

Fitting Paillard's Three State Model Parameters via CMA-ES and GA.

Andrew James Reagan^{1,*} and Morgan R. Frank^{1,†}

¹ *Department of Mathematics and Statistics, Center for Complex Systems, Computational Story Lab,
& the Vermont Advanced Computing Core, University of Vermont, Burlington, VT, 05401*

(Dated: September 17, 2013)

The explanatory value of simple climate models relies upon their ability to mechanistically explain past variation while providing a reasonable fit to available data. We investigate the proper fitting of Paillard's three state model with past data to test its accuracy directly. The parameter fit, our main effort, is achieved via a covariance matrix adaptation evolution strategy (CMA-ES) and a general genetic algorithm (GA). We find that Paillard's model's explanatory value is mainly empirical, as the model is robust to parameter changes and the fit to data from $\delta^{18}O$ records has minimum RMSE 1e-5.

PACS numbers:

INTRODUCTION

Understanding the past climate of the earth allows for us to begin to understand what effects we are having upon it, and to predict what the future will bring. From the analysis of isotope ratios in ice cores and marine sediment cores, we obtain knowledge of past temperature from the ratio of $\delta^{18}O$, a proxy for combined temperature and global ice volume. This time series is constructed over the past 2 Myr, with higher values indicating a colder climate as the ocean is enriched with $\delta^{18}O$ during the colder conditions. The variation in the orbital parameters of our planet affect the amount of solar insolation recieved, and the timescale variation of these quantities is one possible forcing for past climatic change. This was hypothesized as early as the nineteenth century, and further elucidated by Milankovitch, whose name the study of these variations bears. The eccentricity of our elliptical orbit is changing slowly in time, as is the orientation of the planet's spin axis and the longitude closest to the sun at the vernal equinox. We refer to these parameters as e , ϵ and $\bar{\omega}$, respectively.

The main questions arising about this time period from Dijkstra [1] that any satisfactory theory of the Pleistocene Ice Ages should explain are:

- Why did the glacial-interglacial cycles appear in the Pleistocene?
- Which processes in the climate system caused the glacial-interglacial changes in global mean temperature, $p_{CO_2}^o$, and ice sheet extent?
- What caused the transistion (the MPT) from the 41-kyr world to the 100-kyr world about 700 kyr ago?

The Milankovitch theory of orbital forcing is now the 'null hypothesis' answer to these questions, and various models are available that relate this orbital forcing to global mean temperature and ice sheet extent have been proposed. The fit between the changing

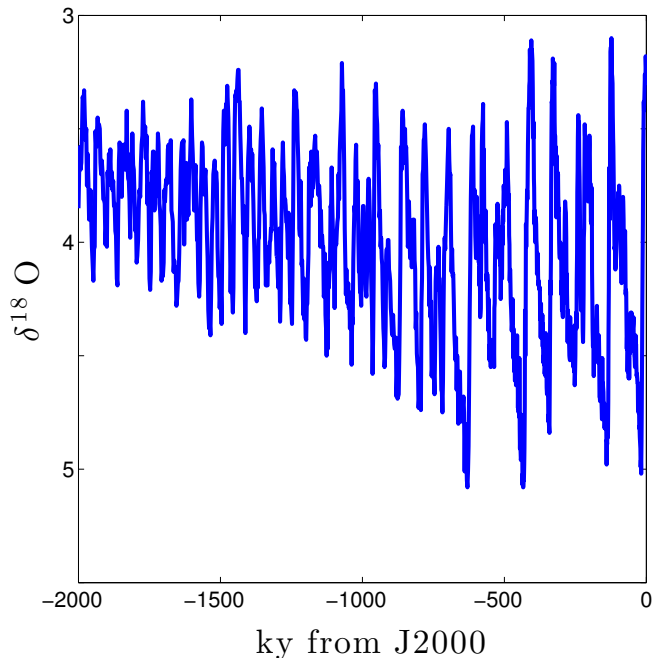


FIG. 1: The LR04 benthic $\delta^{18}O$ stack over the Pleistocene, constructed by the graphic correlation of fifty-seven globally distributed benthic $\delta^{18}O$ records versus time (data from Lisieki [2]).

solar insolation and $\delta^{18}O$ has been demonstrated not to be linear, opening the door for more complex explanations. The potentially important feedbacks of ice-albedo, height-mass balance, load-accumulation, and precipitation-temperature have all been combined into quasi-equilibrium and internal oscillation models, with the latter successfully demonstrating a reasonable fit to the data. In addition to these more complex models is a switching equilibria model from Didier Paillard that fits the data empirically, and does "remarkable" considering the simplicity of the model.

