How Should Mathematical Models of Geomorphic Processes be Judged?

Richard M. Iverson

U.S. Geological Survey, Vancouver, Washington

Mathematical models of geomorphic processes can have value as both predictive tools and precise conceptual frameworks. Well-posed mechanistic models have great conceptual value because they link geomorphic processes to universal scientific principles, such as conservation of energy, momentum, and mass. Models without this linkage (e.g., models based exclusively on cellular rules or empirical correlations) have less conceptual value but offer logical methodology for making practical predictions in some circumstances. Clear tests of the predictive power of mechanistic models can be achieved in controlled experiments, whereas natural landscapes typically have uncontrolled initial and boundary conditions and unresolved geological heterogeneities that preclude decisive tests. The best mechanistic models have a simplicity that results from minimizing assumptions and postulates, rather than minimizing mathematics, and this simplicity promotes conclusive tests. Optimal models also employ only parameters that are defined and measured outside the model context. Common weaknesses in geomorphic models result from use of freely coined equations without clear links to conservation laws or compelling data, use of fitted rather than measured values of parameters, lack of clear distinction between assumptions and approximations, and neglect of the four-dimensional (space + time) nature of most geomorphic processes. Models for predicting landslide runout illustrate principles and pitfalls that are common to all geomorphic modeling.

> "The real voyage of discovery consists not in seeking new landscapes but in having new eyes" [*Marcel Proust,* translated 1981].

1. INTRODUCTION

Assessment of mathematical models is a critical part of the scientific method, because mathematical models represent hypotheses cast in their most precise form. When expressed in the language of mathematics, well-posed hypotheses yield quantitative predictions that are subject to unequivocal tests. However, as noted by *Oreskes et al.* [1994, 2001], model testing in the Earth sciences can present special difficulties. Many Earth processes involve settings, rates, or scales that

are not accessible to direct observation or experimental replication, and uncertain initial conditions or undocumented heterogeneities in geological materials may limit the predictive power of mathematical models even if the equations are essentially correct. In geomorphology, however, these difficulties can be overcome to a considerable extent.

Geomorphology differs from many geosciences because it involves materials and processes that are accessible to direct measurements and manipulative experiments. The time and length scales characteristic of a flowing river or a failing slope are not far removed from human experience: a river can alter its course in a single flood, and a slope can fail before our eyes. This accessibility means that mathematical modeling of geomorphic processes can benefit from relatively stringent quantitative tests that are similar to those of classical physics

and dissimilar from those of deep-Earth and deep-time geosciences, in which material properties and initial and boundary conditions must be inferred rather than measured or controlled. However, despite the ready accessibility of geomorphic processes and the recent proliferation of mathematical models that describe them, agreement is lacking on protocols for model assessment. Opinions vary widely about the purpose of mathematical models in geomorphology and the standards by which such models should be judged.

In this paper I suggest criteria for judging the merit of geomorphic models, with an emphasis on mathematical models of geomorphic processes that operate on time scales directly observable by humans. Forecasting the effects of geomorphic processes on observable time scales is critical because it provides a basis for informed decisions about land management and hazard assessment, and these applications gain urgency and importance as human habitation and exploitation of landscapes grow. In addition, applied problems highlight the need for rigorous standards of model appraisal; the relatively relaxed standards that are sometimes unavoidable in deep-Earth and deep-time geosciences are unsatisfactory when human welfare and ecosystem viability in the present century are at stake. Sound models of observable geomorphic processes also provide the firmest foundations for models of long-term, unobservable landscape change: just as biologists cannot draw strong inferences about evolution of life forms without observing and understanding the workings of molecular genetics, geomorphologists cannot draw strong inferences about evolution of landscapes without understanding the workings of observable erosion and sedimentation events [cf. *Platt*, 1964].

Two overarching principles, which concern *prediction* and *parsimony*, can guide development and appraisal of mathematical models of observable geomorphic processes. The prediction principle has been widely acknowledged and accepted in all sciences: models should forecast the outcome of as-yet unobserved events. Although relevance of this principle is clear, its application is muddied by questions about what constitutes a valuable geomorphic prediction [*Shreve*, 1979; *Haff*, 1996]. For example, does a prediction of experimental results have value that transfers to the field? How valuable are *a priori* model predictions in comparison to calibrated fits? What are the relative merits of mechanistic models versus empirical statistical models as predictive tools? What distinguishes phenomenological predictions from chronological predictions aimed at foretelling the future, which may involve sequences of contingent events?

The principle of parsimony is more subtle than that of prediction, but the significance of parsimony resonates in writings of Newton and Einstein, who emphasized that models should minimize the number of axioms or postulates invoked

to explain the physical world. Modelers ought not to invent new laws where none are warranted, and models should employ the fewest possible assumptions—even if this paucity of laws and assumptions imposes difficult mathematics or experimental tests. It is important to recognize that accurate and elegantly simple concepts can spawn complex mathematics, and that simple algebraic equations can express tortuous concepts that violate physical laws.

The next two sections of this paper consider the prediction and parsimony principles in light of underlying philosophical and methodological issues, and a subsequent section illustrates the principles in a more concrete way by discussing models for predicting landslide runout. The final section of the paper distills the content of previous sections into questions for modelers and model users to consider.

2. APPRAISAL OF MODELS AS PREDICTIVE TOOLS

2.1. Phenomenological Versus Chronological Prediction

To evaluate models as predictive tools, it is useful to distinguish between phenomenological and chronological prediction. The traditional purview of science is phenom-enological prediction—for example, Einstein's famous prediction of the equivalence of energy and mass—whereas foretelling the future has traditionally been the province of clerics and mystics [*Raymo*, 1998; *Sarewitz and Pielke*, 2000]. In the case of regularly cyclical phenomena that involve negligible energy dissipation (e.g., planetary motion), chronology can be inferred from phenomenology, but many phenomena studied in Earth sciences are nonlinear, highly dissipative, and contingent on prior events—attributes that break the clear connection between phenomenology and chronology. Nonetheless, in recent decades Earth-science modelers have focused much effort on prediction of chronological events, such as the timing of earthquakes and progression of climate change. The great difficulty of successfully predicting future events appears deeply rooted in the irreversibility of Earth processes and their sensitivity to contingencies. This sensitivity does not imply that useful predictions are impossible, however. Rather, it implies that models will yield the most successful and testable predictions if they focus chiefly on phenomenology rather than on chronologies that involve sequences of contingent events.

