TFW-PR10-99-001

Comparison of GIS-based Models of Shallow Landsliding for Application to Watershed Management

by

Susan C. Shaw and Laura M. Vaugeois

TIMBER FISH & WILDLIFE

July 9, 1999

PROJECT 10

FINAL REPORT

Comparison of GE-based Models of Shallow Landsliding for Application to Watershed Management

by

Susan C. Shaw and Laura M. Vaugeois

Washington Department of Natural Resources
Forest Practices Division
P.O. Box 47012
Olympia, WA. 98504-7012

Report prepared for the:
Cooperative Monitoring. Evaluation, and Research (CMER) Committee,
Timber/Fish/Wildlife Agreement
Washington Forest Protection Association
Washington Department of Natural Resources

July 9, 1999

EXECUTIVE SUMMARY

Project Purpose

The need for testing and improving GIS-based slope stability models, for use in forest management and forest-practices regulations, has been identified by the Timber/Fish/Wildlife (T/F/W) Cooperative Monitoring, Evaluation, and Research (CMER) Committee and the newly released Forests and Fish Report to the Washington Forest Practices Board (WFPB) and the Governor's Salmon Recovery Office (USDI Fish and Wildlife Service et al., 1999). The original T/F/W agreement (1987, p. 31) called for "... moving toward a hazard zonation mapping system to better identify areas of instability", and efforts began soon thereafter to design mapping systems, both manual and GIS-based, for screening shallow landslides. Likewise, the Forests and Fish Report has called for "a project to identify the best available topo/geographic model to flag landforms that have significant potential to initiate shallow rapid landslides" (p. 37), in anticipation of the completion of the study described herein.

Over the past eight years, CMER has funded or partially funded research to develop GIS-based models. These models, however, have not been tested rigorously or adapted for statewide application to management and regulation of commercial forest lands in Washington. Consequently, the CMER Committee recommended, and the T/F/W Policy Committee approved, Project 10 ("Erosion Effects from Forest Practices") for the 1997-99 biennium, the primary intents of which were to:

(1) evaluate the performance of GIS-based slope-stability models that are readily available and have been developed with support from T/F/W and its cooperators;

(2) select one or more models that meet stated criteria for scientific accuracy, technical accessibility, and applicability to forest management and regulation in Washington; and,

(3) further refine the selected model(s) and make recommendations to the T/F/W community

regarding its/their use as a screening tool, particularly where regulatory watershed analyses or landslide inventories have not been completed.

Model implementation would be promoted by making software and documentation available to all T/F/W cooperators, or by creating a publicly accessed, regional landslide screen to replace the current WDNR-GIS slope-stability layer.

In May 1998, we contracted with the CMER Committee and the Washington Forest Protection Association (WFPA) to carry out Project 10. This technical report and accompanying recommendations describe the methods, results, and conclusions of our year-long analysis. We thank the T/F/W group, WFPA, and Washington Department of Natural Resources (WDNR) for their generous support of this project.

During the course of this study, our focus expanded from evaluating models for use in regulatory watershed analyses and routine forest management, to include an assessment of their potential as statewide landslide-screening tools. This shift was driven primarily by the Forestry Module negotiations and the resulting commitments of the Forests and Fish Report to promote the development of a statewide screen. Hence, we provide recommendations for model use at the local and regional scales. This project has focused on western Washington, due primarily to time constraints. Consequently, we are developing a similar test for watersheds in each of the distinct geomorphic provinces in eastern Washington, as groundwork for creating a statewide screen of shallow landsliding. This test should help determine whether any of these GIS-based models can accommodate the geology and climatic regimes east of the Cascades Range.

Summary of Study Methods and Report Conclusions

We evaluated three shallow-landslide predictive methods that have been used previously in Washington forest management and regulation: the current WDNR-GIS slope-stability screen (referred to in this report as the SOILS screen; tested at the request of CMER Committee members); the SMORPH or DNR-SL model (Shaw and Johnson, 1995); and the SHALSTAB model (Montgomery and Dietrich. 1994). We originally proposed to test the WHPM model (Wu and Abdel-Latif, 1995, 1997); however, this model has not been fully developed and programmed by the authors and, hence, was unavailable during the course of this study. Other published models (e.g., Wu and Sidle, 1995; Pack et al., 1998) also were not tested fully due to availability and software-development issues.

We tested the three selected models in eight watersheds (i.e., nine Watershed Administrative Units (WAUs) and portions of four others), representing the major geomorphic provinces of western Washington (see Figure 1 in the Technical Report) and incorporating 2524 known, existing shallow landslides. The test was conducted by executing the model programs, creating GIS covers from model results, comparing them statistically with landslide inventories and hazard-zonation maps produced for this project or as products of regulatory watershed analyses, and verifying model predictive accuracy in the field.

For maximum test accuracy, we found that we had to verify and update most of the landslide inventories in the field, and make corrections or additions to the digital databases (i.e., we encountered problems with the watershed-analysis GIS products). We also modified the SHALSTAB program, with assistance from one of the authors and staff, such that it functioned correctly on the WDNR UNIX computer system. In addition, we needed to create a method for converting SHALSTAB model output, given as critical rainfall amounts necessary to initiate landslides (in mm/day), to management criteria (i.e., low, moderate, and high "hazard" potentials) in order to compare the model results with those of the SOILS screen and SMORPH model. The latter two models yield results in terms of management criteria, as defined by

WFPB regulations. While conducting this study, we developed some additional documentation and computer help tools that will improve the "user-friendliness" of the SMORPH and SHALSTAB models. This documentation, as well as the algorithm for converting SHALSTAB output, are available by obtaining copies of the computer programs from the WDNR.

Based on statistical comparisons of model results and existing landslide data, we have concluded the following regarding the management applicability of these models:

(1) The SMORPH model generated spatial predictions of shallow landslides that correlated most closely with the pattern of known, existing landslides (i.e., landslide inventory databases) and the landslide hazard-potential maps (e.g., Mass-Wasting Map Unit maps from regulatory watershed analysis). This model correctly predicted 97% of the total existing landslides, compared with 92% for the SHALSTAB model and 68% for the SOILS screen. Compared with the landslide hazard-potential maps, the SHALSTAB model over-predicted by an average 7% the area considered to be "high hazard", whereas the SMORPH model similarly over-predicted by an average 3%. The SMORPH model also performed substantially better than the other models in the least appropriate terrain for GIS-based model applications (i.e., continental glaciated basins).

(2) Using the landslide databases as a measure, the difference in predictive capability of the SMORPH and SHALSTAB models appears to be marginally significant statistically, whereas the difference between either of these models and the SOILS screen is very significant. Hence, SMORPH and SHALSTAB agree fairly well with observed landslide distributions and either conceivably could be developed to produce a regional or statewide GIS cover of shallow landslide potential, contingent on their calibration needs.

(3) The SHALSTAB model is less readily applicable in the current management decision-

making framework because it contains no mechanism for converting model output to management criteria (i.e., low, moderate, high "hazard" potential). Hence, using this model in a watershed-analysis or regulatory context (e.g., creating a statewide cover) would require that the model authors create an algorithm or verify that the one we created is acceptable. Our conversion algorithm was designed to yield the most conservative estimate of slope instability and reproduce most closely the spatial distribution of existing landslides, so we believe that it is a viable approach to solving this application problem. We estimate that it would take the model authors at least three months of concerted effort to make these and other desirable model modifications (e.g., addressing model calibration issues on a statewide level).

(4) The SOILS screen is the least preferable option for management applications because of its comparative inaccuracy, the inability of the user to calibrate model input variables to site-specific physical conditions, and the large gaps in geographic coverage due to lack of comprehensive, digital soils-survey data The SOILS screen is relatively more "user-friendly" than the other models because it is delivered to the user as a pre-compiled GIS cover. Contingent on further testing in eastern Washington, either SHALSTAB or SMORPH programs could be executed to yield a statewide cover that would alleviate the need for individual users to run the model. A new cover could be made available in the public domain by the WDNR.

(5) The SHALSTAB model contains more input variables than SMORPH and, consequently, has the potential for producing relatively more model errors associated with using input values that are unrepresentative of the study area. The soil-property and hydrology input variables are assigned constant values in the model. Few published methods exist for determining appropriate constant values for soil properties that can vary considerably

in space and time. Collecting sufficient data in the field also can be problematic. Reasonable values potentially can be back-calculated by running the model with a range of possible values and choosing ones that yield landslide predictions most comparable with existing landslide databases. This approach might be less labor-intensive than field sampling but requires reliable landslide inventories in a sufficient number of representative watersheds that the calibrated values can be extrapolated to basins without inventories. This calibration might hinder the speedy development of a statewide GIS cover and could inhibit the use of this model in watersheds with no viable analogs (e.g., geomorphically similar watersheds with completed inventories).

(6) The SMORPH model contains relatively fewer input variables (i.e., management criteria for different combinations of hillslope gradient and curvature), relying on the assumption that topographic factors primarily drive landslide initiation. Gradient threshold values corresponding to each criterion (i.e., low, moderate, high "hazard") are set using existing landslide inventories and/or hazard-zonation maps from geomorphically similar watersheds. Hence, this model also requires calibration and suffers correspondingly when no viable analogs exist. We found that the SMORPH model is relatively less sensitive to variations in the gradient thresholds than SHALSTAB is to variations in soil-property values (i.e., magnitudes of the estimated soil cohesion and internal friction angles). As a result, SMORPH likely can accommodate somewhat greater error in the choice of input values than SHALSTAB. In addition, gradient data are more readily accessible than soil-property information; the former can be derived from landslide inventories and topographic or DEM maps, whereas the latter are obtained from field measurements or from soil surveys and geoengineering literature.

(7) The SMORPH model, in its present form, cannot be adjusted to include site-specific soils

and hydrology data. Hence, the model might not function as well in regions where topographic factors are secondary to other hillslope processes. The same is true of SHALSTAB in its present form because it also assumes topographic forcing of landslides. Alternatively, the SHALSTAB model contains placeholders for algorithms that treat variability of soil properties (i.e., substituting computational routines for the constants) and, hence, eventually could prove to be more robust and versatile when such algorithms are added. With respect to western Washington, the comparatively better predictive capability of the SMORPH model suggests that including algorithms for soil and hydrologic factors might not be as critical as modeling the fine-scale variations in topography. This result also implies that, for western Washington, the values of input variables required in the SHALSTAB model might be adequately represented by default values currently set in the computer program.

(8) The SMORPH model runs approximately 80% faster than SHALSTAB on a computer workstation (e.g., WDNR UNIX system), which might be important to managers with limited computer resources and large data requirements. SMORPH also requires about five times less data-storage volume than SHALSTAB and several times less storage volume than the existing SOILS screen.

(9) The SHALSTAB model requires relatively more training to instruct users on executing model programs and interpreting results. The assistance of technical specialists also might be needed more frequently than with other models, to calibrate input variables and interpret model results, particularly if no uniform method exists for converting model output to management criteria.

(10) Both SHALSTAB and SMORPH perform significantly better using 10-m. versus 30-m, resolution DEM data. Hence, the finer-resolution data should be used wherever

possible. Computer programs for both models can be run using default values of the input variables, if time and/or budget precludes more lengthy calibration efforts, although the model results vary accordingly.

Recommendations for Model Adoption as a Landslide Screening Tool

Based on the conclusions presented in this summary and the technical report, we offer the following options for selecting a preferred model as a screening tool for shallow landslides. As summarized previously, the SOILS screen was determined to be the least preferable based on its predictive capability and, hence, is not offered here as an option. These recommendations are the same regardless of whether the model is employed at a watershed scale (e.g., for forest-practices-application reviews, timber-harvest planning, and preliminary hazard-zonation mapping) or at a regional scale (e.g., for creating a statewide or regional GIS cover).

OPTION 1: Choose the SMORPH model as the preferred screening tool.

The advantages of this option are that the SMORPH model performs slightly better than the current version of the SHALSTAB model and yields results that are consistent with observed landslide data. Its output is given in terms of management criteria (i.e., low, moderate, and high "hazard" potential) that are commensurate with the regulatory definitions and management decision-making process. SMORPH requires relatively less calibration, with readily available input data. This model could be incorporated with other model algorithms that address additional key factors known to influence landsliding (e.g., soil properties). The program runs substantially faster, requires less storage space, and can be implemented with less training and technical assistance.

This model has been used by a variety of private, state, and federal entities in the Pacific Northwest to create preliminary screens of landslide potential. It can be readily implemented, although additional testing should be conducted in eastern Washington to assure that it performs accurately in terrain less comparable to watersheds analyzed on the west side. We estimate that a statewide GIS cover could be developed within a few months in western Washington and in about nine months for eastern Washington.

OPTION 2: Choose the SHALSTAB model as the preferred screening tool.

The advantages of this option are that the SHALSTAB model performs nearly as well as the SMORPH model and yields results that are consistent with observed landslide inventories, if our algorithm is used to convert output data to management criteria. This model potentially offers rnore versatility in terrains where topographic controls are confounded by spatial and temporal variations in soil and hydrologic variables, although algorithms to address such variability have not been made available. The SHALSTAB model could be adapted for management and regulatory use if the output conversion algorithm used in this study were refined, replaced, or corroborated by the model authors. Whereas using the model in the current regulatory arena would require establishing management criteria, its use by analysts in watershed analysis would not necessarily depend on these criteria, given that the standard output (i.e., critical rainfall values) can be interpreted by scientific specialists. We expect that model modifications (e.g., refining management criteria) would take a number of months and potentially require funding of the authors to complete. The model requires a fair amount of calibrating with existing landslide databases or soil and precipitation data. This test, however, suggests that using the default values of the input variables is reasonable for western Washington terrains similar to those tested in this study. Although the model program is relatively less time- and storage- efficient, it nevertheless could be run, perhaps in a series of smaller geographic areas, to create a regional or statewide cover. This model requires relatively more training and assistance from technical specialists, especially in calibrating input variables and interpreting results. The SHALSTAB model or its variations have been used by several private, federal, and academic entities to produce GIS covers of landslide potential in its native units of measure (i.e., to our knowledge, no uniform method exists for interpreting results in a management context). It could be used to build a statewide GIS cover, pending refinement of management criteria and further testing in eastern Washington. We estimate that it might take about one year to develop a statewide cover, using the management criteria presented in this study, and potentially longer if other criteria need to be developed. Model modifications would be subject to funding and availability of the model authors, which could influence the completion of a statewide cover by December 2000 (i.e., anticipated deadline for implementation of the Forests and Fish Report).

OPTION 3: Choose the SMORPH model as an interim tool while the SHALSTAB model is being further developed and tested.

This option accommodates the needs of implementing a reliable statewide GIS cover by December 2000, while allowing for further development and testing of the SHALSTAB model. The SHALSTAB model is more sophisticated, although in its current version (i.e., with variables held constant), it is reduced to its most essential element (i.e., a topographic analysis). Hence, there currently is little functional difference between the current SMORPH and SHALSTAB models. This option basically takes care of the present needs while exploring potential advantages of a more complex model. The disadvantage of this option would be the time and money spent developing two GIS coverages. Switching from one model to the other in mid-stream, however, would not necessarily affect users because one coverage could be substituted for another, as long as the management criteria were defined similarly.

OPTION 4: Choose the SMORPH model as an interim tool while other promising models are being refined.

This option is similar to Option 3, although SHALSTAB would be replaced in favor of one of several other promising, GIS-based models. The advantages of these models are summarized in the Technical Report. One such method, currently being developed and tested by the USDA Forest Service and its cooperators in Oregon, couples a variation of the SHALSTAB model with a debris-flow-runout algorithm, to assess not only the spatial distribution of predicted shallow landslides but the "deliverability" of landslide materials to downstream areas with sensitive public resources (D. Miller, Earth Systems Institute, pers. comm.). Hence, this option considers the possibility that more advanced tools would be available in the near future.

We have been asked by members of the CMER Committee to recommend a preferred option. We have selected Option 1. The deciding factors for us were the slightly greater predictive capability of SMORPH, despite the conceptual simplicity of the model, and the immediate accessibility of the operating program to users with a basic knowledge of GIS and mass-wasting mapping techniques (i.e., it does not require any modifications to be implemented). In addition, this model contains fewer variables that need to be calibrated for the watershed of interest, it runs faster, and it yields output in terms of decision-making criteria that currently are being used in the Washington forest-management arena.

We recognize, however, that a GIS screen built with one model could be replaced relatively easily with another, as science and technology advance and better methods are developed (i.e., “adaptive management” in the GIS world). Hence, Option 3 runs a close second, in our estimation. We strongly support the concept of making both models available and implementable, given that each offers some important potential advantages.

Regardless of which model is chosen, we recommend that both SMORPH and SHALSTAB be simultaneously tested and refined for use in eastern Washington, prior to implementing a statewide GIS cover. The possibility exists that one model could perform significantly better than the other in certain types of terrain. To our knowledge, neither model has been analyzed in terms of its applicability to eastern Washington watersheds. Testing both models simultaneously would not cause delay in creating a statewide coverage because the requisite diagnostic test methods have been established as part of this project.

TECHNICAL REPORT

Comparison of GIS-based Models of Shallow Landsliding for Application to Watershed Management

Susan C. Shaw and Laura M. Vaugeois
Washington Department of Natural Resources
Forest Practices Division
P.O. Box 47012
Olympia, WA. 98504-7012

1.0 Introduction

Land managers and regulators in the Pacific Northwest historically have possessed limited means for evaluating landslide potential where land-management activities are proposed. Existing information on site characteristics and failure potential typically has been confined to small geographic areas (e.g., 20 km^2 or less) in which landslide inventories, geomorphic research, or serni-empirical stability analyses have been conducted. More recently, private landowners and natural-resource agencies in Washington State have initiated a regulatory form of watershed analysis (Washington Forest Practices Board, 1995) for specific landscape units (i.e., Watershed Administrative Units (WAUs), usually less than 200 km^2 or 78 mi^2 in size), in which landslide inventories are developed largely with the aid of aerial photographs and limited field reconnaissance. Landslide assessments in only about 60 of the 764 Watershed Administrative Units, however, have been finalized and approved by the state during the last seven years (Washington Department of Natural Resources (WDNR), 1999). Furthermore, incomplete and often imprecisely mapped state soil surveys and their slope-failure ratings still constitute the main source of information used by state regulatory foresters to evaluate management proposals in areas outside of those where reliable landslide assessments have been performed.

In recognition of these and other management needs for improved methods of identifying landslide sites in Washington State, the Washington Forest Practices Board (WFPB) and Governor's Salmon Recovery Office recently adopted new measures for forest-management activities that include the use of a GIS (Geographic Information System) -based topographic model as a statewide screen for predicting potential unstable slopes (USDI Fish and Wildlife Service et al., 1999). GIS-driven models, using digital elevation model (DEM) data, typically combine empirical and theoretical methods for evaluating the relative role of topographic control (e.g., gradient and slope form) on initiating shallow landslides (e.g., Montgomery and Dietrich, 1994; Shaw and Johnson, 1995; Wu and Sidle, 1995; Wu and Abdel-Latif, 1997; Pack et al. 1998; D. Miller, Earth Systems institute, pers. comm.). Depending on the model used, output can vary from spatial distributions of steady-state rainfall predicted to cause slope instability (e.g., Montgomery and Dietrich, 1994), to landslide-hazard potential based on factors of safety (e.g., Wu and Abdel-Latif, 1997), to landslide-hazard rankings based on management criteria defined by the WFPB (e.g., Shaw and Johnson, 1995). These maps can be useful to managers for screening potential landslide areas and determining where land-use or habitat-restoration activities should be concentrated, to regulators as a replacement to the soil surveys for assigning forest-practices class designations (i.e., determining whether environmental checklists or impact statements are required), and to analysts for developing preliminary hazard-zonation maps that reflect initial hypotheses regarding the location and density of shallow landslides, Isolated tests of GIS-based models in the Pacific Northwest have suggested that preliminary landslide-failure or hazard-zonations maps can provide more accurate slope-stability information than customarily can be interpreted from topographic, geologic, or soil maps alone (e.g., Shaw and Johnson, 1995; Montgomery et al., 1998).

In this paper, we present the results of a comparative test of GE-based models of shallow landsliding for use in a management context. This test was conducted under contract to the Washington Timber/Fish/Wildlife (T/F/W) Program (i.e., a cooperative group of regulatory, tribal, environmental, and industrial sponsors who collectively makes recommendations to the WFPB on matters related to forest management; T/F/W, 1992) and Washington Forest Protection Association (WFPA), as a precursor to developing the statewide slope-stability screen required by the WFPB. For the purposes of comparison, we use data on existing and potential shallow landslide sites from eight watersheds in western Washington (i.e., west of the Cascades Range crest) to examine the ability of each model to predict the spatial distribution of shallow landslides. A similar test currently is being developed for watersheds in each of the distinct geomorphic provinces in eastern Washington, as groundwork for creating a statewide screen of shallow landsliding. In addition to evaluating method accuracy and limitations, we discuss management applicability and several technical criteria important in making models accessible to natural-resource managers and technicians.

2.0 Description of Test Models

Three GIS-driven models have been selected for this evaluation, based on their current availability, potential for adaptation to management decision-making, and/or use by T/F/W cooperators in field applications or previous tests of model performance. They are the current statewide soil-stability screen, maintained by the WDNR and herein labeled SOILS; the shallow landslide model of Montgomery and Dietrich (1994), nicknamed SHALSTAB by its authors; and the shallow landslide model of Shaw and Johnson (1995), herein referred to as SMORPH.

The three selected models have a number of elements in common. They use geographic irrformation systems (GIS) to couple DEM data with assumptions regarding topographic attributes that influence slope destabilization and with algorithms for calculating slope stability. Whereas the SHALSTAB and SMORPH models assume that topographic relief (i.e., hillslope gradient) and form (i.e., slope curvature) are the principal driving factors in promoting shallow landslides, the SOILS screen assumes that only gradient is a critical variable. These assumptions derive from previous studies suggesting that shallow landslides occur most often above a threshold gradient and in topographic convergences where shallow subsurface flow concentrates, such as hollows and channelized depressions, with consequent effects on soil moisture and strength (e.g., Dietrich and Dunne, 1978; Swanson et al., 1981; Swanson and Fredriksen, 1982; Sidle et al., 1985; Montgomery and Dietrich, 1994). This simplifying assumption permits a number of key slope-stability factors to be treated implicitly, including substrate type, bedrock structure, rainfall duration and intensity, soil depth, soil conductivity and strength, plant transpiration, root strength, and subsurface drainage properties.

In addition, each model is limited similarly by the accuracy of the DEM data; that is, these models are only as good as the DEMs on which they are based. Much of western Washington is mapped with DEMs at a IO-meter resolution. For regions in which DEMs are available only on a 30-meter grid, however, all models suffer correspondingly in their precision and accuracy, as discussed in section 3.2 of this paper.

The three model differ primarily in the sophistication with which independent physical parameters affecting slope stability are addressed. The SOILS screen relies on hillslope gradient and soil type to rate slope-stability potential (WDNR, 1988). The SMORPH model explicitly treats gradient and slope curvature, while the SHALSTAB model treats these topographic attributes as well as several key soil physical and hydrological properties. From the standpoint of practical application, there are advantages and disadvantages to each approach. Simpler models in which key influencingfactors are treated implicitly can be employed readily (i.e., with little to no data collection) and for larger geographic areas. The level of site-specific: accuracy, however, might be reduced by assuming static or invariant hydrologic and geomorphic conditions, and by extrapolating local data on soil and hydrologic properties to the basin or regional scale. The advantage of explicitly treating parameters such as rainfall, subsurface hydrology, and soil properties is that the model might identify patterns of potentially unstable ground at a higher resolution. Consequently, such models are useful for predicting site conditions in the local area for which the input data apply. Conversely, employing local data might limit the ability of the model to predict accurately the spatial distribution of unstable slopes at a landscape scale. This approach also requires considerably more data collection in the field. Some factors, for example subsurface hydrologic and soil strength properties, might be very difficult to analyze and measure due to their spatial and temporal variations and their complex physical interactions.

The following paper sections summarize the salient features of the three test methods, in order of relative sophistication, and current knowledge of the authors regarding their application to forest management.

2.1 SOILS screen

This GIS cover, crested by WNDR staff in 1988, expresses for each DEM cell, the relative potential for slope destabilization (i.e., low, medium, high, very high potential for shallow landsliding). It is based on the state soil survey classifications of soil type as stable, unstable, or very unstable (WDNR, 1984) and differentiation between steeper slopes (30% to 65%) and less steep slopes (less than 30%). For example, soil mapping units are rated as having low potential if they are classed as stable soils and fall on hillslopes with maximum gradients of

30%. Soil mapping units are assigned a very high mass-wasting potential if they are classed as unstable soils and fall on hillslopes greater than 65% (Table 1).

The SOILS screen was built as a GIS cover for Washington State in 1988 and has been used since that time in a variety of management contexts, including timber harvest planning by private landowners and state-land managers. The SOILS screen remains the primary database used by state regulators in evaluating the slope-stability potential of areas for which forest practices have been proposed. It is available in the public domain on the WDNR-GIS system.

2.2 SMORPH model

The SMORPH model outputs, for each DEM cell, the relative potential for shallow landsliding in terms of hazard ratings of low, moderate, and high (Shaw and Johnson, 1995). This model assumes that hillslope gradient and form are the primary driving factors for shallow landslides and that other critical influencing factors are treated implicitly by calibrating the model with observed landslide densities. For example, it assumes that the greatest density of landslides occurs on steeper, more convergent slopes; hence, a high hazard rating is given to slope segments with the largest area of unstable ground per unit basin area. The model combines an analysis of digital elevation models with an empirical algorithm that expresses stability classes on the basis of measured landslide densities, as obtained from mass-wasting inventories in terrain with similar geologic, climatic, hydrologic, and vegetative regimes. Required model inputs are DEM data and a histogram of slope gradient versus density of shallow landsliding for the geographic area of interest. The model is used most effectively to extrapolate from areas with mapped landslides to those with little or no landslide data.

A modified version of the Arc/Info^{tm} GRID curvature tool (Environmental Systems Research Institute (ESRI), 1992) is used to evaluate slope gradient and form (planar, concave, convex), and a slope-morphology matrix is formed by the union of gradient and curvature (Table 2). This tool calculates the curvature of a surface at each cell center of a DEM grid, by evaluating hillslope gradient, aspect, pianform curvature (i.e., measured transverse to slope direction and influences subsurface flow concentration or dispersal), and profile curvature (i.e., measured normal to slope direction and governs flow acceleration and deceleration). The mathematical derivation of curvature used in the ESRI package is developed by Zevenbergen and Thorne (1987), in which curvature is given as the divergence of the gradient, or the LaPlacian of the topographic surface, Z, as described by a fourth-order polynomial of the form:

$$\text{Curvature} = \nabla^2 Z = \nabla^2(Ax^2y^2 + Bx^2y + Cxy^2 + Dx^2 + Ey^2 + Fxy + Gx + Hy + I). \qquad [1]$$

The 9 elevations of a 3x3 matrix of surface cells are used to calculate parameters A through I. Matrix elements are assigned management hazard calls of low, moderate, and high based on criteria defined in the landslide inventory used to calibrate the model (e.g., hazard ratings assigned by the analyst during watershed analysis). Hence, model output comprises a preliminary hazard-zonation map, with DEM-scale resolution, that can be used in management decision-making or as a tool for planning a thorough field investigation of landslides.

This model was created specifically as a preliminary screening tool for field foresters and managers to use in landscape and timber-sale planning (Hoh Tribe and WDNR, 1993). It has been tested fairly extensively on the Olympic Peninsula (Shaw and Johnson, 1995) and less rigorously by other T/F/W cooperators elsewhere in the state. This model also was employed in an economic analysis of the habitat conservation plan for state-managed lands in western Washington (WDNR et al., 1997), to estimate the percentage of watershed areas that could be classified as having potentially unstable ground. Several model versions also have been distributed to government agencies and private timber companies in five western states; to date, however, no test results have been reported in a statistically meaningful manner

SMORPH is available to the public from the WDNR:

2.3 SHALSTAB model

The SHALSTAB model outputs, for each DEM cell, the relative potential for shallow landsliding in terms of steady-state rainfall required to fully saturate the soil mass (Montgomery and Dietrich, 1994). It couples DEM data and a planar infinite-slope stability model with a hydrologic model (TOPOG; O'Loughlin, 1986) that predicts near-surface throughflow in topographic elements identified by the intersection of topographic contours. Critical rainfall, Q_c, necessary to saturate soils and initiate soil movement is expressed as:

$$Q_c = T\sin\theta\ (a/b)" \left[c'(\rho_w g z \cos^2\theta \tan\phi)^{-1} + (\rho_s/\rho_w)(1 - \tan\theta/\tan\phi)\right] \qquad [2]$$

where T is the depth-integrated soil transmissivity, 8 is the local slope, a is the upslope contributing area, b is the slope length across which subsurface flow is accounted for, c' is the effective soil cohesion as governed by root strength, ρ_w is the bulk density of water, g is gravitational acceleration, z is soil thickness, ϕ is the internal angle of friction of the soil, and ρ_s is the bulk density of the soil (see Montgomery et al., 1998; their equation 5a).

This model calculates a numerical value of Q_c required to cause landsliding for each DEM cell. Analogous to the factor of safety, Q_c values are assigned a slope-stability risk factor (i.e., unconditionally stable, unstable, stable, and unconditionally stable; Table 3). DEM cells are classified as unconditionally stable when they occupy fully saturated soils on slopes less than some value that is dependent on the soil friction angle and bulk density specified in the model (e.g., $\phi = 33°$, $\rho_s = 2000\ kg/m^3$ in model tests described in Montgomery et al., 1998): tan8 $\leq \tan\phi\ [l - (\rho_w/\rho_s)]$ Conversely, DEM cells are designated unconditionally unstable when soils are dry and slopes are greater than the gradient threshold value: tan8 > tan@. For practical applications in which a hazard-potential rating is desired, the user must translate these stability terms into rnanagement criteria (e.g., low, moderate, high) based on an empirical knowledge of instability or other diagnostic criteria. In this paper, we offer one option for creating management criteria from model output.

Most model applications that have been published to date (Montgomery and Dietrich, 1994; Montgomery et al., 1998) have held soil properties and hydrologic variables constant (i.e., soil depth, internal angle of friction, and transmissivity, and effective soil cohesion; see paper section 3.3 for additional discussion). This method reduces the functional elements of the model to those related to topography (i.e., gradient and curvature) and area (i.e., contributing area upslope of each topographic element).

Model results have been compared by the authors with landslide inventory maps for small coastal catchments in northern California, central Oregon, and the western Olympic Peninsula (Montgomery and Dietrich, 1994). In addition, Montgomery et al. (1998) have tested model performance in 14 watersheds for which landslide inventories have been compiled. SHALSTAB is available from the authors and at the Internet Web site of the University of Washington.

2.41 Other models not selected for this study

A number of other models were considered but not chosen for this comparative test because of availability and software-development issues. They include shallow landslide models of WLI and Sidle (1995), Wu and Abdel-Latif (1995, 1997). Pack et al. (1998), and (Earth Systems Institute, pers. comm.). Other methods were too site-specific to be applied over large geographic areas, as required of a watershed analysis or statewide landslide screen (e.g., LISA and DLISA; Hammond et al., 1992). For a general review of analytical methods other than GIS-based modeling, see literature reviews in papers by Montgomery and Dietrich (1994)' and Wu and Sidle (1995).

Although not available or testable in their current form, these models show promise for future management applications, in that they explicitly treat a number of the problematic spatial and temporal distributions in critical slope-stability factors. Better physical characterizations of these factors could improve predictive capability of GIS modeliing techniques beyond those employed currently in SHALSTAB and SMORPH. Any of these models reasonably could be developed as a GIS slope-stability cover should they prove in future to yield more accurate predictions of landslide potential. Similar to SHALSTAB, several of these models (e.g., Wu and Sidle, 1995; Wu and Abdel-Latif, 1997) would require an additional algorithm that instructs the user on translating model output into management criteria (e.g., low, moderate, high hazard). These models are summarized in subsequent paragraphs, to illustrate their similarities and dissimilarities with the models used in this comparative test (i.e., SHALSTAB and SMORPH).

The dSLAM model (Wu and Sidle, 1995) currently is not available for public use and, hence, could not be evaluated fully. It couples DEM data with a planar infinite-slope stability model, a hydrologic algorithm that simulates groundwater movement as kinematic waves through topographic elements similar to those constructed in the SHALSTAB model, and an algorithm that explicitly characterizes root strength. Whereas contributing rainfall is treated as steady-state in the SHALSTAB model, this model can accommodate spatially constant but temporally varying rainfall input (i.e., single or multiple storm events). Hence, the model must calculate a factor of safety in time steps to simulate the measured rainfall patterns. The model requires as input site-specific data on soil properties, vegetation type and age, and individual storm hyetographs (e.g., actual or simulated). Consequently, this model is computationally more complex and labor-intensive than the SHALSTAB model. Outputs of these model simulations are shown as landslide and debris-flow-path location maps, factor-of-safety distributions, and distributions of failure (i.e., hazard) potential. Management criteria (i.e., low, moderate, high "hazard") must be assigned by the user based on local knowledge. The dSLAM model has been evaluated by the authors on its ability to reproduce physical characteristics of measured landslides in a small tributary drainage in the Oregon Coast Range.

The shallow landslide model of Wu and Abdel-Latif (1995, 1997) currently is not programmed to run on one operating system (T.H. Wu, pers. commun.) and, hence, was not accessible for the purposes of comparing GIS models in the Arc/Info"' environment without additional programming work. This model operates similarly to SHALSTAB, by calculating water-table heights in hillslope elements based on DEM data, and applying them to infinite-slope calculations of factors-of-safety. The slope units in which the water-table heights are derived can be of varying size and are chosen by Microlmage MIPS (Map and Image Processing System). The hydrologic model component (Wu et al., 1993) is based on a lumped-parameter, kinematic storage model using a first-order, second-moment approach to allow for stochastic soil-hydrologic properties (Reddi and Wu, 1991), in which the mean and variance of model output are determined from the mean and variance of model input. Rainfall and/or snowmelt is used to generate piezometric levels of corresponding recurrence intervals. The piezometric input is added to soil-strength properties to generate probabilities of failure for each slope element. Model output is a map showing ranges of failure probabilities (e.g., <0.01, 0.01-0.05, 0.05-0.10, >0.10) for water inputs of a given recurrence interval. Such maps can be improved by using smaller slope elements, more data on soil properties, and updating with empirical landslide information. This model has been used to generate hazard maps for two USGS 7.5' quadrangles in Lewis County, Washington, and compared with landslide maps of that area generated by Dragovich and Brunengo (1995).

The SINMAP model of Pack et al. (1998) is freely available on the Internet. We were not able to resolve with the authors a potential problem with the hydrologic component of the model during the course of this study and, hence, could not complete a test of model performance. This model is operated in GIS ArcView and couples the infinite-slope model with a topographically based, steady-state, hydrology model. The model requires no input data for soil, vegetation, and geologic factors known to influence slope stability; rather, these critical parameters are modeled as uniform distributions between empirically derived limits. The user may "pick" appropriate values for a specified watershed based on the ability of the resulting output to "capture" a high proportion of observed landslides and minimize the number of incorrectly identified sites (i.e., areas in which no landslides have been observed). Hence, model calibration requires the use of landslide inventory data, similar to the SMORPH model. Slope stability classes (e.g., low, moderate, high) are assigned based on a slope plot of landslide and non-landslide points.

The shallow-landslide prediction method of Miller (Earth Systems Institute, pers. comm.) was not fully developed in time to be included in this comparative test. Their method couples a modified form of the SHALSTAB model with a debris-flow-runout algorithm (Benda and Cundy, 1990) that predicts the potential delivery of landslide materials to the stream-channel network. This algorithm adds substantially to the management applicability of this method. For example, in the Washington State regulatory context, a management "hazard" is defined as the "likelihood of deliverability and adverse change to public resources" associated with a forest-practices activity (WFPB, 1995; Chapter 222-22). Assigning a management rating, therefore, requires that the identified landslide be assessed to determine whether mass-wasting debris entered stream channels and was delivered to a reach with sensitive public resources (e.g., fish habitat), Hence, the debris-flow component might assist managers, particularly in addressing landslide impacts on downstream resources. The model author also has modified the steady-state rainfall criterion to calculate critical rainfall intensity as a function of storm duration. Similar to the SINMAP algorithm for assigning landslide hazard calls, the model defines hazard on the basis of critical rainfall intensity and soil-parameter values required to "capture" 90% of the observed landslides in a given basin. Hence, this method also requires the use of landslide inventory data to calibrate slope-stability predictions and assign management criteria.

3.0 Methods

3.1 Study areas and landslide data

We chose eight areas in western Washington (Figure 1) for this comparative test. The test basins range in size from 81 km^2 to 331 km^2 (Table 4). Existing Watershed Administrative Units (WAUs) were used as the test-basin boundaries, wherever possible. WAUs, defined for the purposes of regulatory watershed analysis, typically follow major drainage divides; the larger-order river systems, however, may be divided into several WAUs to limit the watershed analyses to a maximum acreage that reasonably could be assessed in the limited time period permitted by law (WFPB, 1995). Hence, *some* of our test basins comprise *only* the upper or mid- sections of a major river system (e.g., Chehalis Headwaters WAU, Middle Hoh WAU). Preference was given to those WAUs with recently completed watershed analyses, to utilize existing databases and to take advantage of the standardized format of data collecting used in this regulatory process.

One test basin (i.e., Morton) was created from portions of two existing WAUs (i.e., East Fork Tilton and Nineteen Creek) to accommodate data restrictions imposed by one of the models (i.e., Wu and Abdel-Latif, 1995, 1997) that was to be tested. As described previously, this model was incompletely programmed at the time of this study. Nonetheless, we continued using the Morton basin for tests of other models, since all landslide data had been compiled in preparation for testing the Wu and Abdel-Latif model.

We attempted to include at least one test watershed in each of the major geologic provinces in western Washington (Table 4; Thorsen, 1978). Parent materials range from glacial till/outwash and lightly metamorphosed sediments to volcanics and igneous intrusives. Test basins also vary in topographic relief (i.e., lowest to highest elevation points) from 818m., in the Chehalis Headwaters basin, to 1941m. in the Jordan-Boulder basin. Figure 2A and 2B demonstrate the relative relief differences between two test basins and show the spatial distribution of shallow landslides identified in recent watershed analyses. Our intent was to examine model performance in areas with different combinations of relief and parent materials, as a means for exploring model versatility and the feasibility of using each model as a management tool in diverse topographic and geologic settings. An apparent gap exists in our selection of test basins, between the North Fork Stilliguamish and Morton watersheds (Figure 1). The central Cascades Range, roughly from the Snoqualmie River basin south to the Morton area, however, generally contains similar geologic units (i.e., rhyolitic to dacitic volcanics with associated clastics, intrusives, and scattered sedimentary basins; Schuster, 1992). Hence, we chose the Stilliguamish, Hazel, and Jordan-Boulder basins to represent the Cascades geologic units north of the Snoqualmie basin, and the Morton and East Fork Lewis basins to represent those to the south.

The eight test basins contain a total of 2524 known landslides (Table 4). including shallow and deep-seated landslides (i.e., earthflows). We retained data on deep-seated landslides (e.g., earthflows) in the test database to evaluate the ability of each model to predict shallow landslide features that often are superimposed on more areally extensive earthflows.

Predictions of unstable slopes made by each GIS model were compared with existing landslide inventories and, where possible, hazard-zonation maps. Figures 3 A, B, and C illustrate three different test basins and show the existing SOILS screen with landslide-inventory data superimposed, the hazard-zonation maps from watershed analyses, and the model outputs from SHALSTAB and SMORPH. Note the differences in the geographic extent of the SOILS cover (Figures 3A versus 3C), and the variations in mapping styles used in hazard-zonation maps (Figures 3A versus 3B).

Existing digital landslide inventories were acquired from the appropriate landowners in the test basins where watershed analyses had been performed (Table 5). Where inventories were not current or were spatially incomplete (i.e., original inventories covered only portions of the test area), we conducted aerial-photograph and field surveys to fill in data gaps. Aerial-photo series extended from the mid-1940's through 1996, in most instances. All inventories were updated chronologically to include, at a minimum, the most recent storm event known to have triggered widespread landsliding throughout Washington State (i.e., the high-intensity, long-duration storm of February, 1996; Gerstel, 1996). In addition, most inventories were checked in the field to verify database accuracy (e.g., landslide type, location, size). Road-related failures were retained in the test database, to evaluate the theory (e.g., Montgomery et al., 1998) that their locations are governed largely by hillslope gradient and topographic convergence. Standardized field data-forms were designed similar to the those used in the mass wasting assessment of the regulatory watershed analysis (WFPB, 1997, Appendix A). Newly identified landslides were mapped on to 1:24,000 scale topographic maps and then digitized into the GIS (Arc/Info^{tm}, version 8.0, for UNIX on a Solaris platform), coded, and edit-checked for positional and tabular accuracy.

In some cases, we updated the landslide inventories to include small landslides (i.e., less than $100m^2$) that might have been omitted due to time and mapping-resolution limitations that customarily constrain the regulatory watershed-analysis process. We increased the number of recorded landslides on these inventories by an average 12%, during our field and aerial-photo verifications of the databases. In the Upper East Fork Lewis River watershed, for example, our reanalysis of the GIS landslide-inventory cover maintained by the USFS resulted in a 70% increase in the number of recorded landslides. Hence, the watershed-analysis-derived landslide inventories really only provide a lower limit on the number of landslides present during the time period evaluated by the analyst (i.e., typically coinciding with the aerial-photo record). Consequently, landslide inventories were used here only as a common basis for comparing model abilities to predict known contemporary landslides, recognizing that other shallow landslides have been overlooked or perhaps no longer can be discerned in the field and photo records due to such obscuring factors as vegetation regrowth. Additionally, we assumed that hazard-zonation maps, if carefully constructed, capture a fair percentage of topographic features that could have influenced landslide initiation in the more distant past.

All inventory data were projected into Washington State Plane, south zone, North American Datum 1927. Having all data in the same projection allowed us to easily incorporate other existing data (e.g. hydrography, transportation), as well as provide a uniform projection from which to work.

We encountered a number of problems with existing landslide data while updating and verifying mass-wasting inventories from the completed, regulatory watershed analyses. These included incorrect basemaps on which landslides were recorded, as well as incorrectly mapped landslides. Discrepancies between USDI Geological Survey (USGS) topographic maps and basemaps created from GIS for use in watershed analysis typically included differences in topographic-contour delineations and stream-channel positions. Keying landslide locations to these features on USGS topographic maps, for example, apparently cause a positional offset when data are transferred to GIS DEM-based topography. A number of mapping errors also appeared to be related to inaccurate transfer of field data onto basemaps or incorrect digitizing from basemaps. In the Sol Duc watershed, for example, we determined from a reassessment of aerial photographs that several landslides were mapped in tributaries adjacent to the ones in which they actually exist. Hence, we remapped and redigitized landslides wherever we encountered such discrepancies during field or aerial-photo verification.

Another common mapping problem is related to landslide size. Mapping techniques used by watershed analysts ranged from representing landslides as a point or symbol (e.g., circle) to delineating slides as polygons of finite area. The latter technique also included a range of mapping styles, from mapping the failure scarp separately to delineating the entire portion of slope involved in landsliding (e.g., some combination of the contributing area, initiation point, transport zone, debris-flow runout track, and depositional area), generally accompanied by little or no explanation of mapping style. In addition, landslide mapping is prone to some amount of inaccuracy, given that data are transferred between a number of different media (e.g., photos, maps, digital databases) with varying levels of resolution and precision, and often between different workers (e.g., field technicians, analysts, cartographers).

To address problems of mapped landslide location and size, we created a buffer around landslides mapped as points or symbols, or polygons smaller than $100m^2$. The buffer, mapped as a polygon of radius 15m. (50 ft.) around the presumed center of the landslide feature, assured that landslides registered in a $100m^2$ DEM grid cell when inventory data were compared with GIS model output. In many cases, landslide scarps and bodies were remapped, during aerial-photo and field verification of the existing databases, to exclude associated features (e.g., contributing areas and debris-flow runout tracks). The landslide polygons then were joined with

Smorph cells

Most Sensitive to Mass Wasting.

Moderately Sensitive to Mass Wasting.

Least Sensitive to Mass Wasting.

Figure 10. Map of the Jordan-Boulder basin with SMORPH plotted against SHALSTAB using management criteria. In order to directly compare models, the SHALSTAB model output was categorized into high, moderate, and low hazard. The areas within these categories were then compared with those of SMORPH, which outputs management criteria automatically.

Shalstab using management criteria

Unstable

Moderately Unstable

Stable

3.2 DEM data

Where available, we used DEMs with 10m. grid resolution for the comparative test (Table 5). As discussed in paper section 4.5, we also compared model output using 10m. and 30m. DEM data, to evaluate the relative percent change in area predicted to fall in each slope-stability class and to quantify the increase in computational time that accompanies the use of finer-resolution DEMs.

All GIS-based models described in this study depend heavily on DEMs. DEM problems commonly reported in the literature, and also evident in this study, include resolution and mapping artifacts. DEM data usually are distributed as datasets with borders approximating the boundaries of the original USGS topographic quadrangles, referred to cartographically as tiles. “Tiling” artifacts can occur along the seams between adjacent sets of DEM data (Figure 4, lower left), interrupting the actual represented surface with artificial cliffs along the tile edges. Tile edges often are interpreted by the GIS shallow-landslide models as representing areas of instability. This type of error only occurs at tile edges and does not propagate into the dataset. Tiling artifacts were observed most frequently in the 30-m. DEM data used in this study.

“Edge effects” occur when the outermost grid cells of the study area (i.e., the clipped edges of the DEM) do not have the same general values as adjacent cells (Figure 4, lower right). This phenomenon only affects the outermost two or three grid cells at the edges of the DEM, and it does not propagate into the dataset. To eliminate edge effects in the test databases, all GIS shallow-landslide models were run on a DEM grid larger than the basin area. The model output then was clipped along with the basin boundary.

In the 10-m. resolution data, the elevation values appear to have a slightly stepped pattern, resulting in model output with elevation bands (e.g., contours) of similar predicted velue. Typically, banding results in slope-parallel arcs of one hazard-potential class, within broader polygons of a different hazard-potential class (Figure 4, upper left). The cause of this elevation banding is unknown; however, it may be related to the original elevation-value collection scheme. Elevation data were created by scanning USGS 7.5minute topographic quadrangles, vectorizing the scanned data, coding the vectors, and then assigning x, y, and z coordinate values on a 10-m. grid using linear interpolation. This process may result in some elevation banding. DEM data with 10-m. resolution are not subject to many smoothing filters, as smoothing tends to degrade original topographic data. This lack of smoothing may also have some effect on elevation banding.

Resolution of available DEMs also can be problematic for precisely locating terrain features. As discussed in paper section 4.5, the relative resolution of 10-m. versus 30-m. elevation data creates a better representation of the actual ground surface (e.g., see Figure 5), especially in resolving small stream channels emanating from zero- and first- order basins, common initiation sites for shallow landslides (Dietrich et al., 1986). For example, 30-m. DEMs have a resolution in the x- and y- planes of 30.5 m. (100 ft.) and 45.7 m. (150 ft.). Hence, landslides digitized onto 30-m. DEMs from USGS topographic maps can be positioned more than one DEM grid cell from their true location, resulting in mismatches between spatial distributions of inventoried landslides and DEM grid cells with predicted unstable slopes. In addition, contour splines fit through 30-m. DEM elevation points can lack curvature more characteristic of USGS topographic maps, causing an artificial angularity in topographic features and resulting in relatively poor matches between contour crenulations and stream courses overlain from GIS hydrology layers. The finer resolution of 10-m. data (i.e., 12.2 m. (40 ft.) in the x- and y- planes, and 15.2 m. (50 ft.) in the z-plane), which is similar to that reported for USGS 7.5minute quadrangles, results in a nearly accurate match between DEM-derived map contours and those on USGS maps. Consequently, the potential for matching errors between DEMs and landslides digitized from topographic maps is considerably less when using 10-m. versus 30-m. DEMs.

3.3 GIS model calibration and database development

The SOILS screen required no adjustments to be employed in this study, and in fact cannot be adjusted to accommodate any new information, including altered soil classifications or gradient classes, without significant revamping of the GIS cover. The digital soils database for federal lands, maintained by the USDA Forest Service on the Internet, was merged with that maintained for state! and private lands by the WDNR (1988). Nonetheless, six of eight test basins had incomplete digital soil covers (Table 5), due largely to gaps in soils-layer coverage on federal property (e.g., Figure 6). For statistical analysis of comparisons between the digital landslide inventories and soils slope-stability cover in these test basins, an existing landslide was given a “no data” value where the soils cover was lacking.

The SMORPH model was calibrated in each test basin with its respective landslide-inventory data to adjust the critical slope classes and their hazard-rating designations in the gradient-curvature matrix (Table 2). A slope map derived from the DEMs was intersected with the landslide inventory to determine the maximum gradient found in each landslide polygon. A curve of maximum gradient ‘versus cumulative frequency percent was created (Figure 7), with the lowest gradient at which a landslide occurred being used to determine the lower class limit of the moderate hazard rating. The lower class limit of the high hazard rating was established al: a value for which 15% of the landslides occurred (Table 6), to guarantee a model-prediction rate of at least 85% of observed landslides.

For consistency with other published tests of the SHALSTAB model (e.g., Montgomery et al., 1998), we used the following soil-property values: soil depth (z) = 1.0m; soil bulk density

(ρ_s) = 2000 kg/m^3; internal friction angle (ϕ) = 33"; effective cohesion (C') = 2 kN/m^2; and transmissivity (T) = 65 m^2/day. These values were selected by Montgomery et al. (1998) based on extensive field measurements in a small catchment in coastal Oregon (Montgomery et al., 1997), and the authors felt that they gave reasonable results for their test watersheds in western Washington, including the Chehalis Headwaters WAU that we also use as a test basin. We then compared predictions of unstable-slope potential for the range of ϕ angles and effective cohesions set internally in the model to yield a standard range of outputs (i.e., default parameters; ϕ = 33° and 45°, and c'= 0, 2, 5, 8, 15 kN/m^2), to evaluate the effect of modifying these parameters. In section 4.2 of this paper, we discuss the sensitivity of model output to variations in input values.

Comparing SHALSTAB with the other GIS models required that we reduce all model outputs to a common denominator. SMORPH and the SOILS screen yield output in terms of management hazard ratings (e.g., low, moderate, high), in which the more subjective determination of what constitutes "hazard" and "risk" previously has been made in the policy arena. For example, the SMORPH slope matrix is calibrated with landslide-inventory and hazard-zonation databases created during regulatory watershed analyses for which definitions of hazard and risk have been set by T/F/W policy and WFPB regulations (WFPB, 1995, Chapter 222-22 WAC). Likewise, the SOILS screen hazard designations are derived from unstable-slope ratings in the state soil surveys. In the absence of another mechanism for converting all model outputs to the same units of measure, we therefore elected to assign hazard ratings to the SHALSTAB model output values of predicted critical rainfall, by using rainfall intensity and duration as the diagnostic criteria.

Given that SHALSTAB model output is expressed as rainfall in mm/day, we created "precipitation rules" for each test basin by clipping the two-year, 24-hour storm isohyte data

(VVDNR-GIS; Miller et al., 1973) and computing the-minimum, maximum, and mean precipitation values for each basin. A high hazard rating was given to each DEM grid cell in which the predicted critical-rainfall value fell in the model-defined Q_c -stability class occupied by the mean precipitation value calculated for that basin (Table 7). A high rating was also given to any predicted Q_c less than the minimum two-year, 24-hour calculated precipitation. A moderate hazard rating was assigned to a DEM cell in which the critical rainfall value occupied the Q_c -stability class corresponding to the maximum calculated precipitation. A low hazard rating was assigned to all other Q_c stability classes. See Table 7 for the precipitation rules and slope-stability hazards created for each test basin.

The two-year, 24-hour recurrence interval was chosen as the precipitation regime for which data were readily available and which yielded the most conservative estimate of failure potential. The SHALSTAB model is configured such that the less frequent rainfall event yields a greater percentage of the basin area predicted to fail (Montgomery and Dietrich, 1994). Theoretically, then, a higher-intensity storm event characteristic of a longer recurrence interval, and/or a longer-duration rainfall, would result in greater spatial distribution of potential shallow landslides.

This method of assigning management criteria to SHALSTAB output was chosen in the absence of established techniques or direction provided by the authors (e.g., see discussion of management applications in Montgomery et al., 1998). A preferred approach might be to.adjust the model in each test basin by using measured values of input parameters (e.g., soil transmissivity, bulk density, cohesion, internal friction angle), and calibrating predicted distributions of slope stability with observed landslide inventories and/or associated hazard-zonation maps in which management criteria have been assigned (i.e., similar to the approach used by SMORPH). Adjusting input parameters in the current version of the SHALSTAB model is problematic, given the relative paucity of soil-property data and the current lack of published algorithms for modelling stochastic elements or calibrating them from landslide inventories. Obtaining sufficient soil-parameter samples to adequately describe their spatial variability also could be intractable or prohibitively expensive for creating a landscape or regional GIS cover of predicted slope stability.

Calibrating model output with landslide-potential ratings from hazard-zonation maps is problematic. We found, for example, that hazard map units with different management designations (e.g., low and high) might contain DEM grid cells with the same range of Q_c - slope stability class values (e.g., 2 through 7; see Table 3), making it difficult to segregate the eight model-output classes into discrete management categories of low, moderate, and high. Calibrating model outputs solely on the basis of landslide inventories also can be misleading because, as discussed previously, they typically represent only contemporary rates of shallow landsiiding, thus conceivably underestimating the density of potential landslide sites. Landslide density commonly has been a key factor in assigning management criteria to hazard-potential map polygons created from inventories (e.g., WFPB, 1997).

The precipitation rules imposed by this study make a number of assumptions, not the least of which is steady-state throughflow of subsurface water. The SHALSTAB model, however, is founded on the assumption of steady-state rainfall, constant transmissivity, and spatially uniform soil saturation (Montgomery and Dietrich, 1994). Hence, the steady-state precipitation rules are consistent with these assumptions. As described further in report section 4.1, the similarity of watershed-analysis-derived hazard-zonation maps and maps of landslide hazard potential made with SHALSTAB precipitation rules suggests that this approach yields reasonable results. Consequently, we have subscribed to this method in the absence of a proven alternative.

Figure 8 illustrates SHALSTAB model predictions of shallow-landslide potential in the Jordan-Boulder test basin, using the author-defined Q_c criteria (Figure 8A) and predictions in terms of management criteria as defined by the precipitation rules (Figure 8B). The comparatively large amount of area classified as high "hazard" by the precipitation-rule designations likely is the result of the way in which subsurface water throughflow, and hence "soil wetness" necessary to destabilize slopes, is calculated by the model. In the SHALSTAB program, water can Row through any one of a nurnber of flow tubes that might diverge around topographic high points. Hence, the program codes these flow tubes, including the ones over intervening divergent topography (e.g., narrow ridgelines) as relatively unstable, which, in turn, are classified as high "hazard" by the precipitation rules, When the magnitude of effective cohesion is increased, resulting in less area classified as highly unstable, the model incurs relatively greater error in predicting known, existing landslides. Thus, the model has the potential for erring one way or the other depending on the assigned values of the input variables. One approach for resolving this dilemma would be to iterate on the magnitudes of cohesion until a value is achieved that yields model output most closely resembling the spatial distribution of existing landslides. As an additional note, the juxtaposition of high and low "hazard" units in the lower portion of the figure is not an artifact; this terrain contains very steep, generally unstable slopes that terminate on flat, glaciated valley bottoms (e.g., see Figure 2A).

4.0 Test Results and Discussion

For the purposes of testing and comparing models, a number of criteria are used to evaluate the predictive capability and management applicability of each model. Test criteria have been divided into two categories: scientific and technical. Together with the critical questions that we posed for each model, these criteria are:

[A] scientific:

(1) model performance: How do model predictions of shallow landsliding compare with existing landslide inventories and hazard-zonation maps? How do model predictions compare with respect to each other?

(2) method limitations: Do data input requirements, particularly those dependent on fieldwork, limit the utility of the model? Are model assumptions regarding geomorphic processes or input variables relevant to all western Washington watersheds?

(3) geographic applicability: Is the model appropriate for use in all forested watersheds in western Washington, and can a reliable slope-stability map cover be created for regional or statewide use?

(4) management applications: Can the model be applied to management decision-making, and if so, are they accessible to users?

(5) modification requirements: What additional adaptations must be made to facilitate creating management criteria (e.g., low, moderate, high "hazard") from model output? Could and should the model be modified by its author(s) to improve its predictive capability for all terrain types in western Washington?

[B] technical:

(1) computational time: How long does it take to run the model for an average-sized basin (e.g., on the order of a WAU)? How long would it take to create a GIS cover for western Washington?

(2) training requirements: Assuming basic computer skills, how much training is needed to run the model, interpret model results, and apply results to management problems?

(3) data requirements: Can the model be run with default values for input variables, if field data are nonexistent? Does model accuracy improve with increasing DEM resolution?

(4) data storage and retrieval: Do model runs or their output require large disk-storage space? Can the model be run on a personal computer (PC) with GIS software?

(5) modification requirements: Is the model computer code adequately documented to aid users in adjusting input values or programming management criteria? Are further modifications needed to adapt the model for management use?

In this paper section, we discuss issues [A] (1 through 3) and [B] with respect to the three tested models. Management applications and model modification requirements are discussed in report section 5.0.

4.1 Model performance

We evaluated the performance of each model by using the GIS to intersect the updated, digital landslide inventories and hazard-zonation maps with model predictions of slope stability. For each model, output was expressed in terms of management criteria (i.e., low, moderate, high "hazard"), as described in the report section 3.0, so that model performances could be compared directly. We statistically analyzed the following, as a measure of the performance of each model: (1) intersection of the digital landslide inventory with model predictions of hazard potential, expressed as the number of incorrectly identified landslides per total number of landslides in each test basin (i.e., Type I model errors);

(2) intersection of the hazard-zonation maps with model predictions of hazard potential, given as the percent probability that the model predicts a low landslide potential where it is likely that landslides have occurred or will occur (i.e., Type I model

errors); and,

(3) intersection as in (2) but expressed as the percent probability that the model predicts the potential for landslides where they are not likely to occur (i.e., Type II model errors).

inventories of known existing landslides and maps of hazard potential often are used in different management contexts. For that reason, we calculated Type I errors first by intersecting model outputs with the landslide inventories, to evaluate the ability of each model to predict the spatial distribution of existing landslides. We then computed Type I errors associated with comparing model outputs and hazard-zonation maps, to assess model abilities to predict the spatial distribution of existing and potential slope instability. Given that landslide inventories typically provide only a minimum estimate of contemporary landslide rates, the hazard-zonation maps theoretically yield a more complete view of the spatial distribution of past, present, and potential future landslide occurrences.

Table 8 lists, for each model, the number of incorrectly identified landslides per total number of landslides in each test basin (i.e., Type I errors). We assumed that an existing landslide was identified incorrectly if all DEM grid cells overlapping the landslide polygon or its 15m. (50 ft.) buffer (e.g., see report section 3.1) were coded by the model as having a low potential (hazard) for shallow landsliding. Conversely, an existing landslide was assumed to be identified correctly if any overlapping DEM grid cell was predicted to have a moderate or high potential (hazard) for landsliding. DEM cells with no data entry in the SOILS screen (i.e., missing soil-survey data) were coded as an incorrect identification, to account statistically for the incomplete nature of the data coverage. For this test, the SHALSTAB model was run using default parameters ϕ = 33" and C' = 2 kN/m^2 and assuming that the two-year 24-hour storm recurrence interval is a reasonable criterion for assigning hazard-potential ratings to the model output (i.e., see report section 3.3).

A principal assumption of the model comparative tests is that predictions of landslide probability densities can be cornpared even though the GIS covers contain known mapping artifacts (e.g., elevation banding), as described in section 3.2. Given that model predictions of slope stability are evaluated using the same DEMs and landslide databases, the model outputs could be evaluated relative to one another. However, computed statistics (e.g., average number of landslides incorrectly identified by each model) should be viewed as estimates rather than absolute values, because the errors in model predictions associated with database noise (e.g., DEM elevation banding, field-mapping accuracy and resolution).

Table 8 indicates that the SOILS screen did not identify 32% of the total known landslides in all eight test basins, whereas the SMORPH and SHALSTAB models misidentifed 3% and 8%, respectively. Figure 9 displays the relative range of model predictions with SMORPH versus SHALSTAB. shown as histograms of the number of total landslides predicted in each model-output category. According to the precipitation rules, SHALSTAB classes of Q_c = 1, 2, 3, ± 4 fall in the management-criteria class 3 (i.e., high “hazard”). Montgomery et al. (‘1998) also reported from their test of the SHALSTAB model that it predicted unconditionally stable slopes in 24% of the area containing known existing landslides, although they discounted approximately half of these failures as being road-related or undistinguishable on 30-m. DEMs and, hence, outside the realm of model predictive capability. The use of more accurate DEMs could account, in part, for the relatively smaller fraction of landslides undetected by SHALSTAB in this test. The significantly higher percentage of landslides missed by the SOILS screen can be attributed to the lack or near lack of soil-survey data for two of the test basins (i.e., the North Fork Stilliguamish and Upper East Fork Lewis watersheds; see Table 5), given that missing data were coded as undetected landslides for the purposes of comparing model performances (see report section 3.1). Where the SOILS screenwas complete (e.g., Morton and Chehalis Headwaters watersheds), however, it misidentified a significantly higher percentage of landslides than the other two models (e.g., for the Chehalis Headwaters watershed, 32% versus 2% each for the SMORPH and SHALSTAB models).

In the Olympic Peninsula test basins, the SOILS screen misidentified more landslides than SMORPH but fewer than SHALSTAB (e.g., in the Hoh watershed, 67 versus 53 and 84, respectively). The fact that these were the only basins for which 30-m. DEMs were used was ruled out as a likely cause. In other test basins for which model results were compared using both 10-m. and 30-m. DEMs, there was no change in the ordering of models based on their predictive accuracy, although the relative magnitudes of predicted landslide occurrence (i.e., number of correctly identified existing landslides) differed between 10-m. and 30-m. DEM test results for each model. Hence, the seemingly better performance of the SOILS screen might be explained by at least two compounding factors. One is that, for the portions of the test basins in which soils data exist, the SOILS screen classes 68% of the Sol Duc and 84% of the Hoh basin terrain as potentially unstable or very unstable, so that the majority of the landscape and its associated landslides fall within the high-hazard-potential category. Although this result lends the appearance that the SOILS screen more closely reflects the spatial distribution of known landslides than does SHALSTAB. it also tends to over-predict significantly the percent of watershed area predicted by field-derived, hazard-zonation maps to be potentially unstable (see further discussion of the SOILS screen in this paper section).

Another compounding factor is that the SOILS screen and SMORPH model consider hillslopes as being potentially unstable at gradients somewhat lower than the threshold gradient defined in the SHALSTAB model. In the latter model, slopes are considered unconditionally stable when $\tan\theta \le \tan\phi\,[1 - (\rho_w/\rho_s)]$ which, for $\phi = 33°$ and $\rho_s = 2000\ kg/m^3$, means any slopes less than 18" (32.5%). Field evidence suggests that non-road-related shallow landslides have occurred in this region on slopes closer to 25% (e.g., Shaw and Johnson, 1995; D. Parks, WDNR, pers. comm.), particularly in gently sloped, groundwater-seepage areas whose downslope margins coincide with the top of steep, inner-gorge slopes, which are quite common in this terrain. Hence, the SHALSTAB model has the potential for under-predicting the spatial distribution of unstable ground on hillslopes with gradients less than the threshold value set internally by the model.

The SMORPH model predicted an average of 22 times fewer Type I errors than the SOILS screen and five times fewer than the SHAL.STAB model. The greatest discrepancy in SMORPH and SHALSTAB model predictions occurred in the Hazel watershed (1% versus 32% incorrectly identified; Table 8). Given that the Hazel watershed is dominated by deep-seated landslides in thick glacial deposits (Table 4). we expected the predictive capability of both models to diminish correspondingly, with respect to locating earthflow-influenced topography. It appeared, however, that SMORPH was better able to distinguish the local slope and curvature of numerous shallow-landslide headscarps superimposed on the larger earthflows. Hence, the polygons representing deep-seated failures effectively were identified by SMORPH predictions of high hazard potential on the basis of these smaller secondary features.

This variation in results might be explained by the manner in which the two models identify "hazard" potential in adjoining DEM grid cells. The SMORPH model analyzes variations in topographic relief between adjacent cells based on their relative steepness and curvature, then assigns a value according to the slope matrix (Table 2); hence, the model can discern topographic changes between a flatter upslope cell and a steeper downslope cell (i.e., a landslide headwall). On the other hand, the SHALSTAB model can smooth (i.e., not detect) subtle variations in topographic relief at the DEM-cell scale, by assigning a given flow tube a Q_c

value depending on the flow across its upper boundary (i.e., variable “a” in Equation 2) from upslope contributing areas, which, in turn, is governed by the way in which flow is dispersed from that contributing area to any one of a number of downslope grid cells. Hence, if the upslope contributing area has a lower gradient and requires a relatively higher water flux to create “wet” soils, then a relatively steeper cell downslope (e.g., a landslide headwall) might not be predicted to fail until the same “wetness” is achieved. Hence, the grid cell downslope of the contributing area is given a lower slope-stability rating, whereas SMORPH assigns a higher value based solely on topographic factors.

Although Table 8 indicates that SMORPH yielded 43% fewer Type I errors in predicting known landslide occurrences than SHALSTAB (Table 8), we wanted to evaluate whether these differences in model performance, based on a comparison in eight watersheds, were significant statistically. We used a non-parametric test for non-normally distributed, small, independent samples to evaluate the hypothesis that there is no difference in the average performance of the SMORPH (SM) and SHALSTAB (SH) models, in terms of their ability to predict the spatial distribution of known landslides. The null hypothesis is that the means (μ) of the population of Type I errors for each model are equal when only eight independent samples (i.e., test basins) exist; H,: $\mu_{SM} = \mu_{SH}$. Equality of means was tested with the Wilcoxon rank-sum statistic for two populations (Walpole, 1974; MathSoft 1998), in which the null hypothesis was true if:

$$Pr\,[W \leq w = (a - n(n+1)/2)] > a,$$

where Pr is the probability distribution, W is the test statistic, a is the smaller of the summed ranks for each population, n is the number of observations corresponding to a, and a = 0.01, 0.05 is the level of significance. Table 9 indicates that the test statistic is significant at a confidence level of 95%, permitting rejection of the null hypothesis, which suggests that the models differ somewhat in their ability to predict known landslide distributions; that is, $\mu_{SM} < \mu_{SH}$.

However, the test statistic proved insignificant at the 99% confidence level (Table 9), allowing. acceptance of the null hypothesis and implying that the difference in model predictive capability is relatively small. A similar statistical comparison of SHALSTAB and the SOILS screen indicated that the test statistic was significant at the 99% confidence level, implying that the screen and model are considerably different in their ability to predict existing landslide distributions.

Tables 10 and 11, respectively, give the estimated Type I and Type I I model errors for the SMORPH and SHALSTAB model based on a comparison of model output with hazard-zonation maps. Error distributions were not computed for the SOILS screen, given that soils-survey data were complete in only two of the test basins, neither of which had usable hazard-zonation maps. Type I errors were calculated, for each model in each test basin, by intersecting the low-hazard DEM cells predicted by the model with the moderate- and/or high-hazard map units produced via watershed analysis (i.e., incorporating all map units intersecting with known landslides in the GIS inventory layer). This database intersection was expressed numerically as a percentage of model-predicted, low-hazard areas (in km^2) overlapping field-mapped hazard areas. Type II errors similarly were analyzed by intersecting the high-hazard cells predicted by the model with the low-hazard map units and computing respective areas, These estimates were made for the four basins in which we had access to complete, digitized, hazard-zonation maps. To facilitate comparison (see Table 10 and 1 I), the percent error for each model (A/M) in each basin was normalized by the basin area in a given hazard class (A) divided by the total A for all four basins (T), that is: E = (A/M)(A/T).

Analysis of Type I error estimates with respect to hazard-zonation maps indicates that the SMORPH and SHALSTAB models similarly under-predict the percent area of hazard map units determined to be of moderate and/or high failure potential, by an average 6% and 5%, respectively. Using the Wilcoxon rank-sum statistic for two populations, as described previously, the computed test statistic proved insignificant at the 95% confidence level (Table 9), implying that the models perform similarly in predicting areas of relatively low hazard potential inside mapped landslide-hazard areas.

Whether the observed discrepancies between model predictions and hazard-zonation map units represent true " Type 1 errors" in the statistical sense is debatable, given that three of the four hazard-zonation maps (i.e., Jordan-Boulder, Hazel, and Sol Duc River) were drawn using broad map polygons (e.g., Figure 3B, lower left) that incorporated both unstable slopes and intervening stable ground. In the Jordan-Boulder basin, for example, hazard-zonation units intentionally were drawn to include potential landslide sites (e.g., hollows, groundwater seeps, inner gorges) and intervening divergent topography (e.g., ridge lines) because it was not possible to delineate them on 1:24,000 scale maps (Coho, 1997). Hence, the GIS-based models might discriminate, more accurately than the hazard-zonation maps, the topographic features potentially influencing shallow landslide initiation in finely dissected terrain.

As a test of the influence of mapping resolution on hazard zonation maps, we intentionally created the hazard-zonation map units in the East Fork Lewis test basin with as fine a resolution as possible on 1:24,000 scale maps. This allowed us to compare model predictions with two different scales of hazard-map resolution (e.g., the Jordan-Boulder basin, Figure 3B, lower left; and East Fork Lewis basin, Figure 3A, lower left). Type I "errors" generated by SMORPH and SHALSTAB decreased substantially, from 14% and 9% for the Jordan-Boulder basin, respectively, to 1% and 2% for the East Fork Lewis basin (Table 10; values normalized as described previously). One implication of this result is that GIS-based model predictions of slope-stability potential could be used advantageously by analysts in drawing hazard-zonation maps with higher resolution than demonstrated, for example, in Figure

3B.

Table 11 shows the distribution of Type II errors generated by the SMORPH and SHALSTAB models, based on comparisons with hazard-zonation maps. As in Table 10, error values are given as normalized relative percent areas. Calculated error estimates for each of the test basins suggest that SMORPH over-predicts the percent area of hazard-zonation map units designated as high landslide potential, by an average amount slightly less than predicted by SHALSTAB (i.e., 3% versus 7%, respectively). In all four test basins, SHALSTAB tended to over-predict, by a factor of two greater than SMORPH, the spatial distribution of high-hazard areas observed on hazard-zonation maps, as depicted in Figure IO. With respect to the East Fork Lewis basin, which we believe was mapped fairly carefully for the purposes of this study, some amount of model over-prediction (i.e., 16% for SMORPH and 43% for SHALSTAB) might be true Type II errors. That is, the models likely do over-predict observed spatial patterns of slope-stability potential, as can be discerned from observed spatial patterns of existing and potential landslides. Particularly in the case of SHALSTAB, however, some portion of this over-prediction might be an artifact of the manner in which hazard-potential criteria were derived (i.e., the Q_c - slope stability classes assigned by precipitation rules to be included in the high-hazard management designation), as discussed previously with regard to Figure 8.

To evaluate the potential for model use in a management context, we developed a ranking scheme to quantify model performance and a number of other comparative criteria (see report section 5.0). We employed a statistical method for ranking models in 'terms of their ability to correctly and incorrectly identify known, existing shallow landslides. A numeric value was assigned to each of the possible database-intersect outcomes:

Type of database intersect&	Assigned value (p)
Landslide overlaps with DEM cell coded by model as high hazard	0
Landslide overlaps with DEM cell coded by model as moderate hazard	1
Landslide overlaps with DEM cell coded by model as low hazard	2

For example, an existing landslide was considered to be identified by a particular model if any superimposed DEM grid cell was coded “high hazard” (p = 0) or “moderate hazard” (p = 1).

The assigned values for all correctly and incorrectly identified landslides in each of the eight test basins were added to yield a cumulative score for each model, which then was normalized by the total number of landslides in each basin, Where landslides occurred in areas for which the soils survey data were missing, the SOILS screen grid cells were given a score of p = 2. These normalized scores then were added to a score sheet including results of other tested criteria, as will be described in report section 5.0.

Table 12 shows the results of this ranked test. SHALSTAB gained approximately twice as many points as SMORPH, reflected in the normalized cumulative scores (i.e., 1.9 versus 0.8, respectively). The SOILS screen received a significantly higher score (i.e., 6.7) than the other two models, due in part to the partial or total absence of soils-survey data in most test basins. SHALSTAB received a greater cumulative score than SMORPH, largely due to more frequent intersections of identified landslide polygons with model-predicted low and moderate hazards (Figure 11). Some of the discrepancy theoretically could be attributed to our assignment of management criteria via the precipitation rules, as described with respect to Figure 8.

At the outset of this study, we posed the following questions with regard to model performance: (1) How do model predictions of shallow landsliding compare with existing landslide inventories and hazard-zonation maps?; and, (2) How do model predictions compare with respect to each other? In summary, test statistics imply that the SMORPH and SHALSTAB models predict fairly well the spatial distribution of known existing landslides in the eight test basins (i.e., error frequency of 3% and 8%, respectively). These models, in general, also compare favorably with maps of shallow-landslide potential produced via watershed analyses (i.e., 6% and 5% Type I errors, respectively; and 3% and 7% Type II errors, respectively). The SOILS screen performed least well, missing 32% of the known existing landslides (i.e., Type I errors) and providing an incomplete cover of a substantial percentage of western Washington terrain (e.g., full data coverage existed in only two of the eight test basins). Test statistics also indicated that the mean differences in predictive model capability between the SOILS screen and either model were statistically significant, whereas the mean differences between SMORPH and SHALSTAB were marginally significant statistically. Hence, we conclude that the SOILS screen is comparatively less accurate and certainly less complete than the two tested models. While the average differences in predictive capability of SMORPH and SHALSTAB were not great, the former model tended to produce slightly fewer Type I and II errors. Contingent on the appropriateness of the precipitation-rule algorithm used to calibrate the SHALSTAB model, we conclude that SMORPH is slightly more accurate than SHALSTAB in predicting existing and potential landslides as represented in our updated landslide-inventory and hazard-zonation-map databases.

4.2 Method limitations

The purpose of this study component was to evaluate the potential constraints placed on management use of each tested model, by: (1) the nature of the key assumptions used to create the model; (2) the type and amount of data required as model input; and, (3) model sensitivity to changes in input parameters or variables. Limitations on model applications resulting from changes in terrain characteristics (i.e., geographic limitations) are discussed in paper section 4.3.

As discussed previously, the principal assumption made by all three models is that topographic controls dominate the spatial distribution of shallow landsliding. The relatively small errors incurred by SHALSTAB and SMORPH in predicting known shallow-landslide occurrences (i.e., 8% and 3%, respectively; Table 8) and potential unstable slopes (i.e., <10% and <6%, respectively; Tables 10 and 11) suggest that this assumption is quite reasonable, because both models on average reproduce fairly faithfully the spatial distribution of unstable slopes as specified in field-derived inventories and hazard-potential maps. Furthermore, the slightly stronger performance of the SMORPH model, in terms of predictive capability, implies that topographic controls are a dominant factor in promoting shallow failures and that inclusion in the model algorithms of other key influencing factors (e.g., soil properties, hydrology, vegetation) might not improve model performance, at least with regard to predicting the spatial distribution of shallow landslides in western Washington and similar terrains with maritime climates.

The relatively simplistic SMORPH model offers some advantages in a management context because it yields results that are comparable to the more sophisticated SHALSTAB model, without having to calibrate input variables (e.g., soil and hydrology properties) with off-site data or needing to collect additional data to run the model. In addition, a simplistic model with fewer data-input requirements contains less potential for Type I and II model errors associated with inaccurate characterizations of the spatial and temporal distributions of input variables. The SMORPH model, on the other hand, might lose substantial predictive capability in terrain where the topographic factors of hillslope gradient and curvature serve less well as proxies for the other key influencing parameters. Models like SHALSTAB, which can be expanded and refined to include algorithms addressing spatial and/or temporal variabilities of soil and hydrologic Factors, might be more appropriate in situations where landslide processes are not governed primarily by topographic forcing of soil- and water- mass fluxes. In its present form, however, the SHALSTAB model uses constant values for soil and hydrologic variables as placeholders for as-yet-undeveloped algorithms that would address problems of spatial and temporal variability. Hence, we conclude that the SHALSTAB model needs to be developed further before testing the hypothesis that factors other than topography might shape the spatial distribution of landslides.

For the test basins in which the soils-data coverage was complete (i.e., Sol Duc, Morton, and Chehalir Headwaters), the SOILS screen incurred the largest error in predicting known landslides (30%; see Table 8). This result suggests that basing shallow-landslide prediction on hillslope gradient and soil stability ratings generated by state soil surveys is less accurate and effective. Moreover, we suspect that the use in the SOILS algorithm of gradient, rather than gradient and curvature, contributes primarily to the greater inaccuracy of this method. The assumption of gradient and curvature as the primary landslide-forcing factors in western Washington is supported by the demonstrably better predictive capabilities of SMORPH and SHALSTAB. Hence, we believe that the SOILS method would be improved substantially by incorporating topographic curvature in the computational algorithm.

Using SMORPH or SHALSTAB in a management context also is affected by the accessibility of data required as input to run the models. The SOILS screen cannot be adjusted to calibrate output with new or more accurate data, without recreating the GIS layer. Table 13 lists: (1) the required input variables; (2) their default values as set internally in the models; (3) the typical sources of data available to the user in modifying default values without additional fieldwork or analysis; and, (4) the relative ease in collecting data when values in the literature . are inappropriate for the watershed of interest or watershed-specific data are nonexistent. For the purposes of comparison, each model was assigned a score reflecting the number of variables and the relative ease in collecting data from a given watershed to adjust the default values assigned to each variable. These scores are used in report section 5.0 to assess, in part, the management applicability of each model. Although the SOILS screen received zero points in this scheme, the score was adjusted later to reflect the relative drawback in using a method that cannot be adjusted to accommodate more accurate information on site or watershed physical variables.

As described in report section 2.2, the SMORPH model requires that slope-stability classes be set on the basis of mapped landslide densities (e.g., a high hazard rating corresponds to slope units in which the greatest landslide number have been measured per unit basin area), which can be ascertained from landslide inventories. Hence, where landslide inventories and/or hazard-zonation maps exist, the model can be calibrated without additional analysis or field work. In addition, management criteria (i.e., low, moderate, and high hazard-potential ratings) are known a *priori* because they are specified in, or can be derived, from watershed-analysis products. The greatest utility of this model lies in extrapolating from watersheds in which inventories have been compiled to areas with similar physical characteristics and no existing landslide databases. The limitations of this method are that it depends on the quality of the landslide database and the appropriateness of data extrapolation to basins where little physical data exist for verifying model predictions of potential landslide densities, Also, as mentioned previously, the simplicity of the model can be a detriment where topographic controls are sub- or co-dominant to other hillslope processes.

As can be seen in Table 13. the SHALSTAB model would require the greatest amount of literature and/or field analysis should the user decide to use values for input variables other than the defaults set internally in the model. Employing values from the literature can be problematic, given that little data exist on key soil input variables in western Washington watersheds. Montgomery et al. (1998), for example, use measured values obtained from a small catchment in coastal Oregon, in which the local geology and precipitation regime are not representative of all of western Washington. This model is limited, as in SMORPH, by the appropriateness of data extrapolation (e.g., from coastal Oregon to western Washington) and the quality of landslide inventories if model calibrations are performed using inventory data. Montgomery and Dietrich (1994) currently do not provide algorithms for addressing spatial and temporal variability in input parameters, nor are there standard methods for designing field sampling strategies and determining a representative value for an input variable if field measurements yield a wide range of values. It is possible computationally to run the model for discrete portions of a watershed which contain relatively homogeneous parent materials. Such an approach, however, might be prohibitively expensive or labor-intensive for landscape or regional management applications. Consequently, published uses of SHALSTAB to date (e.g., Montgomery et al., 1998) have employed the default values specified in Table 13.

An additional limitation of the SHALSTAB model in the management arena, as alluded to by the model authors (Montgomery and Dietrich, 1994), is the current lack of a formula for converting model output (i.e., critical rainfall (mm/day) necessary to initiate shallow landsliding) to management criteria (i.e., low, moderate, and high “hazard” potential). As described in report section 3.3, we chose an approach that utilized existing data and similar units of measure, This method also circumvented needs for additional fieldwork or rnanipulations of landslide inventories to back-calculate appropriate values for input variables, the latter of which appears to require some field effort as well. Our approach, however, might need to be replaced or refined as others begin to work on the problem and find more robust solutions.

As a measure of model sensitivity to input-parameter variability, we ran the SMORPH and SHALSTAB programs for a range of input values. This test did not include the SOILS screen because of its non-adjustability. Table 14 shows the results of modifying the threshold gradient classes in the SMORPH model (Table 14A), and using the range of default values for effective cohesion and phi angles given in the SHALSTAB model (Table 14B). These tables were compiled using the methods employed in Table 12, in which database intersections (i.e., landslide polygons from the inventory database and model predictions of slope stability for each DEM grid cell) were assigned a value depending on their agreement (p = 0 for a high-hazard DEM cell overlying a landslide polygon and p = 1 for a moderate-hazard cell overlying a landslide polygon) or disagreement (p = 2 for no match). As described for Table 12. the cumulative score for each test basin was normalized by the number of existing landslides, and the normalized scores for all eight basins were added to yield a total score for each model. The higher the score for each incremental increase in the magnitude of an input variable, the greater the number of known existing landslides incorrectly identified by the model (i.e., Type I errors). This technique provided a quantitative means for evaluating model predictions of slope-stability potential with changing values of the input variables.

For each model, the values of the input variables were adjusted between those calibrated to yield model predictions most closely resembling the landslide inventory and minimum values at which the hillslopes were predicted to be entirