Saving the Data

Greg Lusk

Abstract: Three decades ago, James Bogen and James Woodward (1988) argued against the possibility and usefulness of scientific explanations of data. They developed a picture of scientific reasoning where stable phenomena were identified via data without much input from theory. Rather than explain data, theories 'save the phenomena'. In contrast, I argue that there are good reasons to explain data, and the practice of science reveals attempts to do so. I demonstrate that algorithms employed to address inverse problems in remote-sensing applications should be understood as attempts to identify phenomena by explaining the data. Thus, this paper furthers understanding of data-to-phenomena reasoning in science, and demonstrates theory may play a more central role in phenomena identification than previously recognized.

1. Introduction

Thirty years ago, James Bogen and James Woodward ([1988]) argued against a widely held view that scientific theories aim to explain or predict facts about observables. To make their argument, they introduced a distinction between data, phenomena, and theory, claiming that theories explain unobservable phenomena which are identified via observable data. The identification procedures they expounded were 'bottom-up' in character. Such procedures turn on the reliability of data, which can be demonstrated through empirical replicability and error control, as well as statistical argumentation. On their view, though theory may furnish motivation to look for a particular phenomenon, and provide the vocabulary to describe it, theoretical explanations of data play little-to-no role in the identification of phenomena. In other words, explanations save phenomena, not data.

Bogen and Woodward's account spurred debates about the ontological status of phenomena (e.g. McAllister [2007]), the value of the data-phenomena distinction itself (e.g. Glymour [2000]), as well as the roles theory can have in scientific reasoning (e.g. Schindler [2007]). However, Woodward ([2010], [2011]) continues to defend the central position of the 1988 paper, that the identification of phenomena is relatively theory-free and does not involve theoretical predictions, derivations, or explanations of data. He laments that critics assume, often without argument, that researchers reasoning from data to phenomena are, in principle, able to detail the production of data. While he regrets putting the original argument in somewhat vague and sometimes exceptionless terms, he claims that commentators have failed to show that derivations or explanations of data 'play some functionally useful role in scientific reasoning' as 'actually practiced' (Woodward [2011], p. 168). According to Woodward, there is no rationale for even attempting to provide such explanations, as they would be largely useless for identifying phenomena.

In what follows, I argue that theoretical explanations of data are not only possible, but part of a viable strategy for phenomena identification used in scientific practice. That scientists can and sometimes do aim to save the data demonstrates theory can be more central to phenomena identification than has previously been admitted. To ground my claims, I distinguish two arguments against theoretical explanations of data: a descriptive argument that claims explanations of data are difficult, if not impossible, to provide, and a normative argument that suggests scientists should not employ such explanations. I then explore a case study that serves as a counter example to both arguments: examining the practice of remote sensing atmospheric trace gases, I show how practicing scientists aim to explain data in order to overcome inverse problems when quantifying phenomena. Without attempting these explanations of data, remote sensing of this kind would not be possible. This example, which relies on maximum a posteriori estimation, reveals a functional use for explanations of data as part of a strategy frequently employed in the Earth and atmospheric sciences for addressing state-dependent sources of systematic error when identifying phenomena. Thus, I show that both of the arguments against explanations of data can be overcome, and that such explanations can play some functionally useful role in scientific practice. This is not, however, a complete rejection of Bogen and Woodward's view. Rather, the analysis presented here furthers understanding of the distinctive roles theory can play in phenomena identification, ultimately enriching Bogen and Woodward's account.

This paper is divided into four subsequent sections. The first, 'Saving the Phenomena', introduces the data-phenomena-theory distinction by providing a detailed explication of the descriptive and normative arguments against the desirability of deriving or explaining data. The next section, 'Saving the Data', introduces atmospheric retrieval using Fourier Transform Infrared Spectroscopy (FTIR) as a case study, and demonstrates why scientists might attempt to explain data. The third section considers some objections and provides responses. The last offers concluding remarks.

2. Saving the Phenomena

In Bogen and Woodward's framework, data are observable public records produced by experiments or measurements (Bogen and Woodward [1998]; Woodward [2011]). Examples of data (here 'data' indicates many independent datum, not a data set and its collective properties) include one of Eddington's 1919 photographs to detect the deflection of starlight during an eclipse or the report of a particular thermometer's indication. Data are characteristically produced by multiple causes, making them 'idiosyncratic to particular experimental contexts, and typically cannot occur outside of those contexts' (Bogen and Woodward [1988], p. 317).

Phenomena—in contrast to data—are often unobservable but repeatable effects or processes that are stable enough, in principle if not in fact, to reoccur under different conditions (Bogen and Woodward [1988]; Woodward [2011]). Phenomena include the actual deflection of star light by the sun during an eclipse and the actual temperature of an object being measured. Phenomena are the result of fewer causes than data. Bogen and Woodward's emphasis is on data and phenomena, and hence, 'theory' is not conceptually filled out. All that is required for our purposes is that theory includes the fundamental equations or statements commonly accepted as playing an explanatory role in science, for example, by describing things like the deflection of light or electromagnetic and gravitational radiation.

Data and phenomena differ in their ontological and epistemic statuses. Ontologically, phenomena are investigator-independent features of the natural world, whereas data are merely the product of a coincidental confluence of factors. Epistemically, data serve as evidence for phenomena: phenomena are typically identified via 'bottom-up' reasoning using data and additional assumptions. Data are thus epistemically prior to phenomena, but have no direct relation to theory. Phenomena, on the other hand, serve as evidence for, and are explained 'top-down' by, theory. Unlike data, phenomena are inferentially justified, and therefore challenges to the existence of phenomena differ from challenges to data. Challenges to claims about a phenomenon assert that erroneous inferences were employed in its identification, whereas challenges to data as such proceed by claiming the public records do not actually depict an instrument's report.

Bogen and Woodward emphasize that the scientist's task is to establish data as reliable evidence for phenomena by identifying and compensating for possible sources of error in their production. For instance, the outcome of any one measurement depends not only on the object measured, but also on the characteristics of the measuring device, the environmental conditions, and other background conditions. Any observed datum d_i can be seen as the result of a function f not only of the phenomenon M of interest, but also sources of error u: $d_i = f(M, u_i)$ (Woodward [2011], p. 167). Assuming the phenomenon is stable and not inherently stochastic, these factors u_i are responsible for differences among data. As Woodard points out, 'the problem the researcher faces is not one of describing the d_i or of providing derivations or predictions of the d_i,' but rather 'her goal is to use the d_i to make an inference about' phenomena (Woodward [2010], p. 795).

