Simulation & Gaming
Volume XX Number X
Month XXXX xx-xx

10.1177/1046878108319640
http://sg.sagepub.com
hosted at
http://online.sagepub.com

Limits to Modeling

Balancing Ambition and Outcome in Astrophysics and Cosmology

Stéphanie Ruphy
Université de Provence (Aix-Marseille I), France

Computer simulations of complex physical objects and processes for which data are very sparse or inexistent have become a major tool of scientific investigation in astrophysics and cosmology. However, one must ask how these simulations acquire their epistemic credentials and whether their realistic ambition is legitimate. A close look at two model-building processes—one in galactic astrophysics, the other in cosmology—reveals heretofore underappreciated features of simulations, such as path dependency. This article argues that such features undermine our confidence in the outcomes of the simulation. Case studies presented here reveal a general tension in computer simulation between realistic ambitions and the possibility of empirical confirmation. The analysis will thus lead to a reassessment of the epistemic goals actually achieved by composite models of complex astrophysical and cosmological phenomena.

Keywords: *astrophysics; Big Bang model; computer simulation; contingency; cosmology; galaxy; limitations; model; modeling; path dependency; pluralism; realism; underdetermination*

As I was working on galactic structure for a doctoral dissertation in astrophysics, I got involved for a couple of years in the improvement of the two—at the time most fully developed—computer simulations of our galaxy. The thrill of immersing myself in the intricacies of galactic modeling soon made way for what I would now qu as epistemological embarrassment. Not only are the two models, which are giving us incompatible representations of the Milky Way, both fitting the data at hand, but they are also displaying a remarkable capacity of adaptation and survival when new data come in. How then can we know which model, if any, is giving us an accurate representation of the galaxy? It seems that the more detailed and realistic we want to make these models, by incorporating increasing numbers and variety of structural components, the more we are losing control of their validity.

Author's Note: I wrote the core of this article in the fall of 2006 during my stay as a research fellow at the Pittsburgh Center for Philosophy of Science. Many thanks to the Center for its support and the benefits of its rich intellectual life.

A more extensive acquaintance with the philosophical literature on models and simulations has not, unfortunately, made the embarrassment fade away. It has even got worse with the multiplication, in the past 10 years or so, of purportedly realistic computer simulations of astrophysical objects and processes. A recent and striking example is the simulated image of large-scale cosmic structures that made the front page of *Nature* with the headlines "Evolution of the Universe" (Springel et al., 2005). Even for those familiar with pictures of filaments and clusters of cosmic matter, it was not obvious at all that the image was constructed using simulated data rather than real data, unless they read the subheadline "Supercomputer simulations of the growth of 20 millions galaxies." With this kind of scientific image, what is obvious at once is the realistic ambition of the simulation that produces them: The simulation aims at mimicking the evolution of a real-world system by producing data that make up for the scarcity of observations; those data are then used to test various theoretical hypotheses. However, on which grounds should we trust the outcomes of such simulations? Is their ambition of realism legitimate? More generally, how do computer simulations of complex real-world phenomena obtain their epistemic credentials? Those questions are all the more pressing because computer simulations have become a major tool of investigation in astrophysics and cosmology. These disciplines are in this respect no exception to a general trend in science vividly summarized in a report to the U.S. National Academy of Sciences a few years ago as follows:

> [But] it is only over the last several years that scientific computation has reached the point where it is on a par with laboratory experiments and mathematical theory as a tool for research in science and engineering. *The computer literally is providing a new window through which we can observe the natural world in exquisite detail.* (J. Langer, as cited in Schweber & Wächter, 2000, p. 586 [italics added])

Acknowledging the importance of computer simulations has become commonplace. Welcome philosophical attention has been paid recently to the details of modeling practice in various branches of science, providing numerous insights into the role that they play in scientific research and into their relationships with theories and experiments (e.g., see Morgan & Morrison, 1999). However, we must give careful epistemological scrutiny to the issue of the epistemic credentials of computer simulations, given that they do not simply inherit the epistemic credentials of their underlying theories (Winsberg, 2003).

My aim in this article is to contribute to filling this gap by offering an assessment, based on a close look at two model-building processes (one in galactic astrophysics, the other in cosmology), of the epistemic goals actually achieved by purportedly realistic computer simulations. In other words, does the computer really succeed in providing us with a new window through which we can observe the cosmos?

Outline

The article is organized as follows: I will first discuss in some detail how the models are built in the two case studies. The discussion will put to the fore underappreciated features of models and simulations (such as path dependency and plasticity), which I will argue undermine our confidence in their results. I will show in particular how these features account for the embarrassing epistemological situation described at the very beginning of this article, to wit, the existence of a persistent plurality of incompatible but equally empirically successful models. Special attention will be paid to (often unacknowledged) pragmatic constraints that bear on the model-building processes. I will also explain why path dependency and plasticity support an unrealistic interpretation of the stability and empirical success of a model or simulation when new data come in. An important lesson drawn from these case studies will be that realistic ambition and the possibility of empirical confirmation pull in opposite directions. The analysis will also shed light on a rather widespread form of representational pluralism in astrophysics and cosmology that follows from path dependency, to wit, permanent incompatible pluralism. I conclude with a call for epistemological prudence and a reassessment of the epistemic goals actually achieved by computer modeling of complex real-world astrophysical phenomena. Note that following scientific practice, I will use indifferently the term galactic "model" or the term galactic "simulation" in this article. In spite of the absence of a dynamic dimension built in galactic models, this terminological versatility can be vindicated by the fact that galactic models, as simulations, aim at representing one particular physical phenomenon, rather than a class of phenomena. Moreover, galactic models also share with simulations the epistemic goal of producing data that make up for the scarcity of observations (for an overview of the relationship between models and simulations, see, for instance, Hartmann, 1996).

A Simulated Universe: The Millennium Run

My first case study is the simulation of the evolution of the universe given in the introduction as a striking example of the realistic ambition of computer simulations. In this cosmological example, the simulation—modestly dubbed the Millennium Run—aims to mimic the formation and evolution of the structures formed by the matter (both dark and visible) in the universe, for the first time on a scale large enough (a cube-shaped region 2.230 billion light years on a side) to make statistically significant comparisons with observations, in particular with recent comprehensive surveys of low red-shift galaxies.

Concretely, the Millennium Run provides, at any time from a few hundred years after the Big Bang to now, the spatial repartition of a very high number (1.0078 ×

10^{10}) of particles of dark matter and a catalog of 20 million virtual galaxies. Just to give an idea of the computer power involved, mimicking the evolution of cosmic structures on such scales took 28 days of wall-clock time, corresponding to about 350,000 processor hours of CPU time. The outputs are then used to construct visual representations such as the one that made the front page of *Nature*.[1]

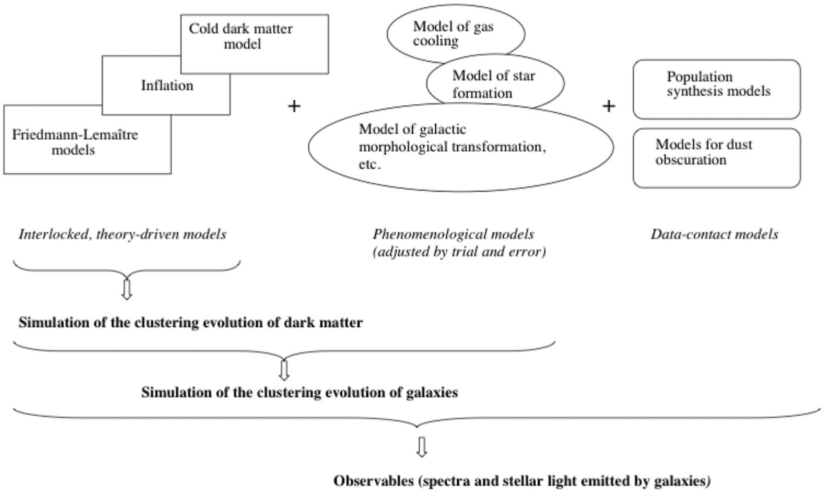
To simplify a bit, the simulation draws on mainly three different kinds of models, forming a hierarchy of interlocking models: what you can describe as "theory-driven" models, phenomenological models, and "data-contact" models (Figure 1). As is well known, cosmology starts by assuming that the large-scale evolution of space-time can be determined by applying Einstein's field equations of gravitation everywhere. And that plus simplifying hypotheses, which I will comment on later, give the family of standard models of modern cosmology the "Friedmann-Lemaître" universes. In itself, a Friedmann-Lemaître model cannot account for the formation of the cosmic structures observed today, in particular the galaxies: The "cold dark matter" model is doing this job. To get off the ground, the cold dark matter model requires initial conditions of early density fluctuations. Those are provided by the inflation model. This first stratum of interlocked models allows the simulation to mimic the clustering evolution of dark matter. But of course, because by definition dark matter cannot be observed, the dark matter distribution must be linked to the distribution of the visible matter. This link is provided by models of galaxy formation. Those models are what astrophysicists call "semi-analytic" or phenomenological models. They integrate the modeling of various physical processes (such as gas cooling, star formation, and morphological transformation of galaxies), and many modeling assumptions and parameters in these models are adjusted by trial and error to fit the observed properties of galaxies. Finally, to make contact with observations, another stratum of models is needed that converts the outputs of the phenomenological models into properties directly observable (such as the spectra and magnitudes for the stellar light emitted by galaxies).

The question that interests us now is the following: At each step of the model-building process, are alternative submodels available? To answer this question, I will turn to what cosmologists themselves have to say, by drawing on a very recent and comprehensive survey of present-day cosmology (Ellis, 2006). Consider first the choice of a Friedmann-Lemaître model as the basic framework for further cosmological studies. When resolving Einstein's field equations of gravitation, you need, in order to obtain a Friedmann-Lemaître model, the assumption that once you have averaged over a large enough physical scale, the universe is spatially homogeneous as well as isotropic. But how do you justify this assumption? Is it empirically justified? Well, the answer is . . . not really!

After having reviewed different arguments in favor of spatial homogeneity, Ellis concludes:

> Finally the argument for spatial homogeneity that is most generally accepted is: isotropy everywhere.

Figure 1
Main Ingredients of the Millennium Run, Modeled on Verbal Description in Springel et al. (2005)



> If all observers see an isotropic universe, then spatial homogeneity follows. . . . Now we cannot observe the universe from any other point, so we cannot observationally establish that far distant observers see an isotropic universe. *Hence the standard argument is to assume a Copernican Principle: that we are not privileged observers*. . . .
>
> Establishing a Robertson-Walker geometry [i.e., the geometry of a Friedmann-Lemaître model] for the universe relies on plausible philosophical assumptions. *The deduction of spatial homogeneity follows not directly from astronomical data, but because we add to the observations a philosophical principle* [the Copernican Principle] *that is plausible but untestable.* (2006, p. 24 [italics added])

Consider another key ingredient of the simulation, namely, the inflation model. Inflation is today a very popular hypothesis among cosmologists, in spite of several serious shortcomings.[2] In a nutshell, the inflation model suffers not only from the lack of definitive observational proof that inflation indeed took place, but also from the fact that the identity of the proposed inflationary field (the “inflation”) has not yet been established (Ellis, 2006, p. 16). So no link with particle physics has yet been realized that could reinforce our confidence in inflation. Sure enough, inflation did solve a number of puzzles that had been hampered for a long time such as the Big Bang model and the so-called horizon problem. And recently, its popularity has been further bolstered when the model successfully accounted for the anisotropies observed in the cosmic microwave background. But such achievements lose their luster when one realizes that alternative models exist—the topological defect model is one of them (e.g., see Durrer, 2002)—with similar explanatory power and empirical support. Hence, Ellis's conclusion:

> *Inflation is not an inevitable conclusion*, *for there are some alternatives proposed*, and the WMAP [microwave background] results can be reproduced by any scenario where Gaussian scale-free perturbations of suitable amplitude occur at the start of the Hot Big Bang. (2006, p. 16 [italics added])

A similar conclusion can be drawn about another key ingredient of the simulation, to wit, the cold dark matter model, which mimics the clustering evolution of dark matter. Quite obviously, a basic presupposition in this model is that there is such a thing as dark matter, that is, some unknown, exotic form of matter that is not seen but is supposed to dominate the dynamics of the universe. Postulating dark matter is an answer to some puzzling observations of galactic dynamics. But there are alternative interpretations of these observations. Here again, it is worth quoting what cosmologists themselves have to say about dark matter:

> Many attempts have been made to identify its nature, . . . but what it is composed of is still unknown. Laboratory searches are under way to try to detect this matter, so far without success. A key question is whether its apparent existence is due to our using the wrong theory of gravity on these scales. *This is under investigation, with various proposals for modified forms of the gravitational equations that might explain the observations without the presence of large quantities of dark matter.* (Ellis, 2006, p. 11 [italics added])

And the same goes for the so-called dark energy that is another key ingredient of recent cosmological models. There are other ways to interpret observations (in that case, observations of very distant supernovae) than to postulate the existence of a form of energy that we know nothing about, except that its effect would be to accelerate the expansion of the universe. I am well aware, of course, that those scanty remarks do not reflect the intricacies of current scientific debates (for extensive references to the relevant scientific literature, see Ellis, 2006). But the oversimplification is deliberate. The details of the scientific arguments are not relevant to the epistemological point I want to make. The very existence of such debates suffices.

Figure 2 sums up the foregoing remarks: At each stratum of the model-building process, there exist alternative submodels with similar empirical support and explanatory power. And at each step, the choice of one particular submodel among various other possibilities constrains the next step. Inflation, for instance, is appealing once a Friedmann-Lemaître universe is adopted (which requires buying a philosophical principle, to wit, the Copernican principle). When starting, alternatively, from a spherically symmetric inhomogeneous model, inflation is not needed anymore to account for the anisotropies observed in the cosmic microwave background

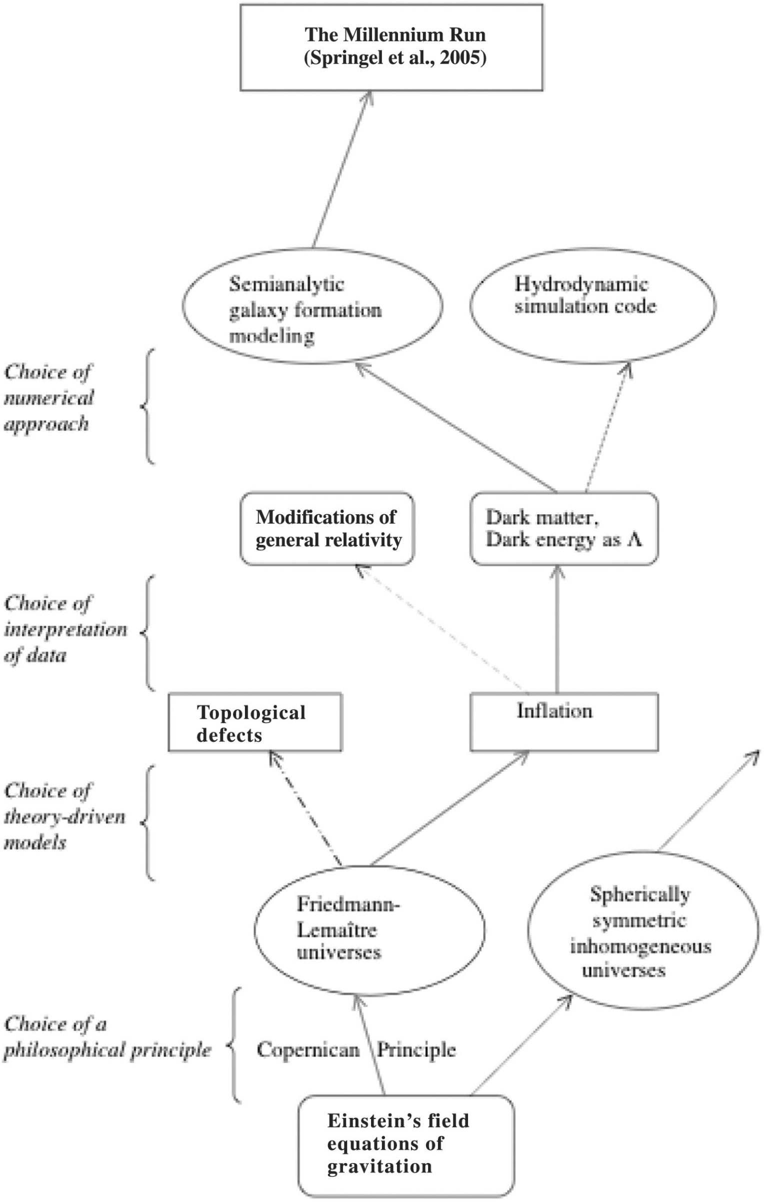
(Ellis, 2006, p. 23). Moreover, even when working in the framework of a Friedmann-Lemaître model, we have seen that at least one alternative scenario, based on the existence of topological defects in the early universe, has also been shown to lead to predictions conforming to observation. So further considerations on the formation of galaxies could have, as a starting point, a model of topological defects rather than an inflation model. This path dependency does not appear only at the level of basic philosophical principles such as the Copernican principle or theory-driven models such as the inflation model. It also manifests itself at the level of the interpretation of astrophysical observations, in particular by the choice of the dark matter and the dark energy hypotheses. Hence, the path dependency of a simulation such as the Millennium Run, illustrated in Figure 2. What I mean here by path dependency is simply the following: Building a simulation of a complex real-world system usually involves putting together specific submodels of particular components and physical processes that constitute the system (hence, the term "composite model"). What may happen is that at each stratum of the model-building process, alternative submodels, equally empirically successful, are available. So the outcome of the simulation turns out to depend on a series of choices made at different levels, from very fundamental hypotheses to more local and pragmatic technical decisions.

Path Dependency and Contingency

Acknowledging path dependency immediately puts to the fore the contingency of a simulation such as the Millennium Run. Had the cosmologists chosen different options at some stages in the model-building process, they would have come up with a different picture of the evolution of cosmic matter. And the point is that those alternative pictures would be equally plausible in the sense that they would also be consistent both with the observations at hand and with our current theoretical knowledge. To deny this would clearly partake of an article of faith. For in light of what has been said on the existence of alternative submodels, there are no good grounds to claim that the path actually taken by modelers is the only path leading to a plausible (in the foregoing sense) simulated universe. Otherwise put, at each step of the model-building process, the choices actually made were not the only possible rational choices, which raises the issue of the kind of nonepistemic constraints that played a role in the choice of the modeling path actually developed (I will come back to this point later).

For the moment, note that if the Millennium Run does not have serious competitors (yet), it is not because alternative paths have also been fully developed and dismissed on empirical grounds. Rather, because of the cost in terms of material and intellectual resources of developing alternative simulations built with different modeling ingredients, only one path has been taken to its end, that is, to a level of detail and to a scale large enough to allow significant comparison with observations. There

Figure 2
Path Dependency of the Millennium Run



are thus no good grounds to exclude that had the cosmologists the resources to fully develop alternative paths, they would come up with different, but equally plausible, representations of the evolution of the universe.

Before discussing in further detail the epistemological consequences of path dependency, let me first lay out another case study in which, by contrast to the previous one, different modeling paths have actually been taken to their ends, leading to a plurality of equally empirically successful but incompatible representations of the same object. This case study is the one briefly discussed at the very beginning of this article, to wit, the case of computer modelings of our galaxy, the Milky Way.

Two Models of Our Galaxy

Our galaxy is currently described as a spiral galaxy, characterized by a concentration of stars, dust, and gas forming spiral arms in a disc and a bulge situated at the center of the galactic disc. Unfortunately for astrophysicists, the sun seems to be located in one of these spiral arms so that no outside vantage point is available for direct observation of the shape, size, and composition of the main structural components of the Milky Way. Imagination thus plays a crucial role in the building of galactic models.[3] Astrophysicists do not derive them from theory, nor do they proceed by abstraction, idealization, or simplification of a phenomenon they would have observed beforehand (recall that there is no outside vantage point to observe the galaxy). Rather, they grope their way toward agreement between predictions and observations by incorporating various ingredients that they imagine are parts of the real object. In that respect, galactic models are perfect examples of composite, autonomous models. Building a galactic model proceeds by trial and error, starting from a minimal number of structural components (usually a stellar disc and a halo) and then adding features (such as a central bulge and spiral arms) so as to obtain a representation empirically adequate in as many respects as possible (which requires comparing in terms of star counts, predictions, and observations in as many directions and domains of wavelength as possible). At each stage of the model-building process, choices are made between equally empirically successful options relative to the physical characteristics of the structural components incorporated in the model, hence, the path dependency of galactic models.

A galactic model is not only a unique tool to learn something about the Milky Way as it is today; it is also a unique tool to investigate its history, that is, its dynamic and chemical evolution. Moreover, our galaxy being considered as a typical spiral galaxy, galactic models aiming at its detailed representation are also expected to shed light on the structure and composition of other, far away spiral galaxies much less accessible to observation. In addition to these central epistemic goals, there are also practical goals. Very often, when an astronomer prepares an observation campaign, she needs to predict what will be, for a given domain of wavelength, the respective contributions of galactic objects and extragalactic ones. A galactic model is in this case used as a tool of prediction.[4]

More precisely, a galactic model aims primarily to represent the different stellar populations that constitute the Milky Way today and how these populations are distributed along different structures. Note that "representing" does not refer here to anything visual. The output of the model is a catalog of "pseudo-stars" from which one can immediately obtain, for comparison with observations, a number of stars (usually per square degree) to be detected in a given direction and in a given magnitude (i.e., brightness) range. For each pseudo-star, the model provides both apparent properties (such as apparent brightness and apparent velocity) and intrinsic ones (such as absolute brightness, color, and spectral type). Catalogs of pseudo-stars are obtained by Monte Carlo simulation from a famous equation in astrophysics, the fundamental equation of stellar statistics:

$$A(m) = \int_R \Phi(M)\rho(R)R\omega dR,$$

where $A(m)$ is an output of the model, to wit, the number of stars of apparent magnitude m, ω is the observational solid angle, R the distance on the line of sight, and M the absolute magnitude. Note that the number of stars actually observed is not simply $A(m)$: Phenomena of interstellar extinction and reddening must be taken into account as well as estimation of the position of the sun, estimated errors in the measurements of stellar magnitudes, and so forth.

The two key components of the equation are $\Phi(M)$, the function of luminosity, which describes how the stars are distributed in absolute magnitude, and $\rho(R)$, the density law, which describes the distribution in space of stars. To simplify, the challenge is essentially to come up with accurate functions of luminosity and density laws for each stellar population and structural component.

Galactic Model Pluralism

There currently exist several models of the Milky Way built along the foregoing approach. Let us focus on the two models I have worked on, which happen to be the two most fully developed ones, the SKY model (Cohen, 1995; Wainscoat et al., 1992) and the Besançon model (Robin & Crézé, 1986; Robin et al., 2003). Three features of these models are directly relevant to our epistemological concern.

First, the two models have roughly the same intended content. Following Kitcher's use of the terminology in the case of maps (2001, chap. 5), I mean by "intended content" a model of the type of entities and the type of relations it is supposed to represent. By sharing the ambition of coming up with a realistic picture of the main structural components of the Milky Way, both models aim to integrate the right structural components with their true physical characteristics (composition and shape in particular). The two models cannot thus be said to differ by dealing with different aspects of the galaxy. Note that if the two models do not differ significantly in their intended content, they do differ in their approach to come up with accurate functions of luminosity and density laws for each stellar population and structural component. The SKY model adopts a strict empirical approach in which functions of luminosity and density laws are free parameters in the model. They are determined by independent observations of the physical properties of stellar populations or by "tuning" the model to adjust predictions to observations. Developers of the Besançon model adopt a more theory-oriented approach: they draw on theoretical knowledge about stellar evolution and dynamic and chemical galactic evolution to obtain the values of some of the free parameters. In this "synthetic" approach, the number of free parameters is reduced. For instance, they use the potential appearing in the Boltzmann equation to constrain the values of the parameters describing the vertical spatial distribution of the stars in the disc, and this potential must correspond to the potential produced by the overall mass distribution of the galaxy, as described by the Poisson equation (Robin et al., 2003).

Second, the two models are incompatible. Having been developed over 10 to 20 years along different modeling paths, they give us incompatible descriptions of what they take to be the main structural components of our galaxy. In the Besançon model, for instance, the laws describing the spatial distribution of the different stellar populations in the disc have been chosen to have the form of an "Einasto" law—a variant of the standard exponential law (Robin et al., 2003), whereas the SKY model uses the standard form. Additionally, the scale length and the radial cutoff of the disc (two key parameters describing its shape and its size) do not have the same value in the two models (Ruphy et al., 1996). Another illustration of the incompatibility of the two models is the fact that they do not integrate exactly the same components: A "molecular ring," for instance, is present only in the SKY model, whereas a "thick" stellar disc is found only in the Besançon model (other significant differences are discussed in Ruphy et al., 1997).

Third, both models enjoy comparable empirical support. This means that both models conform roughly to the set of available and relevant observations. Note, though, that this set is not exactly the same for the two models. Star counts in the far infrared are, for instance, relevant observations only for the SKY model because the Besançon model deliberately does not include a very minority type of star visible only in this domain of wavelength. In other words, the two models do not have exactly the same intended content, but this does not bear on the argument presented here, for the two models disagree on some parts of their common intended content. Thus, galactic modeling presents us with an actual case of plurality of incompatible models of a real-world system with the same intended content and with similar empirical support. How is that possible? Simply because given its composite nature, a galactic model may integrate several inaccurate submodels, whose combined effects lead to predictions conformed to the observations at hand. In other words, it is not unlikely that galactic modelers get the right outcomes (i.e., star counts in agreement with the observations at hand), but not for the right reasons (i.e., not because the model incorporates the “right” submodels). And because astrophysicists cannot test the submodels independently against data (to make contact with data, a submodel often needs to be interlocked with other submodels), there is unfortunately no way to find out if the conformity with observation is accidental and which model, if any, gives us accurate knowledge about the constituents of the galaxy. The situation clearly smacks of some variant of Duhemian holism. To paraphrase the French physicist and epistemologist Duhem (1954, p. 187), the astrophysicist can never subject an isolated submodel to an observational test, but only a whole group of interlocked submodels. Consequently, given the pitfall of accidental conformity with observation, the empirical success of the whole group is a poor guide to the representational accuracy of the submodels.