Successful phenomenological predictions allow observers to anticipate the outcome of specific events, even if the exact timing or sequence of those events cannot be foretold with much accuracy. As a logical construct, a phenomenological prediction states that if a particular condition exists or event occurs, then another condition or event will occur as a consequence. The prediction follows the precepts of classical physics and distinguishes cause and effect [cf. *Cleland*,

2001]. For example, a geomorphic prediction might state that if a particular slope with a particular antecedent pore-pressure distribution is subject to rainfall at a rate exceeding 1 cm/hour for ten hours, the rain will cause the slope to fail [e.g., *Iverson*, 2000]. If such a prediction is accurate, its utility is clear, but the prediction does not generally foretell the future because it requires specification of an initial state (e.g., antecedent pore pressures) and provides no means of anticipating the onset of triggering rain. Despite this type of limitation, phenomenological predictions can provide important information on *rates* of processes. Such rate predictions may explicitly include time, but they differ from foretelling the future because they assume knowledge of a relevant initial state, independent of contingencies.

The difference between phenomenological and chronological prediction is more than philosophical, because it implies practical limits on prediction of geomorphic change that involves sequences of contingent events. In such predictions a pivotal issue is whether a state measured at some time $t = t_0$ contains sufficient information to accurately forecast a subsequent state at $t = t_0 + \Delta t$. For example, in the case of slope failure triggered by rainfall, an initial state measured at time $t = 0$ might contain enough information for accurate prediction of landsliding during a rainstorm at $t = 1$ week but too little information for accurate prediction of landsliding during an identical rainstorm at $t = 100$ years. As time proceeds, a deterministic phenomenological prediction can degenerate into a poorly constrained chronological prediction owing to the cumulative effect of nonlinearities and contingencies that cause memory of the measured initial state to fade [e.g., *Lighthill*, 1994].

Formulation and testing of models that make phenomenological predictions of observable geomorphic events involves many challenges, which are described below. Conclusive testing of geomorphic models that foretell the future on time scales unobservable by humans (or decipher the unobserved past by forward modeling in time from an assumed initial condition) is perhaps impossible [*Cleland*, 2001], and is beyond the scope of this paper.

2.2. Testing Predictions with Data

Quantitative observations of a phenomenon (i.e., data) are rightly viewed as the ultimate test of the predictive capability of any mathematical model, but in geomorphology, where opportunities for data collection abound, what data are best suited to this task? The traditional view in geosciences is that the best test of a model is provided by data collected in the field, where processes operate at full complexity, unfettered by artificial constraints. This field-test paradigm involves assumptions about the nature and purpose of models themselves, and the paradigm may be inappropriate in the case of mechanistic geomorphic models.

If the purpose of a model is exclusively prediction of the final outcome of a phenomenon, without providing insight to intermediary mechanisms that link causes and effects, field tests are probably best. For example, field tests of a statistical regression model that relates suspended sediment load to stream discharge are appropriate because the regression model makes no pretense of mechanistic understanding; its sole purpose is forecasting the value of a dependent variable (sediment load) as a function of an observed independent variable (discharge). In such a model the only postulate about cause and effect involves designation of one variable as independent and the other variable as dependent. Two calibrated parameters, the regression-line slope and intercept, are employed to relate the two variables, but no deeper significance is ascribed to the calibrated slope and intercept.

Models that are more ambitious aim to predict not only final outcomes of phenomena but also mechanistic linkages between causes and effects. To clarify such linkages, model input generally includes specification of an initial state that isolates the phenomenon in time as well as boundary conditions that isolate it in space. Field data from uncontrolled settings are poorly suited for testing such models, because uncontrolled field settings offer few opportunities for establishing unambiguous initial conditions and domain boundaries.

Geological heterogeneity imposes an additional difficulty if field data are used to test mechanistic models. All such models contain parameters, some of which may be nearly constant (e.g., gravitational acceleration at Earth's surface) and others of which may vary significantly within the model domain (e.g., hydraulic conductivity of soil). Heterogeneities that are unresolved by parameter measurements can bedevil any test of predictions because they make it difficult to distinguish the effects of parameter uncertainty from those of erroneous model logic [*Oreskes et al.*, 1994].

For similar reasons, bringing model predictions into agreement with field data through adjustment of parameter values does not constitute a satisfactory test. The predictive and explanatory power of a model diminishes as dependence on adjustable coefficients and calibrated parameters increases. When carried to an extreme, calibration of models with numerous poorly constrained parameters differs little from fitting a polynomial function with N-1 degrees of freedom to a set of N scattered data points; the fit is impressive but the result provides little insight [*Jeffreys and Berger*, 1992].

Rapidly advancing sciences seldom operate with paradigms that emphasize natural experiments in uncontrolled, heterogeneous environments. If high-energy physics employed the paradigm that elementary particles should be

observed in the wild rather than in the highly orchestrated environment of an accelerator, or if molecular biology employed the paradigm that inferences about DNA sequences should be drawn from observations within complex, living organisms, the stunning advances of these fields within the past half century would be implausible. If geomorphology is to make similarly rapid advances, a new paradigm may be required: mechanistic models of geomorphic processes should be tested principally with data collected during controlled, manipulative experiments, not with field data collected under uncontrolled conditions.

2.3. Hierarchy of Data for Model Tests

The suggestion that controlled experiments ought to supercede field measurements as the principal means of model testing in geomorphology contradicts longstanding practice, as noted by *Paola et al.* [2001]. Field observations furnish the inspiration for new models, and allow multiple working hypotheses to be formulated and evaluated in the qualitative manner described by *Chamberlin* [1897], but field measurements obtained under uncontrolled conditions yield decisive tests of only the least ambitious quantitative models, such as statistical models that provide no mechanistic linkage between cause and effect. Field measurements in situations with unconstrained initial and boundary conditions and undetermined parameter values generally leave considerable room for declarations and debate about valid interpretations. However, as summarized by *Shapiro* [1986], "Science works neither by pronouncement nor consensus, but rather by experiment."

