technologically feasible and what can incentivise technological change, delimit how evidence regarding current and future vehicle emissions gets generated, in turn shaping how, when, and where environmental policies get enacted, the future prospects for transportation and energy industries, and the air we eventually breathe.

While this provides estimates of the emissions properties of certified vehicles, it does not offer information on when, where, the speed at which, and the duration of time these vehicles are actually operating. To calculate this, the EPA leverages data collection programmes that count vehicles on roads throughout the US, to which I turn next.

Vehicle-Counting Data Collection Technologies

One of the most important inputs for air emission models is the count of vehicles on federal highways each year, along with the number of miles they have travelled, a measure referred to as Vehicle Miles Travelled. Most US states have several hundred permanent traffic counters installed in or on roadways to produce continuous traffic counts. Inductive road loop counters, for instance, are coiled wires installed underneath or into the surface of roadways that can electromagnetically detect when a vehicle has passed over them. While inductive loop counts are widely considered to be accurate, the technology is expensive to install, even more expensive to maintain, and causes disruptions to traffic, roadway resurfacing, and utility repairs (The Vehicle Detector Clearinghouse 2007). Further, inductive loop counters are susceptible to fluctuations in weather conditions with freezing and thawing causing the loops to break. Thus, continuous traffic count programmes are mainly instituted to collect data regarding overall traffic trends across the state versus counts on *every* highway.

To ensure traffic is accounted for across every highway, state departments of transportation (DOTs) also manage short-term traffic count programmes. The most common method of producing these counts for a given highway is to hire consultants to lay a set of temporary pneumatic tubes on a road segment, a dangerous job that involves managing multiple lane closures (The Vehicle Detector Clearinghouse 2007). Each time a tyre passes over the rubber tubes, a signal is sent to a counter. Pneumatic road tubes are typically left in place for a few days, and then state DOTs calculate the Average Annual Daily Traffic for a roadway segment by averaging daily traffic on the days the tubes were placed and then multiplying the figure by 365. Multiplying the result by the length of the segment yields Vehicle Miles Travelled.

The daily traffic on a given highway can, however, vary drastically over the course of a week, month, or year. In order to plan for and amend these fluctuations in counts when determining annual average traffic, data collectors at state DOTs have become attuned to the cultural contexts of roadways. As a representative from a state traffic monitoring programme described in a 2016 interview with me,

> we collect 72-hour counts, but […] you can't count before Monday at 6 AM and [after] Friday at noon. It has to fall in there. So we consider Monday say to be a typical day, [and] Tuesday, Wednesday, Thursday, Friday morning to be typical. Friday afternoons, a lot of times, I look at the [highway] or something, and everybody's heading north for the weekend […] so we don't collect Friday afternoons.

Discerning what constitutes a 'typical' traffic day becomes more complicated when zooming out to the span of a year. In the interview, the representative discussed another situation where the DOT recognised numerical contingencies and considered options for normalising the count:

> We were looking at traffic counts [in] some real part of [the state] right near a college, … The two previous [counts] were in the 2000s for [average daily traffic]. And the current one we were looking at was 300 and something. But then you look at the dates, and the two previous ones were taken during college. The [third] one was taken in the summer when college wasn't in session. So if it was made in a rural area, colleges—they make a huge difference in the volume of traffic, right?

This anecdote demonstrates one of many ways in which traffic counters account for cultural patterns of migration and vehicle use. Unsuitable for snowy conditions, road tubes are typically not placed down in the winter in regions where there may be snow, and seasonal adjustments must be made so as not to overestimate winter traffic. Dips and spikes in the counts from year to year can signal legitimate changes in traffic conditions or, alternatively, faulty counting equipment. On busy highways, the rubber tubes can wear down, compromising the accuracy of the count (The Vehicle Detector Clearinghouse 2007), while in areas where there is often stop-and-go traffic, it is difficult for the system to distinguish one vehicle from the next (The Vehicle Detector Clearinghouse 2007). State DOTs will use overall trends identified by continuous traffic counters, along with cultural competency, to assess the quality of short-term highway counts and make necessary adjustments.