In the present work we attempt a rigorous fit of Paillard's model to data using the tools of evolutionary com-

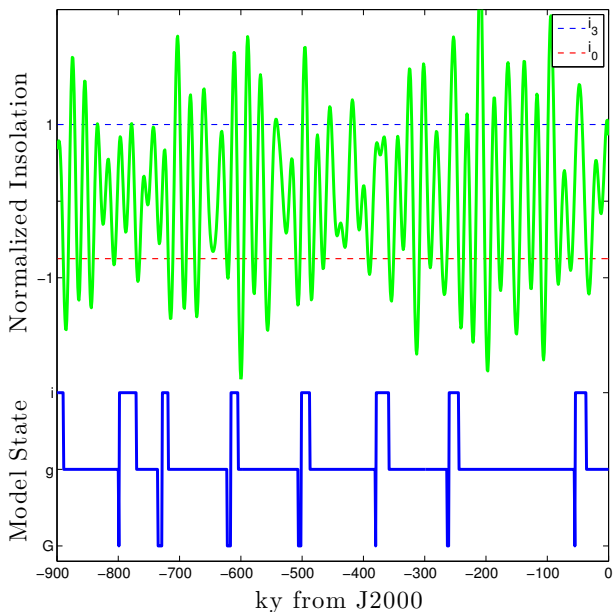


FIG. 2: The discrete Paillard model.

putation.

PAILLARD'S MODEL

With some difficulty understanding his original paper [3], we reconstruct Paillard's model for both versions that he provides. The following two subsections paraphrase [3] directly.

Discrete model

The model has three states called **i** (interglacial), **g** (mild glacial) and **G** (full glacial). This model undergoes an **i-g** transition as soon as the insolation falls below i_0 . An **g-G** transition happens when the ice volume exceeds a threshold v_{\max} and finally a transition **G-i** occurs when the insolation increases above i_1 . These transitions are assumed to be one-way, and the only ones allowed are those just discussed. The **g-G** transition has the constraints that it will happen after some time t_g and with insolation lower than some level i_3 . If t_g is exceeded but i_3 is also exceeded, then the **g-G** transition occurs as the next insolation decrease: when insolation falls below i_2 .

The part which was made least clear was whether the insolation exceeding i_3 necessitates a reset of the time in state **g** until t_g is reached. To achieve similar results, we assume that the answer is yes, the time does reset.

Differential equation model

This model is an extension of the discrete model. There are still three regimes but now we define the ice volume v as a differential equation

$$\frac{dv}{dt} = (v_R - v)/\tau_R - F/\tau_F.$$

where R is the current climate regime ($R = \mathbf{i}, \mathbf{g}, \mathbf{G}$), v_R are the reference ice volumes for the different regimes, F is the forcing and $\tau_{R,F}$ are the time constants. The transitions are the same for **i-g** and **G-g** but **g-G** now happens when $v \geq v_{\max}$, some threshold value (without any condition on the insolation). Paillard also "normalizes the ice volume to unity": $v_g = v_G = v_{\max} = 1$ and $v_i = 0$. For the finale of model weirdness, the forcing F is defined as a truncated insolation signal with truncation function f as

$$f(x) = \frac{1}{2} \left(x + \sqrt{4a^2 + x^2} \right)$$

and setting $a = 1$.

COMPUTING INSOLATION

The only input to the model is the solar insolation, so we are particularly careful to obtain accurate values. Increasingly accurate integrations of the state of the solar system have been performed dating from the time of Lagrange in 1781 through Milankovich's use of Pilgrim's 1904 computations to establish his theory of past climate change to today's direct numerical integrations [4].

Using the orbital parameters of eccentricity e , obliquity ϵ , and longitude of the perihelion $\bar{\omega}$ computed from an integration function `insola.f` provided from Laskar [4] we can reconstruct the insolation at any time. The following derivation is due to Berger [5].

The insolation Q at a given location and time can be written as a function of the solar constant S_0 , the distance from the Earth to sun relative to the average R_0/R_E the solar zenith angle Θ :

$$Q = S_0 \frac{R_0^2}{R_E^2} \cos(\Theta).$$

To solve for the average insolation during one day at a given latitude ϕ , we need to express the solar zenith angle Θ in terms of ϕ and the known orbital parameters. Taking δ as the latitude of the point directly subsolar and h as the hour angle, we use the spherical law of cosines to write

$$\cos(\Theta) = \sin(\phi) \sin(\delta) + \cos(\phi) \cos(\delta) \cos(h).$$

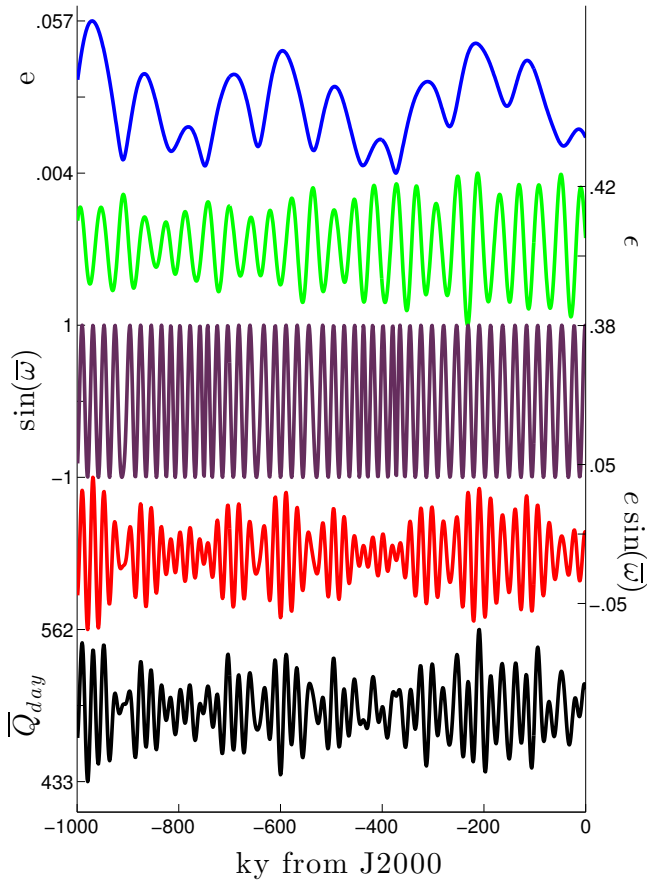


FIG. 3: The average daily solar insolation \bar{Q}_{day} is computed on the summer solstice at 65°N with the provided e , ϵ , and \bar{w} .

Now to have the average daily insolation \bar{Q}_{day} we take the integral of Q from h_o to $-h_o$ where h_o is the hour angle for which the sun first rises. Since we are concerned with the $\phi = 65^\circ\text{N}$, the sun always rises and sets so we won't consider the cases for the sun neither rising nor setting. Integrating, we have

$$\bar{Q}_{\text{day}} = \frac{S_0}{\pi} \frac{R_0^2}{R_E^2} [h_o \sin(\phi) \sin(\delta) + \cos(\phi) \cos(\delta) \sin(h_o)].$$

Defining $\theta = 0$ to be the vernal equinox, we can write δ and R_0/R_E as a function of the orbital parameters in a simple form at the summer solstice ($\theta = \pi/2$):

$$\delta = \epsilon \quad \text{and} \quad \frac{R_0}{R_E} = 1 + e \sin(\bar{w}).$$

EVOLUTIONARY ALGORITHM METHODS

We attempt to search the parameter space within the limits suggested by Paillard [3] in hopes of better fitting the both the discrete and the ODEE models to

the $\delta^{18}\text{O}$ data by using a genetic algorithm (GA) and covariance matrix adaption evolutionary scheme (CMA-ES) to evolve possible parameters. CMA-ES and GA are well-known and documented evolutionary algorithms for evolving real-valued vectors that minimize a particular fitness function when evaluated. For this investigation, we chose to minimize the root mean square error of the normalized model compared to the $\delta^{18}\text{O}$ data.