Controlled experiments can provide a high standard of model testing, but in geomorphology, what constitutes a useful controlled experiment? Good experiments seldom attempt to duplicate complex conditions that exist in the field. Rather, the experimenter establishes initial and boundary conditions that are advantageous for testing hypotheses of interest. Reproducibility and independent constraints on all relevant parameter values are other key attributes of an ideal controlled experiment, whether it is conducted in a laboratory or the field. Manipulative field experiments (such as the 1996 controlled flood in Grand Canyon, Arizona [*Webb et al.*, 1999]) generally have less reproducibility and weaker constraints on parameter values and boundary conditions than do laboratory experiments, but offer the advantage of full scale. In contrast, reproducible laboratory experiments offer full control over model inputs, but scaling is a critical issue. Appropriate scaling can be difficult to achieve in geomorphological experiments, especially if they involve water, wherein surface tension and viscosity can produce disproportionately large effects at miniature scales.

Nonetheless, well designed, properly scaled laboratory experiments commonly provide the most compelling results in science: successful isolation of the influence of a key parameter, initial condition, or boundary condition offers incontrovertible insight.

Critics commonly argue that isolation of phenomena is artificial, and that this reductionist approach does not capture the rich complexity of open geomorphic systems in nature [e.g., *Baker*, 1996; *Werner*, 1999]. Although this criticism has obvious relevance, history shows that it does not constitute a valid condemnation of reductionist science. Virtually all natural phenomena studied in all fields of science occur in open systems, yet it is precisely the reductionist isolation of individual phenomena that has yielded nearly all clear advances.

The power of reductionism and advantages of controlled experimentation may be difficult to embrace in a subject such as geomorphology, in which evolving interconnections and emergent behavior are visibly evident. However, this conspicuous complexity ought to motivate, not deter, investigations of simple cases that yield clear insight. Consider a well-known analogy from classical physics: the complex behavior of nonlinear oscillators (archetypes of chaotic dynamical systems) is illuminating chiefly because the simpler behavior of linear oscillators is understood so well [e.g., *Pippard*, 1985]. The signal from a nonlinear oscillator might seem like unintelligible noise if a clear understanding of the regular cyclicity of linear oscillators were unavailable. Similar benchmarks of understanding are needed to establish a context for complexity in geomorphology, and reductionist experiments and model tests provide the surest means of establishing such benchmarks.

An even higher standard of model testing is available than that afforded by a carefully controlled experiment, however, and this higher standard reveals the full power of classical scientific methods. The standard is characterized well by a rhetorical question posed to all modelers by *Feynman* [1994], which can be paraphrased as, "what *else* does your model predict?" Feynman's point was that models should be pushed to their limits to examine all possible predictions and outcomes, and not merely those that pertain to a particular observation, setting, or experiment. The soundness of a model can in part be gaged by the breadth of contexts in which the model applies. Moreover, when a sound model is used in novel contexts it may predict new and unanticipated phenomena that motivate new observations and new kinds of tests. If a model can withstand repeated testing in diverse applications and experiments, its status gradually rises to that of theory and, ultimately, to that of physical law. A great advantage of reductionist science (and of models that arise from reductionism) lies in its cumulative nature and ability

to identify and exploit physical laws. Models that are synthesized to meet only a particular purpose lack this advantage of cumulation.

2.4. Hierarchy of Parameter Usage

If two competing models perform equally well in predicting phenomena of interest, what additional criteria can be used to judge the merit of the models? One useful guideline is provided by the types of parameters the models employ. It is the *type,* not the *number*, of parameters that generally distinguishes models with a firm physical foundation from "floating" models that are created to reproduce observations mathematically without providing linkage to underlying physical laws [*Savage*, 1997]. (Extreme examples of such floating models include the remarkably lifelike digital animations developed for Hollywood movie productions; the visual appeal of such animations is immense, but scientific content is commonly minimal.)

The weakest type of scientific model contains parameters that have meaning only in the context of that specific model. The fitted coefficients of statistical regression models fall in this category, as do similarly adjustable coefficients that are defined through curve-fitting procedures in more elaborate models. Use of such specialized coefficients limits the potential for broadening the scope of a model to make additional kinds of predictions subject to additional kinds of tests.

A relatively low standard is also met by models that employ parameters with apparent meaning outside the context of the model, but which yield contradictory or even meaningless results if the parameters are measured outside the model context. For example, consider models of subglacial till deformation that use a viscosity coefficient to relate shear stress to shear strain rate. To fit model predictions to field data, till can be assigned a viscosity many billions of times that of water [e.g., *Alley*, 1989]. However, laboratory tests of tills in a ring-shear device (which produces suitably large shear strains) indicate negligible dependence of shear stress on shear strain rates in the appropriate range. The laboratory results cast strong doubt on the relevance of till viscosity as a meaningful parameter, although the meaning of viscosity is unambiguous in other contexts [cf. *Iverson and Iverson*, 2001]. Use and calibration of an inappropriate parameter brings the predictions of viscous till models into agreement with field data, but lends little real insight and little basis for confident prediction if glacier-till systems are perturbed into states that differ significantly from the states in which till viscosities were calibrated.

A better standard of parameter usage is evident in models containing only relevant parameters that are clearly defined and readily measured outside the model context, but which require calibration owing to geological heterogeneity that is unresolved by independent measurements. Some of the best-established models used in geosciences fit this description, including models of seismic wave propagation and fluid flow through porous media. The basic equations in these models are either fundamental (conservation of mass and momentum) or have undergone repeated experimental testing (Hookean elasticity, Darcian permeability). Laboratory and field experiments with diverse Earth materials have uncovered cases in which Hooke's law and Darcy's law are not valid (thus, these formulas are not actually "laws" in the strictest sense), but have reinforced the notion of their wide applicability. Models of seismic-wave propagation and Darcian subsurface flow can therefore provide genuine, testable predictions. Nonetheless, the predictive and explanatory power of these models is compromised if values of elasticity and permeability parameters are calibrated through model tuning rather than measured in independent laboratory or field experiments.

In geomorphology the need for calibration is greatest in models that are coarsely parameterized to aid computation of large-scale phenomena. Coarse parameterization and calibration are justifiable if they explicitly compensate for the effects of conservation laws operating at scales too small to be resolved at the model scale. For example, an approximate, one-dimensional, reach-averaged model of flood-wave propagation in the Colorado River uses a flow resistance parameter that is adjusted with changes in river stage to accurately reproduce observations [*Wiele and Smith*, 1996]. The resistance adjustments compensate for the model's neglect of stage-dependent form drag, an effect of momentum conservation that cannot be represented explicitly in a reach-averaged, 1-D model. Interpretation and justification of the variable-resistance 1-D model hinges not only on its success in fitting data, but also on its clear mathematical derivation from more complete, multidimensional flood-wave models. Such models demonstrate how form drag arises as a consequence of momentum conservation, and they account for flow resistance at smaller, unresolved scales by using parameters with clear links to experimental studies of fluid motion [e.g., *Wiele et al.*, 1996]. In this case, as in all cases, high standards of parameter usage and calibration are attained by establishing links to physical laws and independent experimental data.

The highest standard of parameter usage exists in models that employ only parameters clearly defined and definitively measured outside the context of the model. (A simple example is the model $dv/dt = g$ of the descent velocity v of a boulder in free fall from a cliff. In this model the sole parameter g can be measured using a great variety of techniques, which invariably yield a result similar to 9.8 m/s^2 at Earth's sur-

face.) Such models require no calibration and make unequivocal predictions, as long as knowledge of initial and boundary conditions and domain heterogeneity is available. In contrast to deep-Earth and deep-time geosciences, geomorphology can potentially achieve this standard of parameter usage. Many of the fundamental parameters relevant to geomorphic processes are nearly constant (e.g., the magnitude of gravitational acceleration at Earth's surface, and the density, viscosity, and surface tension of water at standard temperature and pressure), and other parameters vary over ranges that are minuscule by Earth-materials standards (e.g., the densities of sediment grains and friction coefficients of grain contacts). Generally, only heterogeneity provides a serious impediment to full characterization of the fundamental properties of geomorphic materials using independent measurements.

The real challenge in adopting the highest standards of parameter usage in geomorphic models may involve a change of paradigm for model tests: controlled experiments that isolate phenomena by specifying initial and boundary conditions and restricting heterogeneity yield the best, least ambiguous tests of mechanistic models. Effects of realistic heterogeneities and uncertain initial and boundary conditions should be included in models *after* model predictions have been tested in controlled experiments.

3. APPRAISAL OF MODELS AS CONCEPTUAL FRAMEWORKS

Mathematical models can serve a scientific role beyond prediction. When expressed in mathematical form, mechanistic hypotheses, theories and laws have a precise content that leaves little room for equivocation about concepts and assumptions. The role of models as unambiguous conceptual frameworks highlights the importance of logical consistency in model formulation. The criteria below provide a basis for appraising the value of mechanistic geomorphic models as conceptual frameworks. Models that are not mechanistic (e.g., statistical models) have limited value in this context.

3.1. Minimization of Axioms and Postulates

It has often been stated that the ultimate goal of science is finding the most concise yet universal explanation of diverse natural phenomena. Mathematics is well-suited to this purpose, and mathematical models that minimize the number of axioms and postulates invoked to explain observations are inherently superior to those that require additional axioms and postulates [e.g., *Einstein*, 1940]. Such minimization was the great achievement of Newton's reduction of Kepler's laws of planetary motion to a single law of gravitation and of Einstein's general theory of relativity, which demonstrated

that gravitation can be viewed as an *effect* of space-time curvature in the presence of mass rather than as a *cause* of mass attraction. In essence, Kepler's three empirical laws of planetary motion were gradually subsumed by a single geometric postulate.

What do the successes of Newton and Einstein imply for geomorphic modeling? Perhaps the most important lesson is that invention of new laws should be undertaken with caution, because new laws must be consistent with all existing laws and observations. In mechanistic geomorphic models, the length and time scales of interest generally dictate that classical gravitation and conservation of mass, momentum, and energy are the only fundamental laws with universal importance. Additional (constitutive) "laws" are just surrogates for conservation laws operating at scales too small to be resolved at the scale of a particular model. For example, Darcy's law summarizes the effect of momentum conservation as fluid percolates through a porous solid matrix, where microscopic momentum transfer produces a macroscopic Darcian drag force. A wealth of experimental data indicate that the Darcian summary of this process is commonly adequate. However, just as a mathematical model that uses microscopic momentum conservation to describe flow through porous media is inherently superior to a model that invokes Darcy's law, a model that employs Darcy's law is inherently superior to one that invents a new law with weaker links to conservation laws and experimental tests. Minimization of postulates always argues against creation of new laws.

A second lesson from Newton and Einstein is that basic axioms and postulates (i.e., assumptions) of any model should be clearly stated. Too often, mathematical models in geomorphology are presented without clear derivation from established laws and principles, and model logic must be accepted on faith or inferred by prospective model users. Equations presented without derivation can almost seem like subterfuge, and critics are rightly skeptical of models promulgated in this manner. If model results have real value, then it is equally valuable to demonstrate the full logic and derivation of the model, and to acknowledge its limitations.

3.2. Adherence to Classical Conservation Laws

Along with experimentation, physical conservation laws are the most powerful tools in science. Several of them apply to all known phenomena at all length and time scales [*Feynman*, 1994]. In geomorphology the most apt conservation laws are those of classical mechanics and thermodynamics: conservation of energy, momentum, and mass. Forces express momentum transfer at length scales too small to be observable or resolvable in the problem of interest. Thus,

Newton's second law, which characterizes net force as $F = d(mv)/dt$, is simply one way of expressing conservation of observable momentum, mv, as it evolves with time, t.