Notably, the data collection programmes responsible for producing a calculation of Vehicle Miles Travelled are not maintained by the EPA, but by the Federal Highway Administration (FHWA), and their initial purpose was not to measure vehicle emissions but to support transportation planning and to direct the allocation of federal highway aid (Federal Highway Administration, Office of Highway Policy Information 2019). The FHWA classifies counted vehicles into 13 categories (e.g. motorcycles, buses, and combination trucks), information that is important for pavement and bridge designers when considering how to maintain highway infrastructure. Responding to the affordances of road tubes and other vehicle classifying technologies in the 1980s, the boundaries dividing one vehicle class from the next are not determined visually, but by the vehicle classifiers' detecting and calculating the spacing between a vehicle's axles (Federal Highway Administration 2014). When the FHWA first proposed these classifications, the length of the wheelbase could readily differentiate a passenger vehicle from other two-axle, four-tyre vehicles (such as a pick-up truck or van). However, as SUVs and PT Cruisers gained popularity as passenger vehicles in the US, the logic dividing these categories became increasingly fuzzy. To better represent the data that they were actually collecting, in 2007 the FHWA changed the category of passenger vehicle to *light-duty, short wheelbase vehicle*, and the other two-axle, four-tyre category *to light-duty, long wheelbase vehicle*.[5]

The EPA does not regulate vehicle emissions based on wheelbase, however, but on the vehicle's gross weight. Having designed their emissions modelling systems around the inputs available through the FHWA, the EPA had to devise new algorithms for determining what percentage of long wheelbase vehicles were actually passenger carriers rather than commercial trucks (vans) in order to model emissions (US EPA, Office of Transportation and Air Quality 2016). One strategy for this involved analysing the composition of private and commercial truck (van) fleets in the US based on the results of the Census Bureau's Vehicle Inventory and Use Survey, a paper questionnaire mailed to and then collected from US-registered truck (van) owners every five years that gathers data about vehicle use. Megan Beardsley, team leader for the EPA's vehicle emissions model MOVES (to which I turn next), acknowledged this method to be limited since the survey was last taken in 2002.

Counting cars on US highways is not simply a numerical operation. It involves an array of networked people, institutions, calibrations, technologies, and data systems. Producing an 'accurate' count rarely involves relying on vehicle counting technologies alone, but also integrates cultural expertise attuned to when and why driving habits and vehicle purchases change, and is prepared to adjust counts accordingly. State DOTs must balance counting costs against numerical accuracy in the mix of diminishing infrastructure budgets, expensive equipment, and federal air quality regulations. Counting cars is a practice that poses risks to human safety while being designed to improve highway (and air) safety conditions, one that reacts to climatic fluctuations and detection limits as it becomes an input for knowledge systems measuring anthropogenic impacts on our air and climate. In other words, vehicle counts designed to measure and model human impacts on natural worlds are also a product of naturecultural worlds.

Technologies for Modelling Vehicle Emissions

In order to model annual pollution emissions from motor vehicles for their State Implementation Plans, all states (except California)[6] must leverage the EPA's Motor Vehicle Emissions Simulator (MOVES), a computer technology developed and maintained in the EPA's Office of Transportation and Air Quality. In preparing their plans, states input data about Vehicle Miles Travelled, weather conditions, and local demographics into MOVES. The system then calculates estimated emissions of criteria air pollutants, greenhouse gases, and air toxics based on data curated from millions of scientific emissions tests. MOVES is designed to predict future vehicle emissions conditions by ordering data about past and present conditions, along with estimations of how they might change.

The first version of MOVES was released in 1978 as MOBILE and has gone through at least ten major revisions since then 'to reflect improved data, changes in vehicle, engine, and emission control system technologies, changes in applicable regulations and emission standards and test procedures, and improved understanding of in-use emission levels and the factors that influence them' (US EPA 2016). For the MOVES team, designing an all-inclusive model of vehicle emissions, one that can comprehensively account for the array of natural (and more-than-natural) forces impacting emissions, is always a pursuit, that is, always open to further improvement. As improved strategies become available for estimating emissions from different mobile sources (such as from boats, lawnmowers, snowmobiles, and other agricultural equipment), the MOVES team seeks to incorporate the inputs into the modelling technology. As mechanisms become available for more accurately tracking the speed of vehicles on highways and the times highways are most populated (such as satellites and cell phone tracking), the MOVES team seeks to incorporate the inputs into the modelling technology. In October 2016, Megan Beardsley, team leader for the MOVES model, told me that the MOVES team maintains a 'huge laundry list of stuff that [they]'d like to the model to do better'. For example, as manufacturers began improving emissions from already warmed-up vehicles, the MOVES team began diverting their attention to producing better models of the time, place, and quantity of start-up emissions, and how they vary based on fluctuations in outdoor temperatures. In the early 2010s, when the EPA promulgated stricter standards for vehicle emissions and a new standard for gasoline sulphur, the MOVES team adjusted the models to account for cleaner vehicles and fuels on the road. Thus, the knowledge MOVES models are both cumulative and iterative alongside changes in technology and regulation. MOVES inputs are meticulously curated from the expansive corpus of factors impacting emissions. At any given moment, the evidence MOVES produces is acknowledged to be both robust and, to a certain degree, partial. While the MOVES team is judicious in incorporating the latest scientific research regarding vehicle emissions into the models, the selection of inputs is still mediated by what is currently possible to quantify, what is considered a priority for inclusion, and the capacity of the MOVES team (comprising about 20 individuals, many of whom as of 2016 do not work on the project full-time) to make the revisions.

Each time a revised version of the technology is applied in modelling, the quantified history of vehicle emissions in a given region slightly morphs, as does the understanding of present and future air quality conditions. Yet the pacing of revisions to the computer technology is tempered, not only by the timing of technological innovations and scientific advancements, but also by the bureaucratic pacing of research funding, peer review, and EPA rule-making processes. Years can pass from the introduction of new emissions standards until their benefits are understood through scientific research; the same applies to the period from when scientific data becomes available until a new version of the modelling technology is released. While the Clean Air Act requires State Implementation Plans to be prepared based on the most current information and models, the plans are sometimes prepared months to years before they are approved. This means that by the time the EPA approves a State Implementation Plan, there may already be swaths of new evidence repainting the picture of past, present, and future emissions in that state. To make progress towards emissions reductions, the EPA often must make governance decisions based on admittedly outdated estimations.

For example, in the early 2000s, the Sierra Club, one of the oldest and most influential environmental organisations in the US, filed a complaint with the US DC District Court regarding the EPA's conditional decision to approve components of Washington DC's State Implementation Plan (SIP). Since 1991, Washington DC had been classified as an area of 'serious', and at times even 'severe', non-attainment of NAAQS, requiring that it submit a Rate of Progress Plan with its State Implementation Plan that demonstrated 3% reductions in emissions each year leading up to their attainment deadline. The dispute was based, in part, on Washington DC's use of MOBILE5 (an earlier version of MOVES) in measuring the rate of progress towards attainment from 1996 to 1999. While MOBILE5 had been the most recent vehicle emission model available at the time the plan was created, just a month before the plan was submitted, MOBILE6 had become available. It took the EPA another year to approve the plan. The Sierra Club contested the EPA's decision to accept a

Rate of Progress Plan that had not been based on the latest data models. The Court responded:

> Indeed, as its name suggests, MOBILE5 is the fifth generation of this particular model; MOBILE6 is sixth. To require states to revise completed plans every time a new model is announced would lead to significant costs and potentially endless delays in the approval processes. EPA's decision to reject that course, and to accept the use of MOBILE5 in this case, was neither arbitrary nor capricious. (Sierra Club v USA EPA 2004, p. 19)

While this demonstrates the politics of the model's diachronicity, battles of evidence and rival claims over what constitutes sound experimental design can underlie the data inputs that inform MOVES' emissions calculations even in moments of temporary stability. For example, in 2015, the State of Kansas, the State of Nebraska, the Energy Future Coalition, and the Urban Air Initiative filed a suit with the DC Court of Appeals, asking them to review MOVES 2014 in light of a 'flawed fuel effects study' called EPAct/V2/E-89, conducted to test the effect of ethanol on particulate emissions (State of Kansas et al. v US EPA Brief for Respondents 2015). The petitioners argued that in the study the EPA used a method of blending ethanol with gasoline called a Match Blend, while most car manufacturers use a Splash Blend method. With a Splash Blend method, 10% ethanol is simply added to gasoline. In Match Blending, however, aromatic hydrocarbons are added to the mixture in order to ensure the gasoline meets a certain boiling point. These hydrocarbons, the petitioners argued, increase the toxicity of the mixture when emissions tests are run, while simply adding ethanol should reduce the toxicity of the mixture. In MOVES2014, ethanol gets modelled according to the results of this study, and increased ethanol volumes are shown to increase toxic emissions.