To obtain information about phenomena from data, scientists employ strategies that aim to block, remove, or compensate for the influence of u_i. Woodward ([1989]) describes strategies that control for such influences, including the physical isolation of measuring devices, calibration, exploitation of (non-)uniformity, and statistical processing, but not derivation or explanation. For example, scientists might isolate the environment in which data is gathered, shielding it from unwanted error-inducing effects. They may also exploit the uniformity of certain unwanted influences by estimating their effects and altering the data accordingly, or may rule out non-uniform unwanted influences by running duplicate experiments at distant locations when a phenomenon is expected to have a simultaneous effect (as was done recently to detect gravity waves). Additionally, scientists make assumptions that lead to a certain statistical treatment of the data. If systematic error is assumed to be negligible, for example, then it might be safe to assume f represents an additive relationship (i.e., $d_i = M + u_i$) and the u_i are normally distributed; thus, the mean value of a sufficient number of measured d_i will capture the phenomenon with a quantifiable degree of uncertainty. These strategies can be conjoined, and their effectiveness evaluated using calibration or inter-comparison procedures. Bogen and Woodward stress that such strategies rely on empirical assumptions to make reliable inferences about phenomena.

Bogen and Woodward claim that typical reasoning from data to phenomena does not involve explaining data: phenomena are natural candidates for theoretical explanation, whereas data are not. For them, explanation consists of a systematic derivation, 'the provision of a principled, non- ad hoc derivation from the basic laws' of theory (Woodward [2011], p. 167) or from principles of great generality (Woodward [2010]), that specify how an outcome depends on the particular factors that produced it. This systematic-derivative form of explanation is distinguished from non-systematic, or 'singular causal,' explanations, which make no appeal to formalized causal generalizations. Singular causal explanations merely cite some causal factor that plays 'a role in the production of an outcome' but does not specify in any 'quantitative way how the outcome depends on this factor or which other factors influence the out-come' (Woodward [2011], p. 166). It is systematic-derivative explanations of data that purportedly play no role in phenomena identification. Saving the data via systematic-derivative explanation therefore requires: 1) that a piece of data serve as the explanandum, 2) that a derivation of that data be given from general principles, and 3) that the derivation be systematic, in part by demonstrating how the explanandum depends on the idiosyncratic influences that helped produce it. As we will see, Bogen and Woodward will argue that there is no reason to explain a datum, and systematic derivations are nearly impossible to provide. Following Bogen and Woodward's lead, I will use 'systematic derivation' and 'explanation' synonymously.

Woodward claims the aim of systematic-derivative explanation is to provide some degree of unification, where unification consists of 'reference to factors, generalizations, or mechanisms which can figure in the explanation of a range of different phenomena' ([1989], p. 400). Systematicity and unification play two independent functions in the account of data-to-phenomena reasoning: the demand for systematicity undermines scientists' ability to demonstrate dependency between data and theory, while the demand for unification demonstrates that such derivations (if they could be given) would not support the goals of science. I consider these functions separately, addressing the former within what I call the descriptive argument, and the latter within the normative argument. I aim to show that systematically deriving data can be a successful strategy for reasoning to phenomena, and thus, theory may sometimes play a central role in phenomena identification.

I am not, however, suggesting there is a normative demand that scientists provide theoretical derivations of data; Bogen and Woodward rightly point out that derivations of data are not possible in many cases. Consider the Copenhagen interpretation of quantum mechanics (QM): it is only possible to predict the probability with which a particular outcome (e.g., spin state) will be found. This probability represents the phenomenon, and the theory can go no further; by QM's own lights there is no way to predict any particular measurement outcome. Derivations of data are also unnecessary in many cases: Bogen and Woodward convincingly show that the melting point of lead can be established through repeated measurement and statistical inference, with no derivation of data required. Still, the question remains whether theory-based derivations of data can play any significant role in data-to-phenomena reasoning.

2.1 The Descriptive Argument

What I call Bogen and Woodward's 'descriptive argument' claims systematic derivations of data would be too complex to provide in practice, and thus play no role in data to phenomena reasoning. This argument rests on the examination of important historical cases (e.g. the discovery of weak neutral currents), demonstrating that *representational inadequacy, intractability,* and a *lack of contextual knowledge* make derivations of data difficult if not impossible to provide.

In order to deduce the character of a datum from theoretical principles, scientists would need to explicitly represent many of the details of data production, perhaps including the environment, phenomena, error causing factors, and instruments used to produce the data. Representing these details requires knowledge of the causal relations that specify how phenomena and error causing factors interact with instruments, experimenters, etc. These factors need to be explicitly detailed in any equations or premises from which a derivation will proceed. But as Bogen and Woodward indicate, scientific theories do not typically contain information about sources of error or instruments employed in an investigation, and are thus representationally inadequate for formulating generalizations that could support derivations of data. Additional knowledge is required to supplement theory, but such knowledge is often not available to scientists.

Of course, if one did have the resources to construct equations that modeled the entire data production process, they would need to be solved. Bogen and Woodward claim that explicitly representing the concrete details of data production often makes a derivation of the data computationally intractable. Armed with an accurate representation of data production, scientists could have trouble solving for d_i under the relevant conditions, thus preventing detailed descriptions of data production from being any use to scientists.

Even if scientists knew how to describe the complexities of the investigative environment, and the resulting description was tractable, Bogen and Woodward point out that a derivation may require contextual knowledge regarding the system under investigation that scientists often lack. A datum could be the result of a particular phenomenon or could be the result of only other factors that happen to resemble that phenomenon in that investigative context. Differentiating between the two requires substantial knowledge about the actual conditions of data production, possibly including values that represent the phenomena, sources of error, or initial conditions. These quantities are typically unknown to scientists before an investigation is performed. Assuming that initial conditions or confounders may change quickly in an investigative context, quantifying them accurately presents a substantial obstacle to reasoning via derivations of data.

Ioannis Votsis ([2010]) has suggested that the descriptive argument detailed above fails to establish that derivations of data are incapable of potentially supporting explanations. Votsis believes that Bogen and Woodward have forgotten the Duhemian lesson that theory assessment always requires auxiliary hypotheses. He argues that scientists reasoning from data to phenomena have an auxiliary hypothesis like the following in mind: 'evidence or phenomena x implies observation y' (Votsis [2010], p. 272). Such an auxiliary hypothesis can be used to create a disjunction guaranteed to contain the value of the datum, with the proper disjunct selected via statistical properties of the data set and the phenomenon. Since the phenomenon is explained by theory, so are the subsequently derived data. Votsis shows that this is the case for Bogen and Woodward's example concerning the identification of the melting point of lead (a phenomenon).

Vostis' scheme for deducing data from phenomena has two deficiencies. First, as Woodward ([2011]) points out, Vostis is concerned with the in-principle derivability of data from theory, not the in-practice possibility and use of such derivations.[1] Secondly, the form of Vostis' derivation does not provide insight into why particular data points occur when they do; the derivation is based on assumptions of the experimenter rather than a model of the physical production of data. While Vostis provides valuable insight into scientific reasoning, he does not demonstrate how the descriptive argument can be overcome.

[1] Vostis ([2011]) points out that Bogen and Woodward ([2003]) acknowledge it is no objection to the kind of reconstructionist view he offers to merely claim that it does not mirror actual scientific reasoning, as such a claim is undisputed.

The descriptive argument attempts to establish that scientists are not in a position to provide derivations of data. Such derivations are difficult, if not impossible, and since scientists cannot give them, they serve no function in scientific practice.

2.2 The Normative Argument

Having identified the above obstacles, Bogen and Woodward suggest that there are normative reasons to forgo derivations of data, claiming that 'there will be little scientific point' in deriving data 'even if this were possible' ([1988], p. 323).[2] Bogen and Woodward invoke unification as the aim of systematic derivation – systematic derivations achieve understanding by revealing common patterns in what originally seemed like distinct situations – to show derivations of data lack scientific value. For them, 'good explanations in science should unify and connect a range of different kinds of explananda' ([1988], p. 325). If unification secures generality, derivations of data are left lacking. As already noted, data are idiosyncratically produced in specific experimental contexts that are influenced by known and unknown error inducing factors. Scientific explanations will need to account for these idiosyncrasies. As the argument goes, accounting for idiosyncrasies does not produce derivations with a high degree of unification. Thus, there is no reason to provide derivations of data, because as explanations, they would not be very good. Derivations of phenomena, however, avoid having to tell independent local stories, and instead focus on what is stable across different situations. This permits accounting for a wide range of explananda by appealing to general principles.

Though desirable, seeking unification or generality does not provide a normative argument that would undermine the desirability of derivations of data: there are multiple goals of science, and some might be furthered by deriving data. As Bogen and Woodward continually stress, scientists are often interested in calibrating or testing their investigative apparatuses or models to determine their reliability as data producers; these are processes that might benefit from systematic derivations of data. Predictions of data from theory, as well the attribution of differences between theoretical predictions and actual measurements, would be valuable (though not always necessary for) establishing the reliability of measuring devices. Tal ([2009]) points to an example of when such predictions are actually employed. He argues that scientists sometimes establish the reliability of investigative results by locating unique signatures in the data that had previously been predicted by computer simulations of the procedure employed. This use provides a straightforward reason to value the derivation of data; after all, determinations of reliability are crucial to Bogen and Woodward's account of scientific reasoning. To claim that theoretical unification or generality are goals that derivations of data do not fulfill, does not show that such derivations are prevented from playing a useful role in data-to-phenomena reasoning.

[2] Interestingly, there is much in common between Bogen and Woodward's position and that of William Alston ([1971]) when it comes to the explanation of data and the need for generality.

I claim that a stronger normative argument is available. Call the systematic derivation of data a 'd-strategy.' Call any other strategy a 'non-d-strategy.' Non-d-strategies identify phenomena by demonstrating that data is reliable evidence via the processes of managing error that Bogen and Woodward suggest (e.g. physical manipulation of instruments, data reduction, and statistical compensation). The descriptive argument suggests that d-strategies require precise information regarding data production not required by non-d-strategies. Therefore d-strategies require resources, either epistemic (e.g. knowledge of causal processes and conditions) or practical (e.g. computing time for complex calculations), that exceed those required by non-d-strategies. These considerations can ground a stronger normative argument:

> P1) Any scientific task that can be accomplished by d-strategies can be accomplished by non-d-strategies.
>
> P2) D-strategies require more resources to deploy than non-d-strategies.
>
> P3) Scientists should not invest more resources than necessary to complete a scientific task.
>
> C) Since d-strategies accomplish the same scientific tasks as non-d-strategies (P1), but require more resources (P2), then (P3) scientists should not employ them.

P1 is assumed for the sake of argument. The descriptive argument suggests it may be false, but only because d-strategies may not be viable. Something like P1 is suggested when Bogen and Woodward claim that derivations of data are not needed when establishing the reliability of data. The descriptive argument provides evidence that P2 is true. It suggests that derivations of data would require significant resources to overcome the obstacles identified. Furthermore, scientists routinely perform the task of phenomena identification with extant resources, but those resources are insufficient to employ d-strategies. P3 is supposed to capture the intuition that scientists should be efficient in their investigations. If this were not assumed, there would be no reason to prohibit scientists from selecting d-strategies over non-d-strategies for some tasks simply to satisfy their own scientific curiosity or seek the truth about a particular datum. We might attach a clause to P3 permitting the expenditure of extra resources if certain benefits are expected, but such issues need not trouble us. This normative argument shows why explanations of data should be avoided.

Woodward might have an argument like this in mind when he clarifies the 1988 position: 'there is often no obvious scientific rationale or motivation for attempting to provide detailed systematic explanations of data' because in demonstrating that data serve as evidence for phenomena 'researchers do not need to exhibit systematic explanations/derivations of data' (Woodward [2011], p.

167). Extra resources required to derive data would be wasted if the same ends can be achieved by more efficient means.

Articulating the normative argument in this way is advantageous because it clarifies potential defeaters. Either a d-strategy that is successful where non-d-strategies fail, or a d-strategy that uses less resources than non-d-strategies in achieving the same ends, would defeat the argument. The next section considers a case that challenges both the normative and descriptive arguments.

3. Saving the Data: The Case of Trace Gas Profile Retrieval

For atmospheric scientists, establishing the existence of certain phenomena—namely concentrations of atmospheric gases and their patterns of movement—is important for understanding air quality, atmospheric chemical reactivity, and the effects of global warming. One popular method for long-term atmospheric monitoring uses ground-based instruments, in particular spectrometers and Light Detection and Ranging instruments (LIDAR). 'Remote sensing,' as this method is often called, involves quantifying properties of the atmosphere via measurement of their electromagnetic characteristics. Here, I focus on the use of Fourier Transform Infrared Spectroscopy (FTIR) to quantify atmospheric trace gases. FTIR employs a Fourier Transform Spectrometer (FTS) that detects electromagnetic radiation in the infrared spectrum in order to infer a vertical profile of trace gases.

The FTS provides the core data to identify trace gas phenomena. A ground based FTS detects infrared radiation that reaches the surface after interacting with the atmosphere's chemical constitutes. In this interaction, trace gases absorb some of the downwelling radiation at particular wave numbers. The FTS produces records of these interactions in the form of spectra that contain absorption bands—black gaps at particular wavenumbers where trace gases have absorbed the IR radiation— that serve as evidence for gas concentrations in a column above the device. FTSs are often similar in design to Michelson interferometers: they split an incoming beam of radiation into two paths introducing a phase delay between them, and then recombine them to produce an interferogram. A Fourier transform of this interferogram produces a spectrum. A spectrum is represented graphically, as transmission or intensity with respect to wavenumber or wavelength, with peaks and valleys.

Spectra vary with the concentrations of trace gases in the column of atmosphere above the FTS observation platform. Hence, scientists can use spectra to infer vertical gas distributions. The inferential product takes the form of volume mixing ratios (VMRs) – ratios of the mass of particular chemical species in a given volume of atmosphere to the mass of total constituents in that volume – for a number (often ~ 40) of discrete intervals of altitude. The basis of this inference can be a single spectrum, though spectra averages are also employed; the form of the spectral information will differ depending on the timescale of interest and the resolution of the instrument. Here I assume the use of a single spectrum.

Ideally, an empirical correlation would permit direct inference from a spectrum to a vertical distribution of VMRs. However, such a correlation is unavailable. While each trace gas absorbs radiation in a unique way, creating unique spectral line shapes (or signatures), various factors can alter the line shape. Two factors are particularly relevant: Lorentz and Doppler broadening. In short, these broadening effects permit the normally precise, definite, and unique spectral line signature of a gas to widen into another region characteristic of a different molecule's signature, allowing for error inducing 'interfering species'. For example, eight species can interfere with methane's signature, including water vapor and ozone. The resulting overlap permits different VMRs to produce the same observed spectrum. To put it another way, the problem is ill-posed: broadening effects ensure that any equation taking as input the observed spectral lines (suitably quantified) and producing estimates of trace gas concentrations will have no unique solution. This kind of problem is often called an 'inverse problem' because even though scientists can describe mathematically how the spectra are produced, there is no way to invert that description to move directly from a spectrum to VMR estimates.

Instead of trying to find a mathematical relation that facilitates a direct inference from the observed spectrum to VMRs, scientists reason the other way around, moving from the vertical distribution of trace gases (even though their exact values are unknown) to the spectrum using a process called retrieval. In broad strokes, the retrieval process employs an optimal estimation method that relies on previous knowledge of gas distributions to produce a best estimate profile of VMRs. The process can be broken down into two stages, which I will call the choice procedure and the search procedure.

3.1 The Choice Procedure

The choice procedure specifies how to locate the best estimate of the true VMR profile. It draws heavily from inverse theory, which is roughly a set of mathematical techniques applied to inverse problems to help calculate from observations the states that produced them. The techniques described by inverse theory are somewhat general and can be applied to many different kinds of problems. Central to many of the techniques is a forward model. The forward model is an idealized description of the known physics involved in the measurement procedure resulting in the observations. The entire measurement process can be described via the schematic equation $y = F(x, b) + \varepsilon$. Here, x denotes the state vector (i.e. the VMRs of targeted chemical species as well as interfering species at particular altitudes), and y denotes the measurement or data vector (i.e. spectral information). F is the forward model that captures the known physics of the measurement and takes the form of a vector valued function. Also, b denotes other model parameters ('non-retrieved parameters') that represent measurement conditions relevant for producing the observation but that are not the states of interest, and ε is an error term. Call this entire expression the forward equation.

The error term included in the forward equation encompasses not only measurement noise, but also errors associated with the forward model and independently measured non-retrieved parameters. This error term is assumed to take the form of a probability density function (*pdf*) and therefore, the forward equation maps a definitive state vector to a region of measurement space, producing a *pdf* of the measurement vector; in this case, a *pdf* over spectra. Inverse theory specifies how to use the information contained in the measurement, along with knowledge of the state vector, error, and non-retrieved parameters, to calculate a best estimate state vector, denoted $\hat{x}$.

The selection of $\hat{x}$ must be conditioned on prior knowledge due to the ill-posed nature of the problem. Rodgers ([2000], pp. 21-24) demonstrates how the forward model and prior knowledge can be used to characterize all possible solutions to the inverse problem. Assume that there is extant information about the state vector, and thus a prior *pdf* of the state x is available, denoted here as $P(\mathrm{x})$. The forward model maps the state space into observation space and gives information about the probability of the measurement given the state, i.e. $P(\mathrm{y}|\mathrm{x})$. Given this information, Bayes' theorem $P(\mathrm{x}|\mathrm{y}) = P(\mathrm{y}|\mathrm{x})\, P(\mathrm{x}) / P(\mathrm{y})$ can be used to obtain $P(\mathrm{x}|\mathrm{y})$. $P(\mathrm{y})$ is a normalizing factor and often not needed. The full posterior distribution is a characterization of all possible solutions to the problem in terms of all possible states, and provides the sought-after information about the state in terms of the measurement.

Locating $\hat{x}$ requires selecting one state described by the posterior *pdf* as the best estimate of x. Scientists believe that the choice should optimize something, and a popular technique employs optimal estimation to find the maximum a posteriori (MAP) solution defined as the maximum of $P(\mathrm{x}|\mathrm{y})$.[3] Assuming that the *pdf*s involved are Gaussian, one can specify a MAP estimator that can be used to define a cost function whose minimization locates the MAP solution.[4] The cost function is defined such that its value depends on (1) the difference between the spectrum that would have been produced by a candidate state vector and the actual observation and (2) the difference between the candidate state vector and the most likely vector according to the prior. The cost function therefore describes how information from the prior should be combined with information from the measurement when selecting $\hat{x}$. The cost function is essentially a way to select among candidate state vectors that could have produced the observed spectrum. In practice, the minimum value of the cost function is found algorithmically.

3.2 The Search Procedure

An algorithm is used to search the topology of the cost function for the global minimum value through iterative numerical estimation. Two algorithms, SFIT and PROFFIT,[5] are popularly employed in spectrographic retrieval, and both use an optimal estimation technique based on the work of Rogers ([2000]) described above.

[3] While formulated in Bayesian terms, when $P(\mathrm{y})$ is not employed, the optimal estimation is effectively non-Bayesian maximal likelihood inference.

[5] For a detailed description and comparison, see Hase et al. ([2004]).

The algorithms compare the spectrum that would have been produced by a candidate state vector – known as “calculated” or “simulated” spectrum – to the observed spectrum. The forward model used to compute these simulated spectra is essentially a radiative transfer model (RTM) that describes how radiation is affected by absorption, emission, and scattering processes as it interacts with chemical species while being propagated from the top of the atmosphere down to the FTS. Formulating this model for Earth’s atmosphere draws on elements of Radiative Transfer Theory (RTT), including the Beer–Lambert–Bouguer law, an exponential decay law relating radiation attenuation to the materials it passes through, as well as Kirchhoff’s law of radiation, Stefan-Boltzmann Law, and Schwarzschild’s equation (Marzano [2012]).

Comparing the simulated and observed spectra requires the RTM to accurately account for the particular characteristics of the observation instrument employed in the retrieval, such as the instrument’s field of view and resolution (Mariani [2014]). The forward model is capable of accurately predicting spectra that the FTS would have produced under specified conditions. Rodgers summarizes the importance of this forward model nicely: ‘The heart of a successful and accurate retrieval method is the forward model. If it does not accurately represent the physics of the measurement then there is little hope that the retrieval will be satisfactory’ (Rodgers [2000], p. 141).

Both algorithms strategically explore the topology of the cost function in a fairly similar way. Scientists construct what is called an “a priori” VMR profile from existing information, this serves as the prior *pdf* of the state. Starting with the most likely VMRs specified by the a priori profile, the forward model is used to derive a corresponding simulated spectrum, that is, the spectrum that would have been produced by those VMRs had they been actual. Of the parameters in the forward model, only the parameters representing the VMRs of species of interest and interfering species are typically adjustable. The non-retrieved model parameters (i.e. b in the forward equation) representing measurement conditions are not adjustable. These parameters – e.g. representing vertical pressure and vertical temperature – are independently gathered and remain fixed throughout the retrieval process. The relevant error, arising from the forward model, instrument noise, and non-retrieved parameters, is gathered through analyses performed prior to the retrieval.

The retrieval algorithms evaluate the fit between the simulated and observed spectra and compute the value for the cost function. If the difference between the simulated and observed spectra is greater than the signal to noise ratio associated with the instrument, then the difference is attributable to suboptimal VMRs. Thus, the algorithm will attempt to identify another state vector that would lower the value of the cost function, and the process repeats until the fit is within the signal-to-noise ratio of the instrument (or stopping after a prescribed number of steps, in cases of non-convergence). In the case of convergence, the stopping point suggests a zero gradient, which is indicative of a minimum value for the cost function. To ensure the global minimum is identified, the algorithm attempts to make large leaps about the topology by drastically changing the value of retrieved parameters to find sharp gradients that might reveal other minima. In sum, the retrieval algorithm minimizes the cost function by minimizing the residual between the measured spectrum and simulated spectra while being constrained in its search by the a priori profile. Scientists understand the achievement of fit between spectra as evidence of the reliability of the process. Should no simulated spectrum match the observed one, no profile of VMRs can be determined.

3.3 Retrieval: Accounting for the Data

As described in Section 2, to claim that scientists identify phenomena via systematic-derivation in trace gas retrieval it must be shown that: a piece of data serves as an explanandum, that scientists attempt to derive that explanandum from theory, and that the derivation captures how the datum systematically depends on the phenomenon of interest as well as other idiosyncratic influences that helped produce it. The argument that scientists attempt to save the data using derivations in retrieval rests on elements of both the choice procedure and the search procedure.

To help establish the first of the above conditions, let us distinguish between data and phenomena in retrieval. The obvious candidates for the role of data in retrieval are spectra. Spectra are the epistemic foundation upon which these studies are built; they are observable public records produced by the FTS and they bear the idiosyncratic characteristics of their production. The spectra are also epistemically prior to any phenomena claims arising from the retrieval process. If an explanation or derivation is given in retrieval, spectra are the best candidates to serve as explananda, as it is evident from the search procedure that scientists are attempting to predict or re-create them. The choice procedure provides further evidence that the aim of retrieval is to account for spectral observations. Optimal estimation of adjustable model parameters using MAP estimation, which in this case is essentially maximum likelihood estimation (MLE), is commonly described as choosing 'parameter values that are most likely to generate the observed data' (Miura [2011], p. 155). This evidence establishes that scientists' goal in this case is to (at least in part) predict or describe the observation, and thus the observed spectrum could properly serve as an explanandum.

There are three components that might qualify as phenomena: the actual trace gas concentrations reflected by x, the estimated gas concentrations reflected by $\hat{x}$, and the hypothesis $\hat{x}=x$. Of these, only the actual trace gas concentrations for particular situations qualify as phenomena: they are unobservable and ontologically real; they are causally implicated in data production, stable enough to reoccur (in principle if not in fact) in similar form, and are detectable by a variety of procedures. That these gas concentrations are the inferential product of retrieval lends support to their status as phenomena. Therefore $\hat{x}$ should be understood as an attempt describe phenomena whose true value will (if the description is accurate) fall within the ascribed uncertainty. That $\hat{x}$ is an attempt to quantify concentrations of both the target species and interfering species is not worrisome, as phenomena often play the role of error-inducing unwanted influences.

The search procedure indicates that scientists attempt to derive the observed spectrum from theory, and that such derivations include the influence of phenomena and sources of error. Scientists start from the general principles of RTT, which explains the transfer of energy in the form of electromagnetic radiation. They use these general principles to formulate a model of atmospheric radiative transfer, and then tailor that model accommodate the characteristics of their particular FTS. This model permits scientists to calculate – that is, derive on the basis of theory – what the FTS would report given certain conditions. When attempting to compute what the FTS would report, scientists try to overcome obstacles of *representational inadequacy, intractability,* and a *lack of contextual knowledge* identified in the descriptive argument. Scientists have to provide contextual knowledge regarding pressure and temperature profiles, for example, to attempt the derivation. Theory is supplemented by information about the characteristics of the FTS that are gathered during calibration and inter-comparison studies (see Mariani [2014]). Computing simulated spectra line-by-line, and minimizing the cost function, is computationally expensive but tractable; a whole suite of techniques continues to be developed to make the minimization more computationally efficient. Successfully quantifying trace gases helps establish that the models employed in retrieval are representationally adequate.

Systematic dependences between the observed spectrum and the factors that produced it still need to be established. To demonstrate systematicity, it is useful to consider Woodward's counterfactual account of explanation (Woodward [2003]). This account requires that explanations provide information about counterfactual dependencies that exist between explanans and explanandum. To demonstrate this counterfactual dependence a candidate explanation must answer a series of 'what-if-things-had-been-different questions', where these 'w-questions' specify possible causal manipulations of the system (Woodward [2003]). That is, 'the explanation must enable us to see what sort of difference it would have made for the explanandum if the factors cited in the explanans had been different in various possible ways' (Woodward [2003], p. 11). Scientific explanations (though not all explanations) on this account often take the form of detailed derivations from theoretically-based equations. Thus, the account can be used to assess whether a particular attempt at a derivation would establish systematic dependence.

The model used to compute spectra in retrieval is capable of answering w-questions that demonstrate counterfactual dependence between spectra (the explananda) and the forward model plus relevant parameters (the explanans). After all, without the ability to produce a large number of simulated results specifying how spectra would change under different gas distributions and environmental conditions, the search procedure would not locate an acceptable solution to the inverse problem. The derivations performed in retrieval answer w-questions regarding what would happen to a spectrum if the amount of retrieved or interfering species were changed, or were located at different atmospheric heights. Similarly, w-questions about what a spectrum would look like if the pressure changed, or if the instrument had a different field of view, could be answered by using a different pressure or instrument profile. Of course, w-questions that are not specified or related to the parameters involved in the retrieval would not be answerable, for example, how a spectrum would change if the instrument were moved 10 meters to the right, or if a polluting factory came online nearby. Nor would such derivations answer questions about random errors not explicitly represented in the model.

Such w-questions define which dependencies the derivation can capture, and hence the scope of interventions for which an explanation is provided. The accuracy of the answers to w-questions are limited by the accuracy of the forward model, as well as the resolution of the instrument. The point is that Woodward's counterfactual account sanctions the derivations performed in the retrieval as demonstrating systematic dependence between the target phenomenon and data, as well as between interfering species and data. Retrieval should therefore be understood as an example where scientists aim at, and perhaps provide, a systematic-derivative explanation of data.[6]

I say 'perhaps provide' an explanation of data to emphasize that whether a candidate systematic-derivative explanation is successful depends on the accuracy of the forward model, as well as the achievement of fit between simulated and observed spectra. This is to admit that the simulated spectrum is an imperfect match with a noisy observed spectrum: it typically only agrees with the observed spectra within the bounds of error associated with the measurement. At times, Woodward seems to suggest that such imperfections would be intolerable, that is, a systematic-derivative explanation of data would require accounting for all unwanted influences by eliminating typical error terms ([2010]).

But such intolerance for error cannot be maintained on Bogen and Woodward's view. Bogen and Woodward's position is that phenomena can be discovered empirically though bottom-up processes that do not substantially involve theory, and that such phenomena are explainable once discovered. But the results of empirical processes – even, for example, the average of repeated attempts to measure the temperature at which lead melts – will always be subject to uncertainty and error, including when such processes are used to discover phenomena. Since there is no way to eliminate typical error terms when empirically identifying phenomena, systematic-derivative explanations can only save the phenomena within the bounds of error or uncertainty. If such error is tolerated when it comes to explaining phenomena, then it should be tolerated when explaining data. There is no reason not to extend this rationale to cases like atmospheric retrieval: VMR vectors and the forward model can explain observed spectra, error notwithstanding.

It remains to be shown that systematic derivations of data are useful, that is, that the normative argument can be overcome. The way that scientists use the search procedure to find $\hat{x}$ establishes the usefulness of systematic derivations in data-to-phenomena reasoning. Scientists have formulated the cost function in a way that allows its topography to be explored via iterative comparisons between observed spectra and simulated spectra. Systematic derivations of data therefore allow scientists to solve the inverse problem by finding the MAP solution for model parameters that reflect phenomena. Given the non-linear inverse problem, scientists cannot employ strategies that avoid attempts to derive the data in reasoning from spectra to profiles. Thus, P1 of the normative argument fails to hold in a way that favors derivation-strategies. This alone

[6] Note that such derivations are unlikely to explain phenomena. For example, while RTT has the resources to explain radiative phenomena, an explanation of trace gas concentrations would require a theory of chemical transport and interaction.

defeats the normative argument, but P2 fails in ways that support d-strategies as well. Scientists can use other methods that do not rely on attempts to derive data, like sonde- or aircraft- mounted detectors. However, such approaches are resource intensive, expensive, and are impractical for the kinds of near-continuous monitoring performed with an FTS. Contra Bogen and Woodward, then, there are scientific reasons that support the provision of detailed accounts of data production.

Having established that explanations of data can be given and are useful, we may now ask just when data are explained, and how prevalent such explanations are in science. A fairly strong claim can be made, based on the choice procedure in our case study, that any instance of estimating the value of adjustable model parameters via MAP or MLE, when theory provides the structure of the model, is an explanation of data. To see the strength of this claim, consider the similarities that exist between retrieval and more 'typical' cases where MLE is employed. Assume that scientists are able to use theory to describe a data generating process of interest that has structure S, characterized by parameters T. This structure makes it possible to estimate values for T from the data via MLE. Also assume that the value of T reflects a phenomenon and that, just as in retrieval, the value of T is inferred by determining which parameter value confers the highest conditional probability on an observed pattern in the data (i.e. a sample statistic). So long as the fit generated is within the uncertainty or error ascribed to the data, the case seems completely analogous to retrieval, and if retrieval furnishes explanations of data, so would 'typical cases' like this one. While these cases might not be abundant – after all, S is often not derivable from theory or is simply unknown, and determining its form is often the goal of inquiry – explanations of data would be somewhat common in scientific practice on this view.

However, the formal similarities between the two examples are insufficient to conclude that the more typical case always involves explanations of data. In order to have an explanation of data, idiosyncrasies associated with the data cannot be removed without a systematic account. In typical cases of MLE, however, there is often no such attempt to account for idiosyncrasies in this way. To see the point, consider how systematic error might be handled: scientists often aim to remove the effects of systematic unwanted influences from the data before performing MLE, in hopes that the sample will more directly reflect the phenomenon, given certain assumptions. The effects of these unwanted systematic influences are often left unexplained. In these situations, scientists are attempting to remove, rather than account for, certain idiosyncrasies. This approach is part and parcel of the bottom-up strategies that Bogen and Woodward show make data reliable for inferring phenomena without recourse to explanations. Data are characterized by their idiosyncrasies, and removing them without providing a detailed account of their influence makes what is explained – if anything is explained – something other than data (more on this in Section 4).

The retrieval example makes perspicuous the need to account for idiosyncrasies to explain data. Within retrieval, spectra remain idiosyncratic in that scientists are unable to compensate for the unwanted influence of interfering species, which are a systematic but state-dependent source of error. As the search procedure illustrates, rather than attempt to clean the data of its idiosyncrasies, scientists deliberately represent unwanted influences and the atmospheric states upon which they depend. Their strategy is thus different: instead of altering the data so it more directly resembles the phenomenon they wish to infer, scientists infer the phenomenon by altering a systematic-derivative explanation whose product more directly resembles the data. In other words, they are attempting to maximize the likelihood of observing the report provided by the instrument, idiosyncrasies included. The atypical form of the data employed in retrieval helps make this strategy viable: scientists need not look hard for a pattern in the data, *the pattern (i.e. a spectrum) is the data*.

While cases of MLE that are more ordinary than retrieval can employ explanations of data, the practice is unlikely to be as pervasive as MLE itself. Attempts to save the data may be limited to cases where scientists face ill-posed or ill-conditioned problems and scientists have a very good theoretical understanding of the data gathering process. Such situations are often found in the Earth and Atmospheric Sciences, where retrieval-like strategies are widely employed within planetary monitoring via satellite, as well as geological investigations into subterranean structures.[7] A process similar to retrieval is also used in climatological data reanalysis to enhance data consistency.[8] Given the use of retrieval techniques in the Earth Sciences, as well as the prevalence of inverse problems in other fields (e.g. astronomy and medical imaging) it is likely that the actual application of such techniques is wider than the list given here. That such derivations are attempted at all, however, is enough to show that saving the data is a valuable strategy that gives theory a central place in phenomena identification.

The important point is that the identification of phenomena can proceed via a strategy different in character than those Bogen and Woodward discuss. To see the contrast, consider Woodward's example of attempts to measure the deflection of starlight to test general relativity. Woodward emphasizes that, 'Eddington infers *from* features of his photographs (and various background assumptions) *to* a value for the deflection of starlight rather than trying to infer or derive characteristics of the photographs from other assumptions' ([2011], p. 168). In the case of retrieval, examining the search procedure reveals precisely the opposite: scientists attempt to derive the characteristics of spectra *from* theory and other assumptions, in order to make inferences *to* trace gas phenomena. There is thus sometimes good reason to explain data: it is another strategy for reasoning to phenomena.

4. Objections and Responses

One might object that while derivation takes place in retrieval, data is not derived. This charge would vindicate the descriptive argument. Such an objection might claim that in retrieval spectra: (1) must be phenomena, or (2) must not be data.

Claim (1) is untenable. Spectra do not have the characteristics of phenomena: spectra are not thought to be causal agents in the world, but are thought to be records of

[7] Teru Miyake discusses inverse problems similar in form within the context of object tracking and astronomy ([2015]) as well as seismology ([2013]).

[8] See Parker ([2017]).

measurement interactions. Spectra are idiosyncratic in the sense that they will contain errors that are unique to the particular environment and spectrometer that produced them.

Before exploring the objection further, it is worth pointing out that accepting the above response to (1) but maintaining (2) does not completely vindicate Bogen and Woodward. Maintaining this position is equivalent to claiming that there is some intermediary product deployed in reasoning from data to phenomena that is systematically derived in the process of retrieval and plays some functionally useful role in reasoning to phenomena. This would counter Bogen and Woodward's claim that systematic derivations only target phenomena, and do not help identify them. Thus, systematic derivations would still play a useful role in reasoning to phenomena, though Bogen and Woodward's denial that derivations of data can be useful would remain intact.

Responding to (2) requires defending spectra as data. The best argument in favor of (2) is that observed spectra are not 'raw' and are, for example, produced by a Fourier transform of an interferogram. This does not, however, show that spectra are not data. Woodward, for example, characterizes data as 'what registers on a measurement or recording device in a form which is accessible to the human perceptual system, and to public inspection' (1989, 394). Viewed this way, data are not required to be the immediate product of a measurement interaction, but only the public records of such interactions. Maintaining (2) would create inconsistences in many of the examples Bogen and Woodward employ, for instance, it would require arguing that temperature readings from a thermocouple are not data because the immediate product of the measurement is an indication of voltage. Woodward ([2011]) seems to affirm that such temperature readings would be data. If so, given that the effects of the Fourier transform are well known and are taken into account when determining the accuracy of the resulting profiles, it seems reasonable to classify spectra as data.

In a similar vein one might object that the actual vertical distributions of trace gases are not phenomena: phenomena are stable, but a distribution of trace gases is variable across space and time. Such variability might indicate an ontological difference between trace gas concentrations and true phenomena. However, either vertical distributions of trace gases are phenomena, or they are not phenomena but derivations of data are still useful in data-to-phenomena reasoning.

Two considerations suggest trace gas concentrations are phenomena. First, there is no notable difference between gas distributions and examples of phenomena that Woodward employs, like the magnitude of starlight deflection by the sun: both are investigator-independent parts of the physical world and both have similar epistemological grounding in data. Vertical gas concentrations meet the requirements that phenomena be in-principle capable of reoccurrence under different circumstances, be detectable by different methods of investigation, and be ontologically real. Challenges to the accuracy of VMR profiles often resemble challenges to phenomena claims, not data claims. When challenging a phenomenon claim, one does not often question what was observed during detection, but whether the inferences that follow were correctly drawn. These are exactly the kinds of challenges typically raised in the context of FTIR retrieval. Typical challenges include questioning which a priori profile should be employed, or whether atmospheric conditions – excessive cloudiness for example – might invalidate some of the assumptions relied upon during the retrieval. These difficulties concern reasoning to phenomena from data, not the validity of public records.

Second, Bogen and Woodward seem happy to be permissive regarding phenomena, allowing 'all sorts of things, including events, processes, relationships, and instances of property possession, to qualify as instances of phenomena' ([2010], p. 794). Trace gas distributions have a place within (perhaps many) of these categories of things. Furthermore, things that fit into these categories, like the mass of the sun that determines the magnitude of starlight deflection, are constantly changing (though negligibly from our perspective; such changes are still ontologically real). An objector needs to say what differs between cases like this one and trace gas concentrations in order to demonstrate the objection is cogent.

However, the point that theoretical derivations of data have some functionally useful role, or that they are involved in reasoning from data to phenomena, does not rest on the acceptance of trace gas concentrations as phenomena. If one were to understand VMRs as merely some product of data processing, then systematic derivation of data retains scientific usefulness in reasoning to phenomena. VMR profiles are often employed when identifying even more stable processes, for example, diurnal and seasonal chemical cycles (Mariani [2014]), as well as polar vortex intrusions (Whaley et al. [2013]). In supporting the identification of these stable and ontologically real processes, derivations of data also retain their role in data-to-phenomena reasoning.

5. Conclusion

The scientific production of data has increased significantly since Bogen and Woodward wrote 'Saving the Phenomena' in 1988. Now, more so than ever, scientific projects are producing large data sets and embracing data-driven methods. One might think, as Woodward (2011) seems to suggest, that expanding data availability would increase the prominence of theory-independent methods of information extraction and phenomena identification. However, as I have argued here, methods of phenomena identification that involve theory-based explanations of data are also available to scientists. Saving the data can be a viable strategy for phenomena identification that plays a functionally useful role in scientific reasoning as actually practiced. This strategy reveals that theory may sometimes be more central to phenomena identification than was recognized on Bogen and Woodward's account. Ultimately, my argument does not refute Bogen and Woodward's general framework, but rather furthers our understanding of data-to-phenomena reasoning by pointing to an additional strategy for extracting useful information from data. The practices of science are diverse, and at times, saving the data can help meet the challenges of identifying phenomena in otherwise untenable situations.

Acknowledgements

This paper could not have been written without Zen Mariani, whose generosity and patience are unbounded; thanks for taking the time to clarify the intricacies of retrieval for me. I would like to thank Wendy Parker, Margaret Morrison, Joseph Berkovitz, Rasmus Winther, Cory Lewis, Donal Khosrowi and the two anonymous referees for their constructive comments. This paper also benefitted from conversations with Elliott Sober, Michael Goldsby, and John Matelski. In addition, I would like to recognize the Earth, Atmospheric, and Planetary Physics group at the University of Toronto for their willingness to take in an interested philosopher and for inspiring much of this paper. Part of the research for this paper was supported by the Andrew W. Mellon Foundation.

Greg Lusk
Department of Philosophy and Lyman Briggs College
Michigan State University
East Lansing, MI, USA
greglusk@msu.edu

Works Cited

Alston, William P. [1971]: "The Place of the Explanation of Particular Facts in Science." *Philosophy of Science* 38 (1): 13–34.

Bogen, James, and James Woodward. [1988]: "Saving the Phenomena." *The Philosophical Review*, 303–52.

———. [2003]: Evading the IRS. Poznan studies in the philosophy of the sciences and the humanities. In R. Jones & N. Cartwright (Eds.), *Idealization XII: Correcting the model* (pp. 233–268). Amsterdam: Rodopi.

Glymour, Bruce. [2000]: "Data and Phenomena: A Distinction Reconsidered." *Erkenntnis* 52 (1): 29–37.

Hase, F., J.W. Hannigan, M.T. Coffey, A. Goldman, M. Höpfner, N.B. Jones, C.P. Rinsland, and S.W. Wood. [2004]: "Intercomparison of Retrieval Codes Used for the Analysis of High-Resolution, Ground-Based FTIR Measurements." *Journal of Quantitative Spectroscopy and Radiative Transfer* 87 (1): 25–52.

Mariani, Zen. [2014]: "Infrared Emission Measurements of Radiation and Trace Gas Variability in the High Arctic." Dissertation. University of Toronto.

Mariani, Z., K. Strong, M. Wolff, P. Rowe, V. Walden, P. F. Fogal, T. Duck, et al. [2012]: "Infrared Measurements in the Arctic Using Two Atmospheric Emitted Radiance Interferometers." *Atmospheric Measurement Techniques* 5 (2): 329–44.

Marzano, Frank. [2012]: "Radiative Transfer Theory." In *Encyclopedia of Remote Sensing*. [S.l.]: Springer-Verlag Berlin Heidelberg.

McAllister, James W. [2007]: "Model Selection and the Multiplicity of Patterns in Empirical Data." *Philosophy of Science* 74 (5): 884–94.

Miura, Keiji. [2011]: "An Introduction to Maximum Likelihood Estimation and Information Geometry." *Interdisciplinary Information Sciences* 17 (3):155–74.

Miyake, Teru. [2013]: "Underdetermination, Black Boxes, and Measurement." *Philosophy of Science* 80 (5): 697–708.

———. [2015]: "Underdetermination and Decomposition in Kepler's Astronomia Nova." *Studies in History and Philosophy of Science Part A*, Integrated History and Philosophy of Science in Practice, 50 (April): 20–27.

Parker, Wendy S. [2017]: "Computer Simulation, Measurement, and Data Assimilation." *The British Journal for the Philosophy of Science* 68 (1): 273–304.

Rodgers, Clive D. [2000]: *Inverse Methods for Atmospheric Sounding: Theory and Practice*. World Scientific Publishing Company, Incorporated.

Schindler, Samuel. [2007]: "Rehabilitating Theory: Refusal of the 'bottom-Up' construction of Scientific Phenomena." *Studies in History and Philosophy of Science Part A* 38 (1): 160–84.

Small, Gary, Robert Kroutil, John Ditillo, and William Loerop. [1988]: "Detection of Atmospheric Pollutants by Direct Analysis of Passive Fourier Transform Infrared Interferograms." *Analytical Chemistry* 3 (60): 264–69.

Votsis, Ioannis. [2010]: "Making Contact with Observations." In *EPSA Philosophical Issues in the Sciences*, 267–77. Springer.

———. [2011]: "Data Meet Theory: Up Close and Inferentially Personal." *Synthese* 182 (1): 89–100.

Whaley, C., K. Strong, C. Adams, A. E. Bourassa, W. H. Daffer, D. A. Degenstein, H. Fast, et al. [2013]: "Using FTIR Measurements of Stratospheric Composition to Identify Midlatitude Polar Vortex Intrusions over Toronto." *Journal of Geophysical Research: Atmospheres* 118 (22): 12, 766-783.

Wiacek, Aldona. [2006]: "First Trace Gas Measurements Using Fourier Transform Infrared Solar Absorption Spectroscopy at the University of Toronto Atmospheric Observatory." Dissertation. University of Toronto.

Woodward, James. [1989]: "Data and Phenomena." *Synthese* 79 (3): 393–472.

———. [2000]: "Data, Phenomena, and Reliability." *Philosophy of Science* 67 (September): S163–79.

———. [2003]: *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press.

———. [2010]: "Data, Phenomena, Signal, and Noise." *Philosophy of Science* 77 (5): 792–803.

———. [2011]: "Data and Phenomena: A Restatement and Defense." *Synthese* 182 (1): 165–79.