Note that for the practicing scientist, this variant of Duhemian holism is much more preoccupying than the traditional version concerning physical theories. In the latter, the role of the auxiliary hypotheses the physicist makes about instruments when confronting theory and observation admittedly thwarts, in principle, the possibility of testing a particular theoretical hypothesis. In practice, however, as Duhem himself points out, the physicist’s *bon sens* provides a welcome way out, by allowing him to sort out, among all the hypotheses involved in the confrontation with data, the ones that can safely be considered as reliable. (As any scientist could tell, systematically questioning the functioning of the instruments when interpreting a negative experimental test is certainly a logical possibility, but in practice it is not, surely, an efficient way to develop fruitful research.) Such a way out is unfortunately not available for the galactic modeler. The various submodels involved in the confrontation with data are, so to speak, on a par: It is not the case that some of them can get independent credentials (in the way hypotheses concerning the functioning of instruments do).[5] Moreover, the more the model is made realistic, the more it incorporates various submodels and, consequently, the more it runs into the foregoing holist limitation of its testability. In other words, there seems to be a trade-off between the realistic ambition of a model and the reliability of the knowledge it delivers: The more composite a model gets, the more you lose control of its validation. In light of these difficulties, galactic model pluralism comes thus as no surprise.

Temporary Versus Persistent Incompatible Plurality

To grasp why galactic model pluralism is epistemologically problematic, it might be useful to contrast it with other kinds of model pluralism. A first straightforward distinction is between compatible and incompatible pluralism. Compatible pluralism comes in several forms. It may, for instance, reflect different modeling tasks: More or less complex models of a given phenomenon are developed, depending on the epistemic purpose of the modeler (see Parker, 2006, for a discussion of this kind of pluralism in climate modeling). Model pluralism may also come from differences in the intended content of the models. Bailer-Jones (2000) illustrates this form of compatible pluralism in her analysis of the modeling of extended extragalactic radio sources: A plurality of models are developed, each aiming to represent a specific feature of the radio source. The models are then put together to provide a global picture of the phenomenon. Mitchell's "integrative pluralism" (2002) can also be seen as a form of compatible pluralism: Different idealized causal models of a class of phenomena are integrated to account for a particular concrete phenomenon. In all those cases, model pluralism is epistemologically inoffensive. It does not, in itself, undermine the capacity of the models to fulfill their function, namely, to provide reliable insights of what they specifically purport to represent.

Incompatible pluralism might be much more problematic, depending on whether it is temporary or persistent. Scientific modeling is typically seen as an open-ended enterprise associated with research in progress. It is, for instance, often the case that to account for a new phenomenon, scientists come up with several competing models. But the implicit agreement between scientists is that this plurality is temporary—except, of course, when models are considered as purely instrumental tools. Alternative, incompatible models will compete until the conflict is resolved, usually when new data come in that favor one model over the others. Temporary incompatible plurality is also epistemologically inoffensive. Until the conflict is resolved, models just keep a hypothetical status: They are simply considered as tools of investigation providing tentative accounts of observed features of a system, rather than tools of representation providing reliable insights on its underlying physics. The situation becomes epistemologically embarrassing only when incompatible plurality cannot be expected to go away when new data come in. For if you face a case of persistent incompatible plurality, then you have to retreat to a modest view of the epistemic function of your model, to wit, merely saving the phenomena at hand, rather than providing reliable knowledge on its underlying physics. In short, you have to give up on the realistic ambition of your model. We have seen in the previous section that galactic model pluralism is a form of incompatible pluralism. So let us ask now whether galactic pluralism is here to stay.[6]

Plasticity and Stability

Persistent pluralism follows from the plasticity and the resulting stability of galactic models. Plasticity refers to the possibility of adjusting the ingredients of a model so that it remains empirically successful when new data come in. Note, though, that the adjustment processes I will be concerned with do not boil down to the ad hoc tinkering of the model. The scientists' realistic (rather than instrumental) take on galactic models recommends keeping such tinkering as limited as possible.

Ad hoc adjustments are not precluded only on methodological grounds, but also by the fact that in spite of its high number of free parameters, a galactic model cannot be tuned to fit new data by merely playing with the values of its free parameters. The reason for this de facto limitation is that a change in the value of a free parameter often affects outcomes of the model that have already been tested against data, thereby disrupting previous empirical success. In other words, the multiplication of observational constraints (i.e., star counts in different domains of wavelength and galactic directions) progressively "rigidifies" the model by reducing the modeler's leeway for fine-tuning. So adjusting the model when new data come in often requires incorporating new structural components or/and new populations of stars. This takes us back to path dependency: The way models are further developed in response to a given set of new data usually differs from one model to the other, depending on previous choices made along the modeling path. And the point is that in both cases—adjustment by fine-tuning or by additions—fit between new data and outcomes of the model can be obtained without altering previously chosen key ingredients of the model. Hence, the *stability* of the model.

To sum up, plasticity and path dependency suggest an unrealistic interpretation of the stability and empirical success of a model or simulation when new data come in. In other words, the capacity (even in the long run) of a path-dependent model or simulation to account for newly observed features of the system cannot be taken as a reliable sign that it gets the relevant underlying physics right.

Selection Criteria

The situation in galactic and cosmological modeling reminds us of some aspects of Lakatos' views on the development of a research program. Here, too, albeit at a very local level of particular models, a methodological decision is taken by scientists as regards which pieces of the current accepted knowledge are susceptible to be revised in light of new data. Two levels of model selection criteria must be distinguished. At a first level, the criteria concern the choice of one submodel rather than another at a given stage of the model-building process. For instance, to which grounds should one go for inflation rather than topological defect in building a simulation of the evolution of the universe, given that both submodels have comparable (indirect) empirical support and explanatory power? Or, to go back to galactic models, why go for an Einasto law rather than a standard exponential law to describe the spatial distribution of the stellar population in the disc, given that each submodel can lead to an equally empirically well-supported galactic model? At a second level, the criteria concern the methodological decision to further develop a model along a given modeling path. At this second level, a very plausible (but rarely explicitly admitted) criterion is the sheer economy of time and resources. When developing a model to account for new data, the decision not to alter its basic ingredients (i.e., ingredients incorporated at early stages) and rebuild it along another modeling path is understandably driven by pragmatic considerations. There is, in other words, some kind of inertial effect: One just keeps going with the model at hand. This effect is all the more pronounced when a model is developed over many years and involves the expertise of various people. A newcomer (typically a PhD student) in the model-building enterprise does not usually have the time and the knowledge to investigate whether other modeling paths could do better.

Regarding the first level of model selection, the role of pragmatic constraints seems to be less dominant. But from a practitioner's point of view, it is hard to come up with explicit criteria supplementing pragmatic considerations.[7] It would certainly be hard to deny that individual tastes, intuitions, tacit allegiance to dominant views, and so forth play a role in the decision to incorporate one submodel rather than another. But to be more substantial, such claims would call for sociological and psychological scrutiny that would take us beyond the scope of this article. In any case, what matters for our epistemological purpose here is the consequences, as regards the issue of epistemic credential, of this model-specific form of permanent underdetermination that follows from path dependency and plasticity.

Concluding Epistemological Lessons

When discussed with scientists, the underdetermination of theories by evidence does not appear as a big threat to the purpose of their work, not only because actual cases of permanent underdetermination are rather rare (e.g., see Kitcher, 2001, chap. 3), but also because underdetermination leaves untouched many virtues of theories (such as their explanatory and predictive power). Only realists' expectations about theories are threatened by underdetermination. In contrast with underdetermination at the level of theories, persistent incompatible model pluralism (the other face of the model-specific form of underdetermination put to the fore in the previous sections) thwarts the primary and essential goal of this kind of modeling enterprise. Recall that for astrophysicists and cosmologists, one of the primary epistemic functions of computer simulations is to produce data and insights on aspects of a real-world system for which observations are very scarce or inexistent, and then to use them to test various theoretical hypotheses. But as shown above, there is a tension between this realistic goal and the limits set by path dependency and plasticity on the possibility of validating the results obtained: The more composite a model gets to be realistic, the more you lose control of its validation. The tension is manifest when path dependency and plasticity lead to persistent incompatible pluralism. The very fact that several incompatible but equally empirically successful models continue to coexist in the long run calls into question their realistic virtues.

The tension is, however, no less present—albeit certainly easier to overlook—when, for practical, contingent reasons, only one dominant model exists. As we have seen when discussing the cosmological simulation, in such cases, path dependency also undermines our confidence in the realistic virtues of the modeling enterprise. In the specific case of the Millennium Run, epistemological prudence is all the more called for because the outcomes of the simulation are presented in a visual way and at a level of detail and scale that makes them easily mistakable for real observations. It is therefore essential to keep in mind that the story told by the simulation is actually only one plausible story among several other plausible stories that could have been told had the modeling path been different. This caveat is conspicuously absent from the scientific literature and from the presentations of scientific results written for the general public. If (epistemologically inclined) scientists readily warn (and rightly so) against "confusing computer simulations of reality with reality itself, when they can in fact represent only a highly simplified and stylized version of what *actually* is" (Ellis, 2006, p. 35, [italics added]), they unfortunately much more rarely warn against the distinct, additional pitfall of confusing simulations of reality with reality itself, when they can in fact represent only a highly simplified and stylized version of what *possibly* is. All this suggests, on a more general level, that the prospects for realism (as opposed to the more modest epistemic aim of empirical adequacy) seem rather dim in astrophysics and cosmology. It is worth noting that Hudson (2007) has come to a very similar conclusion, albeit on very different grounds (the article deals with discordant experimental results obtained in WIMPs detection experiments).

One last precision might be in order to conclude. My call for epistemological prudence should not be interpreted as a dismissal of the scientific enterprise consisting of developing purportedly realistic models of complex astrophysical phenomena. Rather, it simply emphasizes the necessity of reassessing the epistemic goals actually achieved, in light of the limited possibilities of validation. It is undoubtedly fascinating and useful to learn so much about plausible physical worlds, so long as one keeps in mind that those plausible worlds might be very different from our world. Otherwise put, computer models and simulations might undoubtedly be valuable as tools of intelligibility[8] by telling us plausible stories about real-world systems. But to hold them and present them in a realistic way, as windows through which we can observe the natural world, is often not legitimate in astrophysics and cosmology.

Notes

1. Images and movies associated with the press release of the Millennium Run can be found at http://www.mpa-garching.mpg.de/galform/press or in Ruphy (2005).

2. For a critical, in-depth look at the credentials of the inflation model, see Earman and Mosterin (1999).

3. For more on the role of imagination in model building, see Morgan (2004), who draws a parallel between the work of early astronomers and modeling in economics in the 20th century.

4. For more on the epistemic functions of galactic models, as well for a brief history of their development, see Ruphy (1996).

5. A somewhat similar point concerning the holist limitation of the testability of astrophysical hypotheses has been made recently by Vanderburgh (2003), albeit not at the level of submodels involved in computer modeling. What Vanderburgh's key notion of the "dark matter double bind" shows is that general assumptions such as the existence of a large quantity of dark matter and the validity of general relativity at scales larger than the solar system scale cannot be tested independently of one another.

6. My answer to that question will not take into account the epistemologically uninteresting fact that a galactic model may stop being developed for nonepistemic reasons, such as shortage of funds or career moves.

7. Commonly acknowledged pragmatic constraints on computer simulations include, for instance, issues of solvability (see Humphreys, 1995; Sismondo, 1999). But note that although this kind of pragmatic constraint is an important aspect of computer modeling, it differs, as the two case studies should have made clear, from the (more invisible) kind of constraint that the notion of "path dependency" purports to highlight.

8. See Machamer, Darden, and Craver (2000) for a discussion of the notion of intelligibility.

References

Bailer-Jones, D. (2000). Modelling extended extragalactic radio sources. *Studies in History and Philosophy of Science Part B: History and Philosophy of Modern Physics*, *31B*, 49-74.

Cohen, M. (1995). The displacement of the sun from the galactic plane using IRAS and FAUST source count. *Astrophysical Journal*, *444*, 874-878.

Duhem, P. (1954). *The aim and structure of physical theory* (P. P. Wiener, Trans.). Princeton, NJ: Princeton University Press.

Durrer, R., Kunz, M., & Melchiorri, A. (2002). Topological defects and structure formation. *Physics Reports*, *364*, 1.

Earman, J., & Mosterin, J. (1999). A critical look at inflationary cosmology. *Philosophy of Science*, *66*, 1-49.

Ellis, G. (2006). *Issues in the philosophy of cosmology*. Mathematics Department and Applied Mathematics, University of Cape Town, South Africa. Available from http://xxx.arxiv.org/abs/astro-ph/0602280

Hartmann, S. (1996). The world as a process: Simulations in the natural and social sciences. In R. Hegselmann, U. Müller, & K. Troitzsch (Eds.), *Modelling and simulation in the social sciences from the philosophy of science point of view. Theory and Decision Library* (pp. 77-100). Dordrecht, Netherlands: Kluwer.

Hudson, R. (2007). Annual modulation experiments, galactic models and WIMPs. *Studies in History and Philosophy of Science Part B: History of Modern Physics*, *38*, 97-119.

Humphreys, P. (1995). Computational science and scientific method. *Minds and Machines*, *5*, 499-512.

Kitcher, P. (2001). *Science, truth and democracy*. Oxford, UK: Oxford University Press.

Machamer, P., Darden, L., & Craver, V. F. (2000). Thinking about mechanisms. *Philosophy of Science*, *67*, 1-25.

Mitchell, S. (2002). Integrative pluralism. *Biology and Philosophy*, *17*, 55-70.

Morgan, M. S. (2004). Imagination and imaging in model building. *Philosophy of Science*, *71*, 753-766.

Morgan, M. S., & Morrison, M. (Eds.). (1999). *Models as mediators*. Cambridge, UK: Cambridge University Press.

Parker, W. (2006). Understanding pluralism in climate modeling. *Foundations of Science*, *11*, 349-368.

Robin, A. C., & Crézé, M. (1986). Stellar populations in the Milky Way—A synthetic model. *Astronomy and Astrophysics*, *157*, 71-90.

Robin, A. C., Reyle, C., Derriere, S., & Picaud, S. (2003). A synthetic view on structure and evolution of the Milky Way. *Astronomy and Astrophysics*, *409*, 523-540.

Ruphy, S. (1996). *Contribution à l'étude des inhomogénéités du disque galactique à l'aide des données DENIS* [Studies of inhomogeneities in the galactic disc based on DENIS star counts]. Unpublished doctoral dissertation, Paris VI/Observatoire de Paris.

Ruphy, S. (2005). L'Univers en 28 jours chrono [28 days to compute the history of the universe]. *La Recherche*, *21*, 72-77.

Ruphy, S., Epchtein, N., Cohen, M., Copet, E., de Batz, B., Borsenberger, J., et al. (1997). Stellar populations and inhomogeneities of the galactic plane from DENIS star counts. *Astronomy and Astrophysics*, *326*, 597-607.

Ruphy, S., Robin, A. C., Epchtein, N., Copet, E., Bertin, E., Fouque, P., et al. (1996). New determination of the disc scale length and the radial cutoff in the anticenter with DENIS data. *Astronomy and Astrophysics*, *313*, L21-L24.

Schweber S., & Wächter, M. (2000). Complex systems, modelling and simulation. *Studies in History and Philosophy of Science Part B: History and Philosophy of Modern Physics*, *31*, 583-609.

Sismondo, S. (1999). Models, simulations, and their objects. *Science in Context*, *12*, 247-260.

Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., et al. (2005). Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, *435*, 629-636.

Vanderburgh, W. L. (2003). The dark matter double bind: Astrophysical aspects of the evidential warrant for general relativity. *Philosophy of Science*, *70*, 812-832.

Wainscoat, R. J., Cohen, M., Volk, K., Walker, H. J., & Schwartz, D. E. (1992). A model of the 8-25 micron point source infrared sky. *Astrophysical Journal Supplement Series*, *83*, 111.

Winsberg, E. (2003). Simulated experiments: Methodology for a virtual world. *Philosophy of Science*, *70*, 105-125.

Stéphanie Ruphy, when not busy making the most of Parisian life, thinks about issues in general philosophy of science revolving around the unity/disunity of science debate: reductionism, taxonomic pluralism, feminist criticism of the value neutrality of science, and the epistemology of computer models and simulations. Reflecting interests in both science and the humanities, she has a PhD in philosophy from Columbia University and a PhD in astrophysics from the Paris Observatory. After a few hectic years as a scientific journalist in Paris, she settled down for the quieter joys of academic life as an associate professor in the Philosophy Department of the University of Provence. Contact: Département de philosophie, Université de Provence (Aix-Marseille I), 29, avenue Robert Schuman, 13621 Aix-en-Provence, France; +33 (0)4 42 95 33 21 (t); stephanie.ruphy@wanadoo.fr.