Datasets for fitting

We obtained the necessary data sets for running our experiments. Our $\delta^{18}\text{O}$ data comes from Tiedermann et al. [6] as the result of analyzing San ice core sample taken from up to 273.80 m in 1986. A simple linear transformation is used to scale from depth to the age of the samples. The ice core sample was taken in the North Atlantic Ocean. While this may not be the exact dataset used by Paillard, we believe the data contained here should be sufficiently similar and relevant in our efforts to fit Paillard's models to historical climate behavior.

We employ the same insolation dataset used by Paillard in [3]; namely, we use the data provided by A. Berger in [5]. Berger obtains this curve representing the earth's atmospheric insolation about the latitude 65°N by modeling the earth's rotation about the sun. His calculations account for such variables as the angle of the earth's axis from perpendicular to the sun, and eccentricity. This insolation curve acts as the forcing for our model.

Models in more detail

Paillard [3] suggests two models for obtaining the $\delta^{18}\text{O}$ curves from atmospheric insolation data. These models represent the transformations of three different glacial states: i (interglacial), g (mild glacial), G (full glacial). Both models allow us to impose limits on the insolation, the ice volume that determines the glacial state, and the expected time spent in each glacial state. In both models, it is assumed that the only three possible transitions exists. A $i \rightarrow g$ (interglacial \rightarrow mild glacial) transition occurs when the insolation falls below a parameter i_0 . A $g \rightarrow G$ (mild glacial \rightarrow full glacial) transition occurs when a maximum ice volume v_{max} is exceeded. Finally, a $G \rightarrow i$ (full glacial \rightarrow interglacial) transition occurs when the insolation grows above a parameter.

Paillard first suggests a discrete model for obtaining the $\delta^{18}\text{O}$ curve from insolation data. This model has four insolation parameters. i_0 and i_1 represent the insolation values required for $i - g$ and $G - i$ transitions to occur, respectively. The other insolation parameters i_2 and i_3 identify a range of possible forcing values that would allow a $g - G$ transition to occur. A further parameter, τ_g , is a time constant determining the minimum number

of years for the model to spend in a mild glacial state. In an effort to reduce the search space, we adopt Paillard’s assumption of $\tau_g = 33$ kyr for our discrete model and attempt to evolve values for the parameters i_0 , i_1 , i_2 , and i_3 .

An ODE model is also presented by Paillard, which is given by

$$\frac{dv}{dt} = \frac{v_R - v}{\tau_R} - \frac{F}{\tau_F}$$

where v is the ice volume, v_R and τ_R are the volume and the time constants, respectively, corresponding to the glacial state. τ_F is a time constant associated with a forcing parameter F . The forcing parameter F is given through a truncation of the insolation data described by

$$F = \frac{1}{2}(x + \sqrt{4a^2 + x^2})$$

with x being the insolation and a tunable compression parameter a . i_0 and i_1 are the same as in the discrete model, and the other insolation parameters are omitted. Again, we reduce our search space by adopting the assumption presented by Paillard that $\tau_i = 10$ kyr, $\tau_g = \tau_G = 50$ kyr, $\tau_F = 25$ kyr, and $a = 1$. We attempt to provide good choices for the remaining insolation parameters i_0 and i_1 .

Evolutionary Algorithms

Both of the evolutionary algorithms we used are designed to search through the space \mathbb{R}^n with a population of n length real-valued vectors with each entry corresponding to a parameter in our model. Through common genetic mechanisms, like reproduction and mutation, we are able to evolve vectors representing potentially good parameter choices for the insolation parameters in each model to fit the $\delta^{18}O$ data. Both algorithms require a fitness function with which to assess how well the parameter choices represented by an individual vector represent the $\delta^{18}O$ data through the models. We employ a fitness function which integrates the model and then attempts to minimize the root mean square error of the resulting time series compared to the true $\delta^{18}O$ data. This fitness function allows us to apply selection pressure on the populations of vectors that results in competition within the population to fit the $\delta^{18}O$ data the best.