The petitioners argued that the inclusion of consultants from the petroleum industry in the design of EPAct/V2/E-89 had biased the study and that the use of the computer technology for modelling emissions would injure Kansas and Nebraska by categorising them as areas of non-attainment of air quality standards, depressing ethanol prices, and imposing detrimental effects on tax revenues. In their response, the EPA defended the study, justifying the team involved in the study design and endorsing their own expertise to carry out the research effectively. They also argued that the petition lacked standing because the MOVES technology was non-binding. While the states were required to use it in preparing their State Implementation Plans, MOVES was not a legislative tool, and the EPA could consider the quality of the model's outputs on a case-by-case basis. As they argued, 'applying the model in a particular agency action requires flexibility and the exercise of judgment' (State of Kansas et al. v US EPA Brief for Petitioners 2015, p. 12). The Court dismissed the petition for lack of standing.

In this response, we see how technology can mediate the meaning of evidence-based regulation and what emissions estimates are understood to be 'based on'. While the Clean Air Act specifies that attainment decisions are to be made based on the latest data models, this case reveals that what counts as empirical evidence can shift in response to a number of political, cultural, and judiciary pressures, particularly when data are contested and technological limits on representing and measuring future emissions are recognised. Still, Beardsley acknowledged in our interview that, for the MOVES team, 'there's a real desire to do the best stuff we can', because they always face the risk of a lawsuit when the legitimacy of MOVES as an evidence-producing technology is called into question. Notably, the possibility that the EPA might be sued over their work creates incentives for slowing down the modelling technology's development to allow for more careful study design and more thorough peer review, in turn widening the liminal gap between when new evidence becomes available and when it gets incorporated into the model. In other words, the knowledge about annual vehicle emissions produced through MOVES is both cumulative and co-constitutive of the conditions, cultures, and technologies of knowledge production.

Conclusion

The arguments presented in this case study echo decades of scholarship in the anthropology of technology arguing that naturecultural worlds are co-produced with technology (Downey and Dumit 1997). Data-based representations of nature emerge from situated and routinised human engagements with technologies of data collection and analysis, helping to render complex, pervasive, yet local issues like air pollution a national (and thus, in the US, federally regulable) concern. Decisions about how to calibrate data collection technologies, what inputs to include in data modelling technologies, and how to account for various sources of technological error are made in the face of political, economic, and cultural competencies and pressures, responding to the limits to knowing the world through technologically mediated apparatuses. Corporate interest, environmental activism, and human labour are thus all represented in data about natural worlds—interlaced through standards, measurements, and estimations as data flow between different people, technologies, and institutions. These technological configurations in turn mediate how the air we breathe, the worlds we inhabit, and the technologies available for mediating them come into being and evolve. Thus, for researchers and regulators, environmental decision making often demands critical judgement beyond what automated data collection and modelling technologies can produce and what can be quantitatively measured.

Ethnographically examining technologies of data collection and modelling 'under the hood' reveals far more than merely how they work and the phenomena they are designed to represent. Ethnographies of data-producing technologies provide a unique lens into complex cultures of knowledge production and environmental regulation, along with their technological mediation. They foreground how diverse stakeholders value (in multiple senses of the word) the environment, human health, and technological innovation, how regulators learn how to manage and mitigate pollution in light of acknowledged limitations to its measurement, and how the meaning of empirical evidence gets negotiated. In other words, the anthropology of technologies of data collection and modelling can highlight what makes natural worlds, and the data through which we present them, so human.

Acknowledgements Early versions of this chapter were developed and honed in the context of *The Asthma Files* (https://theasthmafiles.org/) research project, led by Kim Fortun and Mike Fortun.

Notes

1. This was a pithy reference to a 2005 videotape in which Donald Trump, while making a number of vulgar comments about women, told US television personality Billy Bush to 'Grab 'em by the pussy'.
2. I have lost count of the number of times when, in conversation with data analysts in municipal, state, and federal governments about the dangers of an overreliance on data systems and models, I have been surprised to find them nodding in agreement and referencing Cathy O'Neil's (2016) *Weapons of Math Destruction*. Part of this abmivalence has emerged from experience; many experts and policy makers can cite several examples where over-dependence on data-based systems of governance has prevented sound decision making.
3. Federal regulation of air pollution responded to two interrelated concerns: first, that as states compete for new jobs and industry, they have incentives to side-line environmental regulations; and second, that regardless of an individual state's degree of regulation, air does not know state boundaries.
4. Once approved by the EPA, the control strategies outlined in the plan became enforceable at both state and federal levels, and failure to comply with the plans would permit the federal government to take over enforcement.
5. For example, see https://www.fhwa.dot.gov/policyinformation/statistics/2007/vm1.cfm.
6. California, with the worst traffic conditions in the country, has much more stringent air quality regulations and is thus exempt from several federal policies.

References

Anderson, C. (2008, June 23). *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*. WIRED. http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory.

Appadurai, A. (1990). Disjuncture and Difference in the Global Cultural Economy. *Public Culture*, *2*, 1–23.

Aryn, Martin Michael, Lynch (2009). Counting Things and People: The Practices and Politics of Counting. *Social Problems 56*(2) 243–266. https://doi.org/10.1525/sp.2009.56.2.243.

Asdal, K. (2008). Enacting things through numbers: Taking nature into account/ing. *Geoforum*, *39*(1), 123–132. https://doi.org/10.1016/j.geoforum.2006.11.004.

Bachmann, J. (2007). Will the Circle Be Unbroken: A History of the U.S. National Ambient Air Quality Standards. *Journal of the Air & Waste Management Association*, *57*(6), 652–697. https://doi.org/10.3155/1047-3289.57.6.652.

Bowker, G. C. (1998). The Kindness of Strangers: Kinds and Politics in Classification Systems. *Library Trends, 47*(2), 255–292. https://www.ideals.illinois.edu/handle/2142/8207.

Bowker, G. C. (2005). *Memory Practices in the Sciences.* Cambridge, MA: The MIT Press.

Bowker, G. C., & Star, S. L. (1999). *Sorting Things Out: Classification and Its Consequences.* Cambridge, MA: The MIT Press.

Calvillo, N. (2018). Political airs: From monitoring to attuned sensing air pollution. *Social Studies of Science*, *48*(3), 372–388. https://doi.org/10.1177/0306312718784656.

Downey, G. L., & Dumit, J. (1997). *Cyborgs & citadels: Anthropological interventions in emerging sciences and technologies.* Santa Fe: School of American Research Press.

Edwards, P., Mayernik, M. S., Batcheller, A., Bowker, G., & Borgman, C. (2011). Science Friction: Data, Metadata, and Collaboration. *Social Studies of Science*, *41*(5), 667–690. https://doi.org/10.1177/0306312711413314.

Edwards, P. N. (2017). Knowledge infrastructures for the Anthropocene. *The Anthropocene Review*, *4*(1), 34–43. https://doi.org/10.1177/2053019616679854.

Evan, Hepler-Smith (2019). Molecular Bureaucracy: Toxicological Information and Environmental Protection. *Environmental History 24*(3) 534–560. https://doi.org/10.1093/envhis/emy134

Federal Highway Administration. (2014). *Verification, Refinement, and Applicability of Long-Term Pavement Performance Vehicle Classification Rules* (FHWA-HRT-13-091). https://www.fhwa.dot.gov/publications/research/infrastructure/pavements/ltpp/13091/13091.pdf. Accessed 15 October 2019.

Federal Highway Administration, Office of Highway Policy Information. (2019, April 1). *About Highway Performance Monitoring System (HPMS)*. https://www.fhwa.dot.gov/policyinformation/hpms/abouthpms.cfm. Accessed 15 October 2019.

Ferguson, J. (2006). *Global Shadows: Africa in the Neoliberal World Order*. Duke University Press.

Flatt, V. B. (2007). Gasping for Breath: The Administrative Flaws of the Federal Hazardous Air Pollutant Program. *Ecology Law Quarterly*, *34*(107), 107–174.

Fortun, K. (2004). From Bhopal to the Informating of Environmentalism: Risk Communication in Historical Perspective. *Osiris*, *19*, 283–296.

Fortun, K., & Fortun, M. (2005). Scientific Imaginaries and Ethical Plateaus in Contemporary U.S. Toxicology. *American Anthropologist*, *107*(1), 43–54. https://doi.org/10.1525/aa.2005.107.1.043.

Fortun, K., Poirier, L., Morgan, A., Costelloe-Kuehn, B., & Fortun, M. (2016). Pushback: Critical data designers and pollution politics. *Big Data & Society*, *3*(2), 1–14. https://doi.org/10.1177/2053951716668903.

Freeman, S. (2007, September 13). Carmakers Defeated On Emissions Rules. *Washington Post*. http://www.washingtonpost.com/wp-dyn/content/article/2007/09/12/AR2007091202391.html.

Frickel, S., Gibbon, S., Howard, J., Kempner, J., Ottinger, G., & Hess, D. J. (2010). Undone Science: Charting Social Movement and Civil Society Challenges to Research Agenda Setting. *Science, Technology & Human Values*, *35*(4), 444–473. https://doi.org/10.1177/0162243909345836.

Frickel, S., & Vincent, M. B. (2007). Hurricane Katrina, contamination, and the unintended organization of ignorance. *Technology in Society*, *29*(2), 181–188. https://doi.org/10.1016/j.techsoc.2007.01.007.

Garnett, E. (2016). Developing a feeling for error: Practices of monitoring and modelling air pollution data. *Big Data & Society*, *3*(2), 1–12. https://doi.org/10.1177/2053951716658061.

Geoffrey C., Bowker (2000). *Biodiversity Datadiversity. Social Studies of Science 30*(5) 643–683. https://doi.org/10.1177/030631200030005001

Gerard, D., & Lave, L. B. (2005). Implementing technology-forcing policies: The 1970 Clean Air Act Amendments and the introduction of advanced automotive emissions controls in the United States. *Technological Forecasting and Social Change*, *72*(7), 761–778. https://doi.org/10.1016/j.techfore.2004.08.003.

Gitelman L. (2013). *Raw Data Is an Oxymoron*. MIT Press.

Hacking I. (2006). Making Up People. *London Review of Books 28*(16).

Haraway, D. (2003). *The Companion Species Manifesto: Dogs, People, and Significant Otherness* (M. Begelke, Ed.). Chicago, IL: Prickly Paradigm Press.

Jasanoff, S. S. (1987). Contested Boundaries in Policy-Relevant Science. *Social Studies of Science*, *17*(2), 195–230. https://doi.org/10.1177/030631287017002001.

Kirksey, E. (2015). Species: A praxiographic study. *Journal of the Royal Anthropological Institute*, *21*(4), 758–780. https://doi.org/10.1111/1467-9655.12286.

Kitchin R. & Lauriault T. (2014). Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work. *Social Science Research Network*. Available at: https://papers.ssrn.com/abstract=2474112 (accessed 6 November 2017).

Knox, H. (2018). Baseless data?: Modelling, ethnography and the challenge of the anthropocene. In H. Knox & D. Nafus (Eds.), *Ethnography for a Data-Saturated World* (pp. 128–150). Manchester: Manchester University Press. https://doi.org/10.7765/9781526127600.00014.

Lampland, M., & Star, S. L. (Eds.) (2008). *Standards and Their Stories: How Quantifying, Classifying, and Formalizing Practices Shape Everyday Life* (1st ed.). Ithaca, NY: Cornell University Press.

Leonelli, S. (2010). Packaging Small Fact for Re-use: Databases in Model Organism Biology. In P. Howlett & M. S. Morgan (Eds.), *How Well Do Facts Travel?: The Dissemination of Reliable Knowledge* (pp. 325–348). Cambridge, England: Cambridge University Press.

Lorimer, J. (2008). Counting Corncrakes: The Affective Science of the UK Corncrake Census. *Social Studies of Science*, *38*(3), 377–405. https://doi.org/10.1177/0306312707084396.

Loukissas Y. A. (2019). *All Data Are Local: Thinking Critically in a Data-Driven Society*. Cambridge, Massachusetts: The MIT Press.

McCubbin, P. R. (2003). Amending the Clean Air Act to Establish Democratic Legitimacy for the Residual Risk Program. *Virginia Environmental Law Journal*, *22*(1), 1–52.

Murphy, M. (2006). *Sick Building Syndrome and the Problem of Uncertainty: Environmental Politics, Technoscience, and Women Workers*. Durham, NC: Duke University Press.

Nafus, D. (2018). Exploration or Algorithm? The Undone Science Before the Algorithms. *Cultural Anthropology*, *33*(3), 368–374. https://doi.org/10.14506/ca33.3.03.

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy.* (1st ed.). New York City: Crown.
Oreskes, N. (2000). Why Believe a Computer? Models, Measures, and Meaning in the Natural World. In J. Schneiderman (Ed.), *The Earth Around Us: Maintaining a Livable Planet.* W. H. Freeman & Co Ltd. https://doi.org/10.4324/9780429496653-8.
Ottinger, G. (2010). Buckets of Resistance: Standards and the Effectiveness of Citizen Science. *Science, Technology, & Human Values*, *35*(2), 244–270. https://doi.org/10.1177/0162243909337121.
Paul N., Edwards (1999). Global climate science uncertainty and politics: Data-laden models model-filtered data. *Science as Culture 8*(4) 437–472 https://doi.org/10.1080/09505439909526558
Porter, T. M. (1996). *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.
Puschmann, C., & Burgess, J. (2014). Metaphors of Big Data. *International Journal of Communication*, *8*(0), 1690–1709.
Sierra Club v US EPA Petitions for Review of Final Actions of the Environmental Protection Agency, DC Cir. No. 03-1084 (United States Court of Appeals, DC Circuit 2004).
State of Kansas, State of Nebraska, Energy Future Coalition, and Urban Air Initiative, Inc., v US EPA Brief for Petitioners, DC Cir. No. 14-1268 (United States Court of Appeals, DC Circuit 2015).
Stikkers, D. E. (2002). Octane and the environment. *Science of the Total Environment*, *299*(1–3), 37–56. https://doi.org/10.1016/S0048-9697(02)00271-1.
Subramaniam, B. (2014). *Ghost Stories for Darwin: The Science of Variation and the Politics of Diversity* (1st ed.). University of Illinois Press.
The Vehicle Detector Clearinghouse. (2007). *A Summary of Vehicle Detection and Surveillance Technologies used in Intelligent Transportation Systems.* Federal Highway Administration's (FHWA) Intelligent Transportation Systems Program Office. https://www.fhwa.dot.gov/policyinformation/pubs/vdstits2007/vdstits2007.pdf. Accessed 17 November 2020.
Timmermans, S., & Epstein, S. (2010). A World of Standards but not a Standard World: Toward a Sociology of Standards and Standardization. *Annual Review of Sociology*, *36*(1), 69–89. https://doi.org/10.1146/annurev.soc.012809.102629.
US EPA Office of Transportation and Air Quality. (2015). *Exhaust Emission Rates for Light-Duty On-road Vehicles in MOVES2014* (EPA-420-R-15-005). US EPA.
US EPA, Office of Transportation and Air Quality. (2016). *Population and Activity of On-road Vehicles in MOVES2014* (EPA-420-R-16-003). https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P100O7VJ.pdf.
Verbeek, P.-P. (2016). Toward a Theory of Technological Mediation: A Program for Postphenomenological Research. In J. Kyrre Berg, O. Friis, & R. P. Crease (Eds.), *Technoscience and Postphenomenology: The Manhattan Papers* (pp. 189–204). Lexington Books. https://research.utwente.nl/en/publications/toward-a-theory-of-technological-mediation-a-program-for-postphen.
Waterton, C. (2002). From Field to Fantasy: Classifying Nature, Constructing Europe. *Social Studies of Science*, *32*(2), 177–204. https://doi.org/10.1177/0306312702032002001.