

# Credulous Modellers and Suspicious Experimentalists?

## Comparison of Model Output and Data in Meteorological Simulation Modelling

*Mikaela Sundberg*

This article studies the relation between two specialised practices in meteorology, modelling and field measurements. This relation is embodied in a number of joint practices of which model evaluation is one of them. The relationship between theory, model and observation has been of concern for many philosophers of science whereas the relationship between the working practices that underlie theories, models and observations, has received less attention. This paper describes and compares the practices that generate observation data and model output and the way different roles that fieldworkers and modellers have in this process are established. Next, different practice-oriented perspectives on the status of models, data and experiments are analysed and the discussion then moves on to the different trouble-shooting strategies scientists deploy, depending on their position within the discipline. It is shown that these strategies are often based on the ideas and judgements that modellers and experimentalists express about each other, and by implication, about their own role in the meteorological enterprise.

*Keywords:* meteorology, simulation modelling, falsification

Climate change has received a great deal of attention in politics, media and research. In climate research, climate modelling is the most dominant approach. Climate models are originally based on meteorological models of the atmosphere and meteorology in general and numerical weather prediction models in particular are therefore closely re-

lated to the climate change question (see Edwards, 2001). Some studies of climate modelling have focused on the uncertainties surrounding climate models (Shackley *et al.*, 1998; Shackley and Wynne, 1996). Others have touched on or analysed model evaluation. For example, in concluding his discussion on validation of climate models, Edwards

(1999) notes that the current results of climate model evaluation hold that large-scale features are well simulated by current climate models (cf. Norton and Suppe, 2001).<sup>1</sup> A recent paper by Lahsen (2005) centres on the distribution of uncertainty around climate modelling, but it also raises questions related to the different capabilities of experimentalists and modellers to judge the performance of, and thus the uncertainty surrounding, models. The purpose of this paper is to analyse model evaluation, more specifically comparison of model output and data, in meteorology from a slightly different angle by focusing on the different social practices that are underlying these scientific products. Meteorology is an interesting case because of its crucial role in defining the climate change problem, and also because of its reliance on simulation modelling, which is a scientific practice yet to be explored in more depth. This analysis takes its point of departure in the relationship between the two practices, simulation modelling and field experimenting, and the two groups of practitioners that orient themselves around them.

Scientific practice is a heterogeneous enterprise, not only considering the differentiation into disciplines but also within them (Galison & Stump, 1996). If we aim to recognise some of the heterogeneity of scientific practice, it is of fundamental importance to acknowledge the internal diversity in terms of how science is practiced and not to assume disciplinary unity, for example regarding falsification and more generally, how principles and practices contrast. Several studies of scientific practice have indeed noted the various concerns of different groups of researchers within the

same discipline (Traweek, 1988; Galison, 1997; Saari, 2003). This paper seeks to develop further how different activities are related to different perspectives and how this affects the relationship between groups of practitioners in meteorological research. What keeps them together and what draws them apart?

From the point of view of an institutional setting and structural conditions, meteorology as a scientific discipline can be seen as a social world. All activities that address a given set of coherent and cohesive problems constitute a social world. It is a social whole in which people construct shared ideas about how to go about their business, but they also conduct debates about both their own activities and others that may affect them (Becker, 1982; Strauss, 1978). However, if work and problems are emphasised, meteorology as a discipline (conceived of as an institutional structure) is no longer a “given” unit of analysis, especially since segmentation into subworlds is an inevitable feature of social worlds (Strauss, 1978). In relation to segmentation, *boundary work* may take place. Boundary work is the discursive attribution of selected qualities to scientists, scientific methods and scientific claims for the purpose of drawing a rhetorical boundary between science and less authoritative perspectives, but it can also be used to demarcate disciplines, specialties or theoretical orientations within science (Gieryn, 1999: 4f.). Boundary work is not a concept developed within the social world tradition but largely consistent with the approach. In *intersections*, resources, skills or information flow between segments, or subworlds (Strauss, 1978). Thus, boundary work serves to demarcate, while in-

tersections keep segments together.

Social worlds are actor-defined in the sense that they permit identification and analysis of collectives constructed as meaningful by the actors themselves (Clarke, 1991: 135). Thus, commitment to a social world or subworld is an empirical question. In meteorology, field experimentation and simulation modelling contain important collective commitments and constitute two subworlds of meteorology. The concepts boundary work and intersection will be used to analyse their relationship.

### **Methodological Considerations**

The analysis is based on a qualitative study of meteorological research which took place at the Department of Meteorology, Stockholm University (MISU). MISU is the largest meteorological research department in Sweden. Current research activities at MISU represent meteorology in a broad sense and include studies of both the lower parts of the atmosphere (the troposphere including the boundary layer) and higher parts (the stratosphere) by using simulation models to simulate atmospheric processes or by conducting field experiments. The study is limited to research concerning the lower parts of the atmosphere. During 2003, I conducted participant observation, held numerous informal conversations and discussions with research staff about their work and conducted 19 tape-recorded interviews with various researchers, doctoral students and members of the technical staff. The analysis of the paper is based on their stories during interviews and to a smaller extent on field notes from participant observation. The narratives are

analysed either as descriptions of practice or as expressions of a perspective based on practice (cf. Gubrium and Holstein, 1997). During the analysis, the development of themes and codes was both data-driven and theory-driven (Boyatzis, 1998: 51ff.). It involved an inductive approach in identifying themes, but theoretical guidance in their articulation and development. Thus, model evaluation and the different perspectives on it became important primarily during the analytical phase.<sup>2</sup>

### **Field Experimentation and Simulation Modelling as Subworlds of Meteorology**

A major part of experimental field research at the MISU concerns studies of gases or particles and/or their interaction with clouds in the lower part of the atmosphere. These are small-scale processes where changes in time-scales on the magnitude of seconds to hours are significant. Field experimental work often takes place within large projects which aim to organise field expeditions, so-called field campaigns, at different sites. On these occasions field experimentalists with different backgrounds participate and collect many different types of data by using various types of measuring instruments. For example, different types of radiometers are used for measuring radiation and various anemometers are used for measuring wind velocity and fluxes. For particle measurements, there are a variety of commonly used instruments, such as condensation particle counters and optical particle counters. In addition to skills in data interpretation, field experimental work requires much technical and prac-

tical knowledge, often achieved through extended practice. During the phase of interpretation, researchers try to figure out what happened during the measurements and often use data from other researchers in the campaign to support their case. Field experimentation is not theory-free; the design of campaigns and the interpretation of the data campaigns produce are based on theoretical assumptions, but these theories are generally not the same as those used to construct meteorological simulation models. In the following, the use of models in experimentation will not be discussed.

Whereas field experiments tend to be concerned with details of processes, simulation modelling tends to be more concerned with generalities of larger scale processes. In meteorological simulation modelling, different types of simulation models are used, both in terms of the processes they aim at representing and in terms of the complexity of these representations, i.e. the resolution of the model, the amount of process descriptions, and the type and number of equations. Most simulation modelling at MISU is conducted within the field of dynamic meteorology and based on dynamic modelling of the lower parts of the atmosphere. In dynamic meteorology, moving air masses are regarded in the same way as fluids in motion. In this type of dynamic modelling, the full complexity of the atmosphere is reduced to a small number of physical laws. Numerical versions of the models are constructed in order to simulate the processes that these laws describe. In the models, four non-linear differential equations describe atmospheric dynamics in atmospheric parameters, includ-

ing temperature, humidity, pressure, and wind velocities in three directions.<sup>3</sup> These equations cannot be solved analytically, i.e. it is impossible to write closed form equations that would represent a unique solution to the set of equations. In order to enable the construction of a simulation model, the equations are turned into difference equations and by discretising the equations, modellers approximate the solution. Differential equations relate continuous rates of change over infinitesimal intervals and discretisation relate rates of change over finite, or discrete, intervals. This enables computer calculations. The mathematical model (which is also a theoretical model) is transformed into an algorithm, and the dynamic model is turned into a computational (simulation) model (cf. Winsberg, 1999: 281f.). The results of computer simulations are in the form of very large data sets.

Computational models are based upon a large amount of discrete points, a grid, where approximated equations are calculated in each point. The grid results in a one-, two- or three-dimensional model domain. In a one-dimensional model, the points constitute a vertical pillar or slice. In a two-dimensional model, points are spread over an area. The grid net in more advanced models constitutes a three-dimensional volume. General circulation models used in weather and climate prediction have a global grid and exemplify these types of models. Because the models are prognostic they also include a time dimension.

The resolution of a model refers to the distance between the grid points. In dynamic, atmospheric models, it is gener-

ally coarse. In three-dimensional models used for research purposes only, the horizontal distance between grid points ranges from about half to a few kilometers. At MISU, simulation models with different number of dimensions are used and the most advanced models are three-dimensional limited area models. A few of these models are also used for routine weather forecasting at other locations, but this paper focuses on simulation modelling for research purposes.

The researchers refer to themselves as “modellers” or “experimentalists” respectively and very few of them participate in both modelling and experimental activities. Several doctoral students at MISU, mainly from experimental groups, mentioned that they strive to include both modelling and measurements in their dissertation work, but their ability to do so is constrained. For example, limited time makes it difficult to learn both how to use and interpret measurement technology *and* to set up and run a model, because both activities require a long time to master. As a consequence, people become socialised into either the subworld of experimenting or modelling.

The different activities involving measuring instruments and simulation models create divergent practical interests and concerns. Most modelling and experimenting take place independent of each other and many researchers consider the distance between modelling and experimenting as a problem. The quote below is an example.

A big dilemma within meteorological research is that the experimentalists are over here [the researcher points his hands to the left] and do their stuff, while the modellers are here [the re-

searcher points his hands to the right] and do their things. Then they meet sometimes, and it is often so that it is the modellers who meet the experimentalists because they want data so that they can test their models. It is more unusual that the experimental meet the modellers because they want something. (Bill, modeller, interview)

Consequently, modelling requires data as a resource in order to test models. The principal way to evaluate the performance of a model is to compare simulation output with observations, in other words, the products of simulation modelling and field experimentation. This creates an intersection between modelling and experimenting. In the following, I analyse this intersection more in detail focused on the relationship it creates between modelling and experimenting.

### **Comparing Simulations and Data: “Falsification” Re-negotiated**

In simulation modelling, models are used to explore and learn more about features of the atmosphere. In that sense the models are studied in the same way as a natural system and the model plays the role of an epistemic object – a question-generating tool of investigation (Rheinberger, 1997; Merz, 1999; Morrison and Morgan, 1999). However, simulation models used for exploration are also evaluated according to their ability to reproduce observations, even if agreement between output and observations is often secondary for the purpose of gaining knowledge and understanding (simulated) atmospheric processes.<sup>4</sup>

Simulations produce output divided into many variables, for example, temperature and wind velocity. All of them should, at least in principle, be com-

pared to observation data. For example, it is stated in one article:

Since the modelling system is so complex, it is complicated to perform a total validation of the entire system simultaneously. It also requires a very large and detailed dataset of all dependent variables. During the development stage, every part of the model was tested and compared to observational data... A number of validation experiments were performed for the meteorological part of the model for different types of terrain and surface conditions. (Svensson, 1996: 645)

This shows how comparison of all output variables with data is a very time-consuming task. Depending on the type of simulation, different data is used for comparison. I primarily discuss evaluation of research models with experimental data from field campaigns in order to focus on the relationship between experimentation and modelling.

In scientific publications, model evaluation is often referred to as model validation, but recent discussions, both among scientists and social scientists, have illustrated confusions and uncertainties that accompany model validation (Oreskes *et al.*, 1994; Randall and Wielicki, 1997; Shackley *et al.*, 1998; Küppers and Lenhard, 2005). Some modellers at MISU refer to Popper's falsification doctrine, which holds that scientific hypotheses can be proved false, but not proven true, when they speak of model evaluation. For instance, one modeller said, "Popper is very basic. All theories are good until they are disproved and fall. It's the same thing here." (Britt, modeller, interview) Furthermore, some modellers suggest that output should be compared to measurements every time a model simulates a new case

because they argue that it is impossible to show that a simulation model is successful (in reproducing observation values) once and for all. One modeller emphasised:

You can only invalidate a model. It doesn't matter how many datasets you use to interpolate your numerical model with. It's enough with one single dataset that the model can't deal with and the model is out. You can never validate a model. You may know that a model works for this particular lapse of events, at this particular place, during these certain conditions, but one can never know how it would work to apply it on another phenomenon at a completely different place. (Bo, modeller, interview)

However, models do not seem to be "invalidated" or declared inadequate on the basis of observations (cf. Shackley and Wynne, 1996: 284). Although modellers hold that models can be invalidated but not validated in principle, they neither verify nor falsify models in practice. This discrepancy is not surprising considering earlier sociological research about science. First, sociology of science has strived to show that neither the principles of classic epistemology nor Merton's (1938/1970; 1973) norms of science govern how science is actually produced (Mulkay, 1969; 1975). Second, scientific practice cannot be explained by rational criteria only (Bloor, 1976; Bourdieu, 1977: 19). Whereas the rhetoric of *rules* in general (not particularly in science) tends to be invoked for legitimating purposes and often has its own logic, the logic of *practical work* may be another (Bourdieu, 1998). Mere rules cannot explain practice (Turner, 1994). Thus, the need to be critical towards scientists' statements to outsiders about their work

is perhaps most evident regarding scientists' appeals to norms or "rules" of science, such as Popper's falsificationism. However, even if these rules are not followed, it does not mean that the reference to Popper must be regarded as merely empty rhetoric, independent from the scientists' understanding of their own practice.

Mulkay and Gilbert (1981) have analysed scientific practice in relation to Popper's ideas regarding how much scientists know about Popper, if they use Popper's rules and in that case how, and also how to understand scientists' reference to Popper. They conclude that scientists disagree on the meaning of "Popperian" method and rules and that the general issues of how particular acts are related to rules are indeterminate. My analysis shares some of the conclusions of Mulkay and Gilbert's (1981) paper, but it focuses on the relationship between simulation modelling and field experimentation, embodied in the activity of model evaluation, by illustrating how modellers "defend" modelling and simulation models in different ways. Thereby they bolster the belief in models in spite of lack of agreement with data (cf. Evans, 1997: 414ff.; Star, 1985).

### *Attributing Problems to Data and Experimental Work*

There are various problematic issues related to access and use of data in comparison with model output. One problem is lack of "useful" observations, from a modelling perspective.

For example, some parameters of (modelling) interest are very difficult to measure, some even impossible, and these problems are especially evident in

the case of inhospitable locations where field campaigns are difficult to organise. Moreover, whereas many measurements with simple standard equipment are preferable in order to compare model output with data, experimental projects benefit from more advanced instruments as additional equipment in order to study particular processes. Thus, modelling and experimenting do not benefit from the same types of measurements.

Experimentalists also point out that modellers' lack of (useful) data is caused by their ignorance about what measurements experimental practice can offer, both in terms of what, where and how. According to experimentalists, modellers therefore expect results that are incompatible with what measurements can produce, or do not strive to adjust models to that. One experimentalist stated:

It is important that the models are constructed in such way so that it's possible to use existing data... In this case, much work on modelling is made on unhealthy grounds. To test a certain aspect would require a million airplanes at the same time in the air. Many modellers simulate things that we do not have the conditions to... test or support [by using data]. (Eric, experimentalist, interview)

In spite of the fact that it is difficult to compare output from models with existing observations, the modellers I interviewed did not prioritise rectifying this stated deficiency. For example, one modeller said that it is too time-consuming to (re-)construct models with the purpose of making better use of observations. Simulation models develop gradually and contain standardised ways of working with them in their con-

struction. When modellers implement a new part in the simulation model, programming the computer and making the model work as a whole takes a lot of time. Because the model is dynamically consistent, everything depends on everything and new implementations may disturb the previous balance by creating, for example, unforeseen interactions between different parts of the models. In the modelling view, these problems should be considered when campaigns are designed rather than expected to be dealt with by changing models in order to accommodate data.