Covariance Matrix Adaption Evolutionary Scheme (CMA-ES)

CMA-ES is often a good initial choice of evolutionary algorithm for evolving real-valued vectors because the algorithm itself is not expensive and the algorithm often yields a reasonably good result after only a few

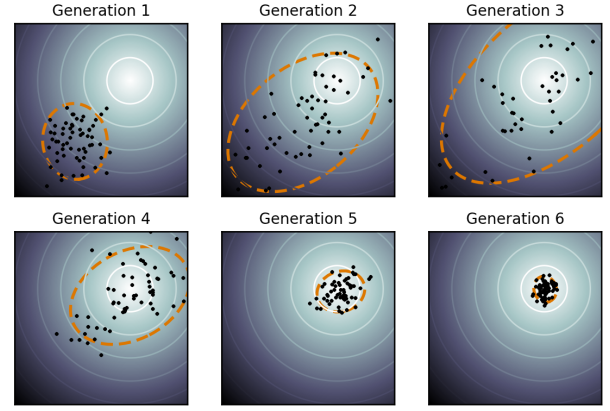


FIG. 4: A cartoon representing 6 generations in a CMA-ES experiment. Here, lighter colors indicate improved fitness. The dashed lines represent the influence of the covariance matrix.

generations. The algorithm begins with a uniformly distributed cloud of points in \mathbb{R}^n . Each point represents parameter choices for the models. We assess the fitness of each point and create a covariance matrix weighted by the fitness values. By taking the average of the fitness values, we can determine weighted center of the cloud of points, and we can determine a direction to move that point that will hopefully yield a better average fitness. Having relocated the average point, we create a new cloud of points around the average, however we now use the covariance matrix to determine how to distribute the individuals points. For example, if the population members all agree at a particular index, then they will have small covariances. This will cause the new cloud of points to be very narrow in that dimension since all of the population members are in agreement. On the other hand, if all of the individuals disagree in a particular dimension, then there will be large covariances representing a need to explore more in that dimension. Thus, the cloud of points should be wider in that dimension. We fix the number of generations to be 500 for our experiments. More information about CMA-ES can be found in [7].

Genetic Algorithm (GA)

GA’s are typically more expensive than CMA-ES, but allow for an increased wealth in ways for population members to combine information. GA’s often yield greater runtimes than CMA-ES but can also yield more robust explorations of parameter space. The algorithm begins with a population of real-valued vectors. Each vector is assessed for fitness. The fitter individuals in the population are chosen as candidates for reproduction in the form of crossover. Here, we use uniform crossover, which means that given two “parent” vectors from the population, we create a “child” vector by ran-

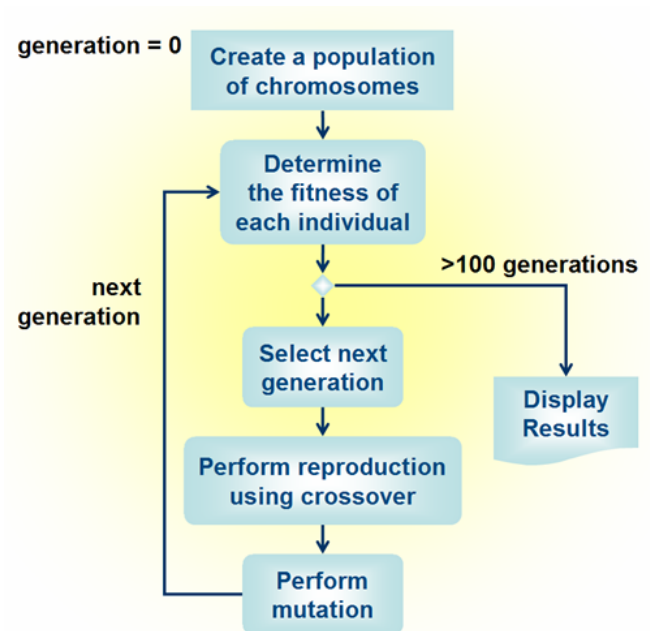


FIG. 5: A schematic explaining the basic steps in a genetic algorithm.

domly selecting indices to contain the value stored in that index by parent 1. The remaining indices are filled with the values stored in parent 2 at those indices. A second child vector is created by taking the same random indices, and swapping the roles of the parents. These two children vectors then replace their parents in the population. Individuals in the population are also subject to mutation. Mutation is performed on individuals where a small amount is added or subtracted from each index in an individual vector with some small probability. Crossover often leads the population to converge on good solutions, while mutation generally causes the population to explore parameter space. Striking a balance between these two forces is vital and often very difficult depending on the noise in the fitness landscape; that is, the fitness values associated with each point in the parameter space. For our experiments, we use a population size of 500 and calculate 500 generations.

RESULTS

Constructing the insolation from orbital parameters, Figure 3 depicts the reconstruction of insolation data back in time from Laskar’s data set. And verifying that our interpretation of the model is the same as Paillard intended, Figure 2 is a direct representation of the second figure of Paillard’s original paper [3].

Having run 100 CMA-ES and GA experiments, we discuss the fit of the best solutions here. Firstly, CMA-ES underperformed. None of the CMA-ES experiments

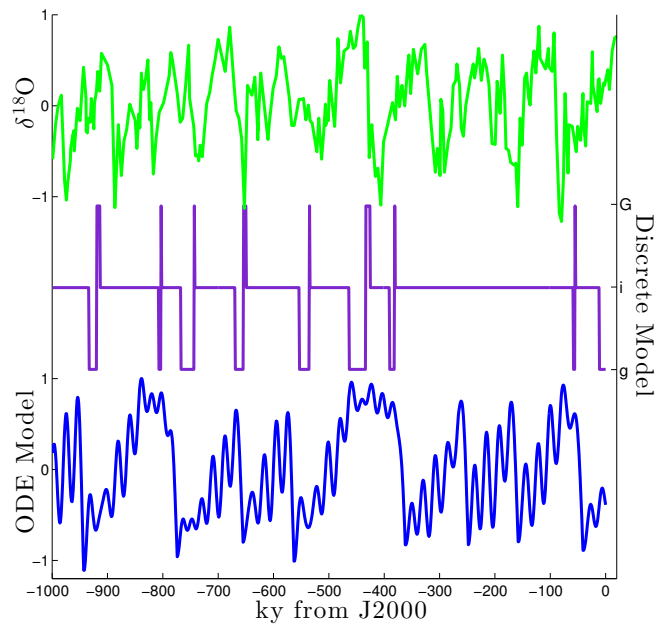


FIG. 6: The best result from the GA experiments attempting to fit parameters to Paillard’s discrete and ODE model, with the $\delta^{18}O$ data used for fitness.

yielded results worth noting here. This is likely due to unresolvable noise in the fitness landscape. We focus instead on the more promising results from our GA experiments.

Figure 6 demonstrates the results of our most promising GA experiments along with their root mean square errors (RMSE) from the $\delta^{18}O$ data. We should note that the particular parameter choices proposed by Paillard in [3] yield RMSE= 1.42 for the discrete model, and RMSE= 1.7186 for the ODE model. Figure 3 shows that our GA experiment ($i_0 = -0.7293$, $i_1 = -0.2287$, $i_2 = 0.0803$, and $i_3 = 0.9862$) was only narrowly outperformed by Paillard’s parameter choices for the discrete model, and that our GA experiment ($i_0 = -0.9677$ and $i_1 = 0.0294$) yields a better fitting ODE model than the model suggested by Paillard’s parameter choices for this model.

CONCLUSION

The fitted Paillard model of climatic variability makes a strong case for differing atmospheric states playing a role in the temperature variation in the $\delta^{18}O$ record. The methods of evolutionary computation are important in climate science to assist in model validation, an important step in understanding. With only forcing from changing solar insolation, this verifies that past climatic variation of larger time scales is a result of the variation of the Earth’s orbital parameters.

The authors would like to thank the poor chaps who

were out in the arctic drilling ice cores for their sacrifices of comfort in the name of science.

* Electronic address: `andrew.reagan@uvm.edu`

† Electronic address: `mrfrank@uvm.edu`

[1] H. A. Dijkstra, *Nonlinear Climate Dynamics* (Cambridge University Press, 2013), ISBN 0521879175.

[2] L. E. Lisiecki and M. E. Raymo, *Paleoceanography* **20**

(2005).

[3] D. Paillard, *Nature* **391**, 378 (1998).

[4] J. Laskar, P. Robutel, F. Joutel, M. Gastineau, A. Correia, B. Levrard, et al., *Astronomy & Astrophysics* **428**, 261 (2004).

[5] A. Berger, J. Guiot, G. Kukla, and P. Pestiaux, *Geologische Rundschau* **70**, 748 (1981).

[6] R. Tiedemann, M. Sarnthein, and N. J. Shackleton, *Paleoceanography* **9**, 619 (1994).

[7] T. Suttorp, N. Hansen, and C. Igel, *Machine Learning* **75**, 167 (2009).