Geomorphic models should take full advantage of mass and momentum conservation, because properly formulated conservation laws involve no assumptions other than identification of the length and time scales of primary interest and stipulation of constitutive formulas that summarize behavior at smaller scales. However, although mass conservation is a common focus in geomorphic models, momentum conservation is commonly disregarded or even violated. Generally this violation is not intentional but is introduced by making inappropriate assumptions. Such errors can result from attempts to use one- or two-dimensional models to draw inferences about three-dimensional phenomena, or from neglect of an important force through misidentification of the boundaries of a system in which momentum is conserved.

Violation of momentum conservation can occur, for example, in use of the infinite-slope stability model to calculate three-dimensional attributes of landslides. The infinite-slope model assumes that all relevant forces are resolvable on planes that parallel the ground surface. This simple model has value as a pedagogical tool, but its one-dimensionality restricts its practical application to cases where forces on all slope-normal surfaces are negligible. Nonetheless, some investigators have applied the infinite-slope model to multidimensional landslides by appending *ad hoc* terms intended to account for resisting forces (e.g., root strength) along landslide margins that do not parallel the ground surface [e.g., *Montgomery et al.*, 2000]. Addition of such terms without derivation from multidimensional momentum (i.e., force-balance) equations breaks the connection of the model to physical conservation laws.

3.3. Approximations Versus Assumptions

Model formulations should distinguish rational approximations from assumptions made mostly for convenience. Some geomorphic models are deemed "approximate" when "assumed" is a more accurate characterization. Approximation is a mathematical procedure. Just as the number 7.2 can be obtained as a rational mathematical approximation of the number 7.23, approximate equations can be derived from more complete, more accurate equations through a rational mathematical process.

Approximation procedures are particularly relevant with respect to application of exact conservation laws. For example, an exact differential equation describing conservation of a continuously differentiable scalar quantity c observed in a deforming, continuous substance with velocity v is $\partial c/\partial t + v \cdot \nabla c + c \nabla \cdot v = S$, where S describes the influence of sources and sinks (phenomena that affect conservation of c but at a scale that is not resolvable). However, if measurements demonstrate that the deforming substance is nearly incompressible and undergoes negligible volume change, then mathematical analysis demonstrates that conservation of mass implies $\nabla \cdot v \approx 0$, and the conservation equation for c thereby reduces to a simpler form, $\partial c/\partial t + v \cdot \nabla c \approx S$. Applicability of this approximate equation has not been assumed *a priori*. Rather, it has been demonstrated through derivation from a more exact statement of conservation.

A wide variety of approximation procedures are used in applied mathematics and theoretical physics, and are beyond the scope of this paper. The key implication these procedures hold for geomorphic modeling is that rational approximations are commonly possible, and they differ fundamentally from assumptions. Only rational approximations reveal how models are connected to physical conservation laws.

3.4. Conceptual Versus Mathematical Simplicity

The criteria described above—minimization of postulates, adherence to classical conservation laws, and use of approximations rather than assumptions—can be summarized by the principle that models should be as simple as possible. This principle is valid, however, only if a clear distinction is drawn between mathematical simplicity and conceptual simplicity. Too often in geomorphology, the adjective "simple" is applied to models that use simple mathematics, not simple concepts. It is important to recognize that models constructed of elegantly simple, logical, and testable concepts can yield complicated mathematics, and that models involving only elementary algebraic equations can express ill-founded concepts that are untestable or incompatible with physical laws. As stated by Fourier, "Nature's indifferent to the difficulties it causes a mathematician," [*Freeze*, 1978] and geomorphologists ought to embrace mathematical challenge as an integral part of science.

4. AN EXAMPLE: LANDSLIDE RUNOUT

Many of the issues described above arise in modeling and predicting runout of rapid landslides. This problem is similar to many in process geomorphology because it has applications in geologic hazards assessment and land-use management, relevance to long-term sediment delivery and landscape change, and intrinsic scientific interest as an observable yet sometimes perplexing phenomenon. The landslide-runout models described below illustrate a spectrum of conceptual frameworks, predictive power, and modeling pitfalls.

4.1. The Coulomb Slide-block Model

The best-known model for predicting landslide runout is the Coulomb slide-block or energy line model first presented by *Heim* [1932] and still used in some practical applications [e.g., *Sheridan et al.*, 1999]. The Coulomb slide-block model is physically based but neglects key aspects of mass and momentum conservation. The model results from formulating and solving a one-dimensional momentum equation (Newton's second law) for a rigid body of density ρ and uniform thickness h sliding down a slope inclined at an angle θ,

$$\rho h \frac{dv}{dt} = \rho g h \sin \theta - \rho g h \, \cos \theta \tan \phi \qquad (1)$$

where ϕ is the Coulomb friction angle that characterizes the ratio of shear to normal forces at the sliding surface. Coulomb friction is a surprisingly simple "emergent" phenomenon that summarizes the macroscopic effect of momentum conservation at innumerable microscopic grain contacts [cf. *Duran*, 2000]. Moreover, Coulomb friction is a readily measurable property of rocks and soils, and is therefore a relevant constitutive parameter to include in landslide models.

Interpretation of (1) is facilitated by canceling the factor ρh from all its terms, which leaves ϕ and g as the only parameters of interest. This simplification makes it clear that (1) is equivalent to an equation describing gravity-driven, frictional sliding of a point mass with zero volume. Therefore, runout predictions derived from (1) tacitly assume that effects of landslide volume are inconsequential.

If centripetal acceleration due to slope curvature is neglected (another tacit assumption in Heim's analysis and in (1)), integration of (1) along a landslide path with decreasing slope shows that the extent of runout is predicted by

$$H/L = \tan \phi \qquad (2)$$

where H is vertical height of landslide descent, and L is the horizontal runout distance. Laboratory experiments with soil and rock samples consistently yield Coulomb friction angles around $\phi = 30\text{-}40°$, and use of these values in (2) yields relatively accurate runout predictions for many small landslides. However, landslides larger than about 10^6 m^3 typically run out further than predicted by (2) with $\phi = 30\text{-}40°$, and this "excess" runout increases systematically as landslide size increases [e.g., *Schiedegger*, 1973; *Legros,* 2002].

The excess runout of large landslides has attracted much attention: why should a straightforward, physically based predictive model such as (2) fail systematically?

Remarkably little of this attention has focused on removing the assumption of zero landslide volume and applying conservation laws to evaluate the effects of mass redistribution and internal momentum transfer. Instead, the point-mass model (1) has generally been retained, and various extraneous phenomena have been proposed to explain long runouts. Some of these proposals are quite plausible (e.g., high fluid pressures that reduce basal fiction) and others less so, but none can be evaluated conclusively without first clarifying the role of mass and momentum conservation in four (space + time) dimensions.

4.2. The Adjustable Resistance Model

Many proposals for explaining long landslide runouts replace the well-supported Coulomb friction term $\rho g h \cos \theta \, \tan \phi$ in (1) with a less exacting expression that allows adjustment of resisting forces [e.g., *Voight et al.*, 1983]. For example, a term that includes constant resistance c plus velocity-dependent resistance can be substituted in (1) to generate the equation of motion

$$\rho h \frac{dv}{dt} = \rho g h \sin \theta - c - \frac{\mu}{h} v \qquad (3)$$

where c is a finite, stress-independent ("cohesive") shear strength and μ is a viscosity coefficient. A wide variety of results can be generated by using (3) and adjusting the values of c and μ to match model output with data. This procedure might yield insight if c and μ had clear relevance and values that were measurable outside the context of the model. However, numerous laboratory experiments with pervasively fractured rocks and soils yield almost no evidence of significant viscous behavior over the range of pressures, temperatures, and deformation rates applicable in landslides, and the experiments yield cohesive strengths too small to be relevant in large-scale landslides ($> 10^6$ m^3).

The effect of scale becomes more apparent if (3) is normalized by dividing all terms by $\rho g h$, which yields

$$\frac{1}{g} \frac{dv}{dt} = \sin \theta - \frac{c}{\rho g h} - \frac{\mu}{\rho g h^2} v \qquad (4)$$

This equation demonstrates that effects of cohesion and viscosity (if present) diminish as landslide size increases, because the landslide thickness, h, appears in the denominators of both the cohesion and viscous terms. One might argue that this scaling explains long runouts because it indicates that fixed resistance due to constant c produces the smallest effects in the largest landslides [cf. *Dade and Huppert*, 1998]. However,

such arguments are untenable because no experimental evidence indicates that c or μ have relevance like that of $\tan \phi$.

4.3. The Mass-Change Model

A different type of untenable argument results from applying a Coulomb slide-block model similar to (1) to cases with changing landslide mass—without properly accounting for the effects of mass change on momentum conservation. For example, if deposition causes progressive loss of landslide mass, then a one-dimensional analysis implies that landslide thinning occurs ($dh/dt < 0$). To account for this effect, one might generalize the momentum equation (1) as

$$\rho \frac{d(hv)}{dt} = \rho gh \sin \theta - \rho gh \cos \theta \tan \phi \qquad (5)$$

Here, the left-hand side can be expanded as $\rho [v(dh/dt) + h(dv/dt)]$, which appears to indicate that loss of mass in transit ($dh/dt < 0$) increases landslide acceleration (dv/dt) and ultimately enhances runout [cf. *Van Gassen and Cruden*, 1989]. This logic is flawed, however, because it assumes erroneous boundaries for the system in which momentum conservation applies. If mass leaves the landslide and becomes part of the stationary bed, momentum conservation applies not to the moving landslide alone but to the landslide and bed together. Viewed another way, if the landslide behaves as a body decoupled from the underlying Earth, the Earth must exert an additional external force to cause deceleration and depletion of the landslide mass, and this additional force is neglected in (5) [*Erlichson*, 1991]. Therefore, predictions derived from (5) do not yield meaningful conclusions.

This example highlights an important mechanical principle in geomorphology: any geomorphic process that involves erosion or sedimentation also involves exchange of momentum between a surficial body (such as a landslide) and the underlying Earth. Such mass and momentum exchange does not, however, imply that conservation laws are rendered useless owing to the presence of an open system. Rather, it implies that the system to be analyzed must be defined with care and precision.

4.4. The Continuum Conservation-Law Model

A reliable path for discovering the controls of landslide runout generalizes the momentum balance contained in the Coulomb sliding model (1) to explore the implications of continuum conservation laws in four (space + time) dimensions [e.g., *Gray et al.*, 1999]. In this approach the mathematics involve a system of simultaneous, nonlinear partial differential equations that are considerably more complicated than (1), but rational approximations facilitate analysis. No parameters other than those in (1) are introduced, and the essential concepts remain very simple: momentum is conserved, mass is conserved, and internal and boundary forces are governed by Coulomb friction.

With this conceptual framework, explicit and testable predictions become possible, issues of scale dependence become clear, and additional effects (such as those of pore-fluid pressure and runout-path geometry) can be investigated systematically, without invoking poorly supported simplifying assumptions [*Iverson and Denlinger*, 2001; *Denlinger and Iverson*, 2001]. In contrast to the single dependent variable $v(t)$ predicted by (1), the continuum conservation-law model predicts three interdependent variables: two velocity components, v_x and v_y, and landslide thickness, h, all of which vary as functions of two map coordinates, x and y, and time t. Simultaneous prediction of these multiple interdependent variables expands the possibilities for stringent model tests [cf. *Furbish*, this volume].

Four-dimensional continuum conservation-law models provide a means for drawing concrete conclusions about the influence of various phenomena affecting landslides (such as landslide volume, Coulomb friction, path geometry, and pore-fluid pressure), but they do not provide a panacea. Such models pose significant computational difficulties, which highlight the importance of accurate solution methods (whether numerical or analytical) as well as rigorous model formulation. Furthermore, a continuum model that invokes Coulomb friction provides information about phenomena only at the bulk scale, and grain-scale interactions are unresolvable. More detailed models, which include mass and momentum conservation at smaller scales, can provide deeper understanding, but generally do so at the cost of less explicit prediction of macroscopic runout.

4.5. The Many-Body Conservation-Law Model

Detailed phenomena in moving landslides can be examined by using micromechanical models that resolve momentum and mass conservation at the scale of individual grains [*Campbell et al.*, 1995]. Such models have undisputed value, but have very large degrees of freedom (e.g., angular and linear momenta and positions of each grain). The great computational demands of such models have, to date, limited their application to two-dimensional landslides consisting of no more than about 10^6 disk-shaped grains (similar to the number of grains in 1 liter of coarse sand). As a consequence, the full macroscopic effects of mass and momentum conservation in landslides have not been predicted. The

feasibility of more realistic micromechanical models will likely grow as computational resources grow. However, in the foreseeable future, such models will serve the greatest purpose in providing guidance for improved continuum modeling.

4.6. The Volume Statistics Model

Gaps in existing mechanistic models leave a niche for empirical models that are less ambitious but more practical. Physical knowledge and dimensional analysis of the landslide runout problem provide guidance for selecting appropriate independent and dependent variables and computing relevant statistics [*Iverson et al.*, 1998].

The systematic decline of measured H/L values with increasing landslide size indicates that any predictive model should account for landslide volume explicitly [*Davies*, 1982] – unlike the models characterized by equations 1-5. Several investigators [e.g., *Hungr*, 1990] have noted that the planimetric area A inundated by a rapid landslide will be proportional to landslide volume V raised to the 2/3 power if geometric similarity applies (because A has dimensions of length squared and V has dimensions of length cubed). Indeed, an empirical predictive equation

$$A = 10 \, V^{2/3} \qquad (6)$$

fits the trend of data that relate inundated areas to landslide volumes ranging over many orders of magnitude [*Kilburn et al.*, 1998; *Dade and Huppert*, 1998; *Legros*, 2002]. When used in conjunction with a similar equation that relates landslide volumes to the vertical cross-sectional areas of runout paths, (6) provides a basis for statistical forecasting of inundation limits [*Iverson et al.*, 1998].

If an empirical equation such as (6) is tested and calibrated statistically, it carries explicit measures of uncertainty and predictive confidence. This feature argues strongly for use of empirical statistical models in many practical applications. Moreover, if elementary physical reasoning such as dimensional analysis is used to posit and constrain empirical models, they can serve as stepping stones toward mechanistic understanding [*Bridgman*, 1922].

5. CONCLUSION: SELF-INTERROGATION FOR MODELERS

The preceding sections of this paper describe criteria that can guide formulation and appraisal of mathematical models of observable geomorphic processes. The list below recapitulates these criteria as a series of ten questions for consideration by model users as well as model builders.

1. Why construct or use a mathematical model? What is the objective?
2. If the chief objective is practical prediction of a specific phenomenon, will an empirical statistical model suffice and involve fewer assumptions and better estimates of uncertainty than those of a similarly directed mechanistic model?
   (If a principal objective is development of a precise and consistent conceptual framework for describing and explaining phenomena, proceed with formulation of a mechanistic model.)
3. Does the mechanistic model satisfy physical conservation laws (for energy, momentum, and mass)?
4. Have axioms and postulates (such as new "laws") been minimized in model formulation?
5. Have the basic axioms, logical development, and mathematical derivation of the model been fully documented and justified?
6. Have mathematical simplifications been attained through rational approximations rather than assumptions?
7. Have values of all model parameters been obtained through independent measurements, insofar as possible?
8. Do predictions of the uncalibrated model match the results of relevant, controlled experiments? If not, reformulate the model.
9. Do predictions of the uncalibrated model match field measurements and observations? If not, is the misfit the result of undocumented heterogeneity or fine-scale phenomena that are not resolvable in the model? Is parameter calibration warranted?
10. What *else* does the model predict? What are its logical and mathematical limits? Is the model sufficiently accurate and universal that it can be used to predict and interpret unobservable phenomena such as landscape evolution?

REFERENCES

Alley, R.B., Water-pressure coupling of sliding and bed deformation: II. velocity-depth profiles, *Journal of Glaciology*, 35, 119-129, 1989.

Baker, V.R., Hypotheses and geomorphological reasoning, in *The Scientific Nature of Geomorphology*, B.L. Rhoads and C.E. Thorn, eds., Wiley, New York, 57-85, 1996.

Bridgman, P.W., *Dimensional Analysis*, Yale University Press, New Haven, Conn., 1922.

Campbell, C.S., P.W. Cleary, and M. Hopkins, Large-scale landslide simulations: global deformation, velocities and basal friction, *Journal of Geophysical Research, B*, 100, 8267-8283, 1995.

Chamberlin, T.C., The method of multiple working hypotheses, *Journal of Geology*, 5, 837-848, 1897.

Cleland, C.E., Historical science, experimental science, and the scientific method, *Geology*, 29, 987-990, 2001.

Dade, W.B., and H.E. Huppert, Long-runout rockfalls, *Geology*, 26, 803-806, 1998.

Davies, T.R.H., Spreading of rock avalanche debris by mechanical fluidization, *Rock Mechanics*, 15, 9-24, 1982.

Denlinger, R.P., and R.M. Iverson, Flow of variably fluidized granular masses across three-dimensional terrain: 2. numerical predictions and experimental tests, *Journal of Geophysical Research, B*, 106, 553-566, 2001.

Duran, J., *Sands, Powders and Grains - An Introduction to the Physics of Granular Materials*, Springer, New York, 214 p., 2000.

Einstein, A., The fundamentals of theoretical physics, *Science*, 91, 487-492, 1940.

Erlichson, H., A mass-change model for the estimation of debris-flow runout: a second discussion, *Journal of Geology*, 99, 633-634, 1990.

Feynman, R.P., *The Character of Physical Law*, The Modern Library, New York, 1994, 167 p.

Freeze, R.A., Mathematical models of hillslope hydrology, in *Hillslope Hydrology*, M.J. Kirkby, ed., Wiley, Chichester, 177-225, 1978.

Furbish, D.J., Using the dynamically coupled behavior of land-surface geometry and soil thickness in developing and testing hillslope evolution models, *This Volume*, 2003.

Gray, J.M.N.T., M. Wieland, and K. Hutter, Gravity driven free surface flow of granular avalanches over complex basal topography, *Proceedings of the Royal Society of London, Ser. A.*, 455, 1841-1874, 1999.

Haff, P.K., Limitations on predictive modeling in geomorphology, in *The Scientific Nature of Geomorphology*, B.L. Rhoads and C.E. Thorn, eds., Wiley, New York, 1996, 337-358, 1996.

Heim, A., *Bergsturz und Menschenleben*, 218 p., Fretz and Wasmuth, Zürich, Switzerland, 1932.

Hsü, K.J., Catastrophic debris streams (sturzstroms) generated by rockfalls, *Geological Society of America Bulletin*, 86, 129-140, 1975.

Hungr, O., Mobility of rock avalanches, *Report of the National Research Center for Disaster Prevention* (Japan), no. 46, 11-19, 1990.

Iverson, N.R., and R.M. Iverson, Distributed shear of subglacial till due to Coulomb slip, *Journal of Glaciology*, 47, 481-488, 2001.

Iverson, R.M., Landslide triggering by rain infiltration, *Water Resources Research*, 36, 1897-1910, 2000.

Iverson, R.M., and R.P. Denlinger, Flow of variably fluidized granular masses across three-dimensional terrain: 1. Coulomb mixture theory, *Journal of Geophysical Research, B*, 106, 537-552, 2001.

Iverson, R.M., S.P. Schilling, and J.W. Vallance, Objective delineation of lahar-inundation hazard zones, *Geological Society of America Bulletin*, 110, 972-984, 1998.

Jeffreys, W.H., and J.O. Berger, Ockham's razor and Bayesian analysis, *American Scientist*, 80, 64-72, 1992.

Kilburn, C.R.J., and S-A. Sorensen, Runout lengths of struzstroms: the control of initial conditions and of fragment dynamics, *Journal of Geophysical Research, B*, 103, 17877-17884, 1998.

Legros, F., The mobility of long-runout landslides, *Engineering Geology*, 63, 301-331, 2002.

Lighthill, J., Chaos: a historical perspective, in *Nonlinear Dynamics and Predictability of Geophysical Phenomena,* W.I. Newman, A. Gabrilov, and D.L. Turcotte, eds., Geophysical Monograph 88, Am. Geophys. Union, Washington DC., 1-5,1994.

Montgomery, D.A., K.M. Schmidt, H.H. Greenberg, and W.E. Dietrich, Forest clearing and landsliding, *Geology*, 28, 311-314, 2000.

Oreskes, N., K. Shrader-Frechette, and K. Belitz, Verification, validation, and confirmation of numerical models in the earth sciences, *Science*, 263, 641-646, 1994.

Oreskes, N., and K. Belitz, Philosophical issues in model assessment, in *Model Validation: Perspectives in Hydrological Science*, M.G. Anderson and P.D. Bates, eds., Wiley, New York, 25-41, 2001.

Paola, C., J. Mullin, C. Ellis, D.C. Mohrig, J.B. Swenson, G. Parker, T. Hickson, P.L. Heller, L. Pratson, J. Syvitski, B. Sheets, and N. Strong, Experimental stratigraphy, *GSA Today*, 11, 4-9, 2001.

Pippard, A.B., *Response and Stability, An Introduction to the Physical Theory*, Cambridge University Press, Cambridge, 228 p., 1985.

Platt, J.R., Strong Inference, *Science*, 146, 347-353, 1964.

Proust, M., *Remembrance of Things Past*, v. 3, translated by C.K.S. Moncrieff, T. Kilmartin, and A. Mayor, Random House, New York, 1981.

Raymo, C., *Skeptics and True Believers*, Walker and Company, New York, 288 p.,1998.

Sarewitz, D. and R.A. Pielke Jr., Prediction in science and policy, in *Prediction B Science, Decision Making, and the Future of Nature*, D. Sarewtiz, R.A. Pielke, Jr., and R. Byerly, Jr., eds., Island Press, Washington, D.C., 11-22, 2000.

Savage, S.B., Modeling and granular material boundary value problems, in *Physics of Dry Granular Media*, H.J. Herrmann, J.P. Hovi, and S. Ludig, eds., Kluwer, 25-96, 1997.

Scheidegger, A.E., On the prediction of the reach and velocity of catastrophic landslides, *Rock Mechanics*, 5, 231-236, 1973.

Shapiro, R., *Origins - A Skeptics Guide to Creation of Life on Earth*, Bantam, Toronto, 332 p., 1986.

Sheridan, M.F., B.E. Hubbard, W, Strauch, B. van Wyk de Vries, A hazard model for the 30 October 1998 debris avalanche and lahar at Casita Volcano, Nicaragua, *EOS*, 80, F1154, 1999.

Shreve, R.L., Models for prediction in fluvial geomorphology, *Mathematical Geology,* 11, 165-174, 1979.

Van Gassen, W., and D.M. Cruden, Momentum transfer and friction in the debris of rock avalanches, *Canadian Geotechnical Journal*, 26, 623-628, 1989.

Voight, B., R.J. Janda, H. Glicken, and P.M. Douglas, Nature and mechanics of the Mount St. Helens rockslide-avalanche of 18 May 1980. *Geotechnique*, 33, 243-273, 1983.

Webb, R.H., J.C. Schmidt, G.R. Marzolf, and R.A. Valdez, eds., *The controlled flood in Grand Canyon*, Geophys. Monograph 110, Am. Geophys. Union, Washington DC., 367 p., 1999.

Werner, B.T., Complexity in natural landform patterns, *Science*, 284, 102-104, 1999.

Wiele, S.M., and J.D. Smith, A reach-averaged model of diurnal discharge wave propagation down the Colorado River through Grand Canyon, *Water Resources Research*, 32, 1375-1386, 1996.

Wiele, S.M., J.B. Graf, and J.D. Smith, Sand deposition in the Colorado River in the Grand Canyon from flooding of the Little Colorado River, *Water Resources Research*, 32, 3579-3596, 1996.

R.M. Iverson, U.S. Geological Survey, 1300 SE Cardinal Ct. 100, Vancouver, WA 98683-9589 phone: 360-993-8920 fax: 360-993-8980 email: riverson@usgs.gov