Furthermore, a lot of information from measurement instruments is not meaningful for modelling purposes, as the following quotation suggests:

Of course there are possibilities to extract information in the models in a way that makes it more comparable to the measurements that are taking place. But... as long as you speak of basically all models from the meso-scale and up, that's in the scale from a few kilometres and more, the information that you measure in detail with, for instance, aerosol instruments; that information does not exist in the model. It doesn't work to modify the model in order to give that information. (Bill, modeller, interview)

This quote suggests that adjusting models in order to enable better comparisons with data from campaigns is, in many cases, not considered to be an option from within modelling. On the other hand, some think that it is experimentalists who should take into consideration if their data is needed for modelling purposes. One respondent, who is one of the few who works both with measurements and modelling (but mostly with the former), stated:

Both camps have a great responsibility to find out what type of data is needed for the existing model. A typical social behaviour on the measurement side is that once you learn to make a certain measurement, you want to do that for the rest of your life and defend this with that you need these measurements, continuously, without having the slightest idea how they can be used for existing models. (Edwin, experimentalist, interview)

This quotation re-addresses the divide between experimentalists and modellers (“both camps”) and their lack of communication. It also implies that experimenting should be more directed towards the interest in modelling. The significance of boundary work lies not only in the position of a putative boundary but also in the kinds of order it implies (Shackley and Wynne, 1996: 293), in this case it can be related to who is considered to be “support personnel” (Becker, 1982). Because modelling to some extent relies on data from experimenting, experimentalists can be seen as support personnel in relation to, and from the point of view of, modelling. From this perspective, it is reasonable for modellers to expect that experimentalists should take modelling work more into consideration when planning their measurements and when processing raw data, especially since modellers sometimes cannot use data because it is not refined or calculated into the quantities that modelling requires. Modellers also generally lack the knowledge to make use of raw measurement data. (Randall *et al.*, 2003). However, these expectations are not in agreement with the self-perception of experimentalists, who do not feel obligated to spend additional time on data processing to support mod-



elling. The experimentalists rather want to start interpreting the measurements they are interested in as soon as possible. In their view, data processing is a boring, routine job, not accountable as “science”. This is illustrated by the fact that if possible, data processing is handed to technical personnel like engineers and laboratory technicians or doctoral students to be done. They act as support personnel in the experimental subworld.

The view on support can also be related to the difference in status between modelling and experimenting in terms of the different types of work it involves. For example, one modeller said the following about modellers’ and experimentalists’ respective views on the other.

I think that those who are more theoretical, or those who are modellers, who sit and run their models, they sometimes think... that those who are out measuring, they do not care about theories about why it is like this. They are just out measuring and they write about what they measure and it is as simple as that. (Bob, modeller, interview)

The view of modelling as “theoretical” and experimentation as “technical” associates modelling with conceptual work and experimentation with less prestigious gadget work (hands-on work) (cf. Star, 1995; Fujimura, 1987) and it is another illustration of boundary work that serves to demarcate between modelling and experimenting.

The discussion above exemplifies how the expectations of modelling are to some extent incompatible with the problems that the experimental subworld seeks to pursue (cf. Clarke and Gerson, 1990). However, when model-

lers actually find measurement data that they consider useful, there seems to be few conditions when data can contest a simulation model. In spite of a principal view that available data should be an authoritative source compared to model output - *if* the measurement technique is successful, if the instrument has been calibrated properly, or if there is not an error in the instrument - there seems to be many reasons to neglect this principle in practice. In fact, modellers tend to emphasise the problems and errors in data, rather than potential problems in modelling (cf. Shackley and Wynne, 1996: 284).

*Shifting Focus: Attributing Certainty to “The Physics” of the Simulation Model*

In addition to data comparison and the replication of observed patterns as a measure of the capacity of the model, modellers point at “the physics of the model” as important in order to judge its qualities (cf. Küppers and Lenhard, 2005). One modeller said that if the simulation model would “absolutely not” reproduce the observations one would ask: “Can the physics of the model really recreate those measurements, the phenomena you want to recreate? We know it can. All the physics are in this model, so it’s cool.” (Betty, modeller, interview) She continued that there is, however, always a question about the resolution and that the resolution can cause problems in terms of data agreement. This statement illustrates two points. First, if the “physics” of the simulation model is considered as good, it is acceptable with a performance not completely in agreement with data. Second,

there seems to be an assumption that lack of agreement is a consequence of the features of simulation models that surround this basic structure, such as the code or, as in this citation, the resolution, not of some problem in the “physics” (Merz, 1999). Not everyone agrees upon this however. One modeller claimed:

I know I can study the problem I want to study with this model, but when you have submitted articles, someone has said that you can't do this with that model. But then you have arguments: Yes I can do it with this model because of this and this. But everyone does not believe in models and you should be sceptical about what you get out of it. But if you can argue from a physics viewpoint why you can use a model, it's okay. (Bob, modeller, interview)

This quote exemplifies how the claim that “the physics” of the model is correct is a claim that mobilises a form of theoretical resource bound to a cumulative notion of scientific knowledge. In a social setting such as a seminar, it is easier to question the agreement between the lines representing model output and data in diagrams and figures than to question accepted theoretical statements and physical laws. Consequently, by referring to “the physics” of the model, modellers refer to statements or “laws” with such a status that single disagreements cannot question them.<sup>5</sup> The possibility to act as a successful dissenter is also limited by the activities in the subworlds and the knowledge and skill they require (cf. Collins and Yearley, 1992). One experimentalist said that the experimentalists are lost “if the modeller starts writing some horrible formula that the measurement specialist perhaps remembers from studies ten years ago” and alternatively, if experimentalists

“start talking about a sophisticated detector that makes a new type of measurement, then there are many who feel lost in the modelling camp” (Espen, experimentalist, interview). This does not mean that researchers are incapable of judging research results based on other methods than their own, but that the way results are discussed, defended or attacked, including what counts as an adequate response, are likely to depend on whether the dissenter is part of the same subworld or not.

Important to note is also the difficult trade-off between producing the best possible correspondence to observations and paying strict attention to the best physical construction of models. One modeller noted “It is a paradox that the [models] which are good at making weather forecasts are not good at making physically correct forecasts.” (Benny, modeller, interview) However, since opportunistic tuning – adjusting the value of parameters and coefficients in order to make model output agree better with “expected” values - sometimes contribute to the better results of forecast models, “too much” agreement with data is actually likely to create suspicion, as opposed to confidence (cf. Randall and Wielicki, 1997: 405). Consequently, because of modellers’ “common knowledge” that the best physical models are not the best forecast models, reference to tuning of operational models can be used to defend simulation modelling for research purposes, instead of questioning it.

### *Problematizing the Comparison*

When modellers have observations they find useful to compare, they make dia-

grams to compare curves from data and output visually, in addition to or instead of using statistical tools. Numbers have to agree at the same moment in time and modellers do not find that correlation measures can deal with this satisfactorily. Modellers suggest that experimentalists and modellers have different ideas about what constitutes acceptable and reasonable agreement with data, including what can be expected from models in terms of this. For example, one modeller showed me a diagram with two curves, representing the result of a model simulation and observations. He said that for him, the model agrees with the observations, but for an experimentalist it would not, because the curves were not identical. According to this modeller, “people with no experience of modelling” do not know what a “good” result is and they complain that models are “bad” if they do not reproduce observed values precisely because they expect models to “exactly reproduce reality” (as they measure it).<sup>6</sup> Modellers perceive that experimentalists have unrealistic ideals about what models can accomplish in terms of data-agreement. From an experimental viewpoint, models reduce too much of the complexity. However, in grant proposals, it is evident that a lot of field experimentalists motivate their research by pointing at the lack of detail in simulation models (in particular climate models) and their aim to improve it (Sundberg, 2005: Ch. 8). Their criticism is thus best described as ambivalent (cf. Lahsen, 2005: 915) and the different scientific groups also live side-by-side; strengthening each other’s credibility and usefulness (cf. Pinch, 1980).

One reason why modellers and

experimentalists have different opinions about how well output and data agree can be traced to different foci on “details” and accuracy within modelling and experimenting (cf. Lahsen, 2005: 899, 913). An example from a workshop serves to show how experimentalists look more into details compared to modellers. The coordinator of a field expedition organised the workshop in order to interpret data collaboratively. A few invited researchers participated, including a modeller, as well as some people from the department who were not part of the project. At one occasion, everyone was looking at a diagram based on measurements of temperature and humidity and the modeller said to one of the experimentalists: “I would say that nothing is happening, but with your level of detail it’s actually an inversion”. On another occasion, the participants looked at the result of a simulation of air mass trajectories, depicted in diagrams and as lines on a map with the measured area in the centre. One of the experimentalists started to develop an interpretation of measurements in relation to the trajectories, but he was interrupted by two participating modellers who claimed that it was impossible to base interpretations on such a “detailed” reading of the model result. Simulations indicate directions rather than exact trajectories. Thus, different interest in details or accuracy is in fact also a reflection of the different kinds of information that models and measurements can offer.

In addition to the different ideas about what constitutes an agreement – in part related to how detailed the analysis should be – it is not self-evident what is being compared. One researcher said that she and her colleagues who work

with satellite measurements joked that comparing clouds detected by satellites and clouds produced by models is like comparing apples and pears: They are both fruits but slightly different. A modeller reflected upon what is measured and said:

What is a measurement point really representing? Is it a point you measure... If you consider putting up a measurement instrument for wind on one side of the house it is obvious that if it blows on the other side, the measurements are not good... But it is the same thing that makes it different when you measure chemical elements too. If you measure close to a road and there's traffic pollution, then you will measure higher values there than what it really is... When you then put it in the model, it is perhaps one time one kilometre over two meters. That's not what you measured over the road. So you can never say that you have *precise* correspondence to measurements. This is usually difficult to understand: What the models say and what the measurements say and how to compare them in a reasonable way. (Bernie, modeller, interview)

The modeller quoted above said that measurements made in close proximity of a road give “higher values than it is really”. But really *where*? The volume in the simulated atmosphere is a consequence of the mutual adjustment of theoretical presuppositions and computer sources. A “representative” observation would be “representative” of the average value of this volume in the simulated world, but to be “representative” in this way has no meaning in relation to field experimental work – it is an internal modelling reference. Thus, this is an additional example of how modelling and experimental perspectives contain different viewpoints regarding what con-

stitutes a “good”, “useful” observation.

## Conclusions

The analysis above has illustrated some of the heterogeneity within the meteorological discipline, based on a qualitative study of a meteorological research department. More specifically, the paper has analysed the relationship between simulation modelling and field experimenting, and the two groups of practitioners that orient themselves around them. My point of departure for this analysis has been comparison of model output and data, which constitutes an intersection between the practices. Partly because of the falsification principle, comparison is important in principle, but it is less so in practice. Modellers identify various problems with data and “defend” the result of models in spite of lack of agreement with data, in part by pointing at the theoretical construction of the simulation model. While it is not openly suggested, the value of comparison sometimes seems doubted. There are also disagreements between experimentalists and modellers about what constitutes agreement between model output and data.

The examples contain different levels of reasoning which highlight different aspects of the relationship within the social world of meteorology. The constant re-production of the boundary between modelling and experimenting is salient in the scientists' accounts. These accounts enforce stereotypes and exemplify boundary work that re-produces experimentalists and modellers as participants in two different subworlds of meteorological research. The stories also show how the rhetoric of positivistic sci-

ence is built into the scientists own understanding of their work, i.e. their perspective. Yet how modellers deal with the falsification principle in practice is affected by the work they do. Modellers' defence of models serves to legitimise modelling practice, but it is not *only* rhetoric to convince outsiders. The defence also expresses practical thoughts, and a shared modelling perspective, based on the common activities that modellers are involved in.

The analysis of comparison of output and data thus illustrates how activities within a social world can be debated and acknowledges the internal differences and patterns of commitment in meteorology. In practice, modellers seem to take the basic credibility of models for granted (cf. Shackley and Wynne, 1996: 285). Lahsen (2005) even implies that modellers *have to* take the credibility for granted for the sake of their careers, which are based on the performance of their models. In his study of artificial life simulators, Helmreich (1998) in fact suggests that the distinction between the model and the real world sometimes becomes blurred for modellers (cf. Lahsen, 2005: 910 ff.). From within an experimental perspective, however, knowledge about the real world is based on measurements and consequently, experimentalists evaluate the representative qualities of simulation models compared to data. To convince experimentalists (and others) that models are adequate descriptions of the atmosphere is a way for modelling to establish itself as legitimate in the larger disciplinary world of meteorology and it is necessary in order to keep the social world intact (cf. Gerson, 1983: 366f.).

While Lahsen (2005) suggests that

experimentalists are (sometimes) better at judging the shortcomings of models, my point is that the different way experimentalists and modellers work with and understand simulation models and data shapes what it means to be a "good" simulation model or "good" data. The contents of these qualities depend on whether you measure or simulate. From a social world viewpoint, it is the practice-based understandings rather than any abstract criteria that are of importance. Philosophers of science question comparison with data as an autonomous criterion by which simulation models are to be judged (Winsberg, 1999; Zeigler, 1976), but questions related to how model evaluation *should* take place are secondary from the sociological perspective employed in this paper (cf. Becker, 1996: 54f.). In addition, attempts to translate philosophical conclusions into rules of scientific action are misleading if it is not recognised that a major transformation of perspective is involved (Mulkay and Gilbert, 1981). While much of philosophical analysis of scientific knowledge aims at reconstructing the logical structure of the end products of scientific work, scientists themselves are rather involved in making a contribution to a science (Mulkay and Gilbert, 1981: 128), or in other words, science in action (Latour, 1987). To analyse the relationship between the social practices of modelling and experimenting from the point of view of comparison of model output and data is a way to understand the practical significance of the falsification principle and to see how it is embodied and negotiated in practice. Hence, this paper contributes to existing knowledge of simulation modelling by addressing how modellers

*relate* to epistemological standards, at the same time as it is clear that epistemological standards themselves are ill-suited for describing what scientists actually *do*.

Finally, if this analysis of meteorology is regarded as a case study of the relationship between simulation modelling and field experimentation, it should be acknowledged that if these practices differ within different scientific disciplines, this probably affects how comparison of output and data is discussed and practiced. For example, in many scientific fields, simulations are performed in order to learn about systems for which data are sparse and comparison with data obviously become more difficult to practice. In other cases, simulations assume a world in which the results of the simulation are implemented. This illustrates an important difference between the simulations in, for example, particle physics (Merz, 1999) and meteorological simulations. Particle physicists may simulate an existing detector and therefore, the detector and the simulation model constitute partial objects of the epistemic object (Knorr Cetina, 2001). In meteorology, it is possible that the close relation between meteorology and weather forecasting has produced an emphasis on the predictive capacities of meteorological simulations, even when simulations do not have a weather prediction purpose. This may be another aspect that keeps field experimentation and simulation modelling practices together. Thus, special characteristics of the discipline should be considered in order to understand simulation modelling practices and perspectives, while more general patterns of simulation modelling are also important to explore.

## Notes

- 1 Thus, the idea that model output “should” agree with data is taken as a point of departure because it is an idea among the researchers, not a statement by the author.
- 2 I use quotes and descriptions from field notes to illustrate and support my case. Most interview quotes have been translated from Swedish. In order to protect the anonymity of my informants, I do not inform about translation. Excluded words or sentences are marked with ... . See Chapter Three in Sundberg (2005) for further details about the study.
- 3 The first equation, the Navier-Stokes equation, is based on Newton's first law (force equation). It is used to tell how much wind velocity changes with time and it can be used for fluids and gas, like air. The second equation is based on the first law of thermodynamics and tells how much the temperature changes with time. The equation of state is the point of departure for the third equation. It indicates the relationship between pressure, temperature and density. The continuity equation tells how much density changes with time, based on the law of the maintenance of mass. These equations are related to one another in a way that requires them to be calculated step-by-step together.
- 4 Compare Boumans (2006) discussion of different types of model use in economics.
- 5 As Boumans (1999) has discussed, built-in-justification is another reason for modellers to believe in their models. In this paper, I only discuss belief in relation to the comparison with data, which follows from my material on model evaluation and my interest in the relationship between modelling and experimenting.
- 6 On the interpretational flexibility of representations see for example Lynch and Woolgar's (1990) collection of essays on representations in science.

References

- Becker, H.S.  
1982 *Art Worlds*. Berkeley / Los Angeles: University of California Press.
- 1996 "The Epistemology of Qualitative Research" in Jessor, R., Colby, A., Shweder, R. (eds.) *Ethnography and Human Development; Context and Meaning in Social Inquiry*. Chicago: University of Chicago.
- Bloor, D.  
1976 *Knowledge and Social Imagery*. London: Routledge and Kegan Paul.
- Boumans, M.  
1999 "Built-in-justification" in M.S. Morgan & Morrison, M. (eds.) *Models as Mediators. Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press.
- 2006 "The Difference Between Answering a 'Why'-question and Answering a 'How Much'-question" in Lenhard, J & Shinn, T. (eds.) *The Yearbook of the Sociology of Sciences*. Dordrecht: Kluwer Academic Publishers.
- Bourdieu, P.  
1977 *Outline of A Theory of Practice*. Cambridge: Cambridge University Press.
- 1998 *Practical Reason: On the Theory of Action*. Oxford: Polity Press.
- Boyatzis, R. E.  
1998 *Transforming Qualitative Information: Thematic Analysis and Code Development*. Thousands Oaks / London / New Delhi: SAGE Publications.
- Clarke, A.  
1991 "Social Worlds / Arenas Theory as Organizational Theory" in Maines, D. (ed.) *Social Organization and Social Processes: Essays in Honor of Anselm L. Strauss*. Hawthorne, NY: Aldine de Gruyter.
- Clarke, A. & Gerson, E.  
1990 "Symbolic Interactionism in Social Studies of Science" in Becker, H. S & Mc Call M.M. (eds.) *Symbolic Interaction and Cultural Studies*. London/Chicago: Chicago University Press.
- Collins, H. M. & Yearley, S.  
1992 "Epistemological Chicken" in Pickering, A. (ed.) *Science as Practice and Culture*. Chicago: University of Chicago Press.
- Edwards, P.N.  
1999 "Global Climate Science, Uncertainty and Politics: Data-laden Models, Model-filtered Data" in *Science as Culture*, Vol 8, No 4, pp 37-72.
- 2001 "Representing the Global Atmosphere: Computer Models, Data, and Knowledge about Climate Change" in Miller, C. A. & Edwards, P. N (eds.) *Changing the Atmosphere: Expert Knowledge and Environmental Governance*. Cambridge, Massachusetts: The MIT Press.
- Evans, R.  
1997 "Soothsaying or Science? Falsification, Uncertainty and Social Change in Macroeconomic Modelling" in *Social Studies of Science*, Vol 27, No. 3, pp. 395-438.
- Fujimura, J.  
1987 "Constructing 'Do-able' Problems in Cancer Research: Articulating Alignment" in *Social Studies of Science*. Vol 17, No 2, pp. 257-293.
- Galison, P.  
1997 *Image and Logic: A Material Culture of Micro-physics*. Chicago/London: University of Chicago Press.
- Galison, P. & Stump, D. J. (eds.)  
1996 *Disunities in Science. Boundaries, Contexts and Power*. Stanford: Stanford University Press.
- Gerson, E.  
1983 "Scientific Work and Social Worlds" in *Knowledge: Creation, Diffusion, Utilization*, Vol 4, No 3, pp. 357-377.
- Gieryn, T. F.  
1999 *Cultural Boundaries of Science: Credibility on the Line*. Chicago: University of Chicago Press.
- Gubrium, J. F. & Holstein, J. A.  
1997 *The New Language of Qualitative Method*. Oxford/New York: Oxford University Press.
- Helmreich, S.  
1998 *Silicon Second Nature: Culturing Artificial Life in a Digital World*. Berkeley: California University Press.

- Knorr Cetina, K.  
 2001 "Objectual Practice" in Schatzki, T. R., Knorr Cetina, K., von Savigny, E. (eds.) *The Practice Turn in Contemporary Theory*. London: Routledge.
- Küppers, G. & Lenhard, J.  
 2005 "Validation of Simulation: Patterns in the Social and Natural Sciences" in *Journal of Artificial Societies and Social Simulation*, Vol 8, No 4, <http://jasss.soc.surrey.ac.uk/8/4/3.html>.
- Lahsen, M.  
 2005 "Seductive Simulations? Uncertainty Distribution Around Climate Models" in *Social Studies of Science*, Vol 35, No 6, pp. 895-922.
- Latour, B.  
 1987 *Science in Action. How to Follow Scientists and Engineers through Society*. Cambridge: Harvard University Press.
- Lynch, M. & Woolgar, S. (eds.)  
 1990 *Representation in Scientific Practice*. London/Cambridge: MIT Press.
- Merton, R. K.  
 1938/1970 "Science, Technology and Society in Seventeenth Century England". *Osiris, Studies on the History and Philosophy of Science*, Part 4. New York: Fertig.
- 1973 *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: University of Chicago Press.
- Merz, M.  
 1999 "Multiplex and Unfolding: Computer Simulation in Particle Physics" in *Science in Context*. Vol 12, No 2, pp. 293-316.
- Morrison, M. & Morgan, M. S. (eds.)  
 1999 *Models as Mediators. Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press.
- Mulkay, M.  
 1969 "Some Aspects of Cultural Growth in the Natural Sciences" in *Social Research*, Vol 36, No 1, pp. 22-52.
- 1975 "Norms and Ideology in Science" in Mulkay, M (1997) *Sociology of Science. A Sociological Pilgrimage*. Milton Keynes: Open University Press.
- Mulkay, M. & Gilbert, N.  
 1981 "Putting Philosophy to Work" in Mulkay, M (1997) *Sociology of Science. A Sociological Pilgrimage*. Milton Keynes: Open University Press.
- Norton, S.D. & Suppe, F.  
 2001 "Why Atmospheric Modelling is Good Science" in Miller, C. A. & Edwards, P.N (eds.) *Changing the Atmosphere: Expert Knowledge and Environmental Governance*. Cambridge, Massachusetts: The MIT Press.
- Oreskes, N., Shrader-Freshette, K., & Belitz, K.  
 1994 "Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences" in *Science* 263, pp. 641-646.
- Pinch, T.  
 1980 "Theoreticians and the Production of Experimental Anomaly: The Case of Solar Neutrinos," in Knorr, K., Krohn, R., Whitley, R. (eds.) *The Social Process of Scientific Investigation, Sociology of Sciences*, Vol. 4, Dordrecht: Reidel.
- Randall, D. A. & Wielicki, B.A.  
 1997 "Measurements, Models and Hypotheses in the Atmospheric Sciences" in *Bulletin of American Meteorological Society*, Vol 78, No 3, pp. 399-406.
- Randall, D. A. et al.  
 2003 "Confronting Models with Data. The GEWEX Cloud System Study" in *Bulletin of American Meteorological Society*, Vol 84, No 4, April. pp. 455-469.
- Rheinberger, H. J.  
 1997 *Toward A History of Epistemic Things. Synthesizing Proteins in the Test Tube*. Stanford, CA: Stanford University Press.
- Saari, E.  
 2003 *The Pulse of Change in Research Work. A Study of Learning and Development in a Research Group*. PhD dissertation at the Department of Education, University of Helsinki.
- Shackley, S. & Wynne, B.  
 1996 "Representing Uncertainty in Global Climate Change Science and Policy: Boundary Ordering Devices and Authority" in *Science Technology and Human Values*. Vol 21, No 3, pp. 275-302.



- Shackley, S., Young, P., Parkinson, S. & Wynne, B.  
1998 "Uncertainty, Complexity and Concepts of Good Science in Climate Change Modelling: Are GCM's the Best Tools?" in *Climatic Change*, Vol 38, pp. 159-205.
- Star, S. L.  
1985 "Scientific Work and Uncertainty" in *Social Studies of Science*, Vol 15, No 3, pp. 391-427.  
1995 "Epilogue: Work and Practice in Social Studies of Science, Medicine, and Technology" in *Science, Technology and Human Values*, Vol 20, No 4, pp. 501-507.
- Strauss, A.  
1978 "A Social Worlds Perspective" in Denzin, N. (ed.) *Studies in Symbolic Interaction 1*. Greenwich: JAI Press.
- Sundberg, M.  
2005 *Making Meteorology: Social Relations and Scientific Practice*. PhD Dissertation, Department of Sociology, Stockholm University.
- Svensson, G.  
1996 "A Numerical Model for Chemical and Meteorological Processes in the Atmospheric Boundary Layer. Part I: A Model Description and A One-dimensional Parameter Study" in *Journal of Applied Meteorology*, Vol 35, No 6, pp. 939-954.
- Traweek, S.  
1988 *Beamtimes and Lifetimes. The World of High-energy Physicists*. Cambridge, MA/ London: Harvard University Press.
- Turner, S.  
1994 *The Social Theory of Practices: Tradition, Tacit Knowledge, and Presuppositions*, Oxford: Polity Press; Chicago: University of Chicago Press.
- Winsberg, E.  
1999 "Sanctioning Models: The Epistemology of Models" in *Science in Context*, Vol 12, No 2, pp. 275-292.
- Zeigler, B. P.  
1976 *Theory of Modelling and Simulation*. Malabar: Krieger.

Mikaela Sundberg  
Department of Sociology  
Stockholm University, Sweden  
mikaela.sundberg@sociology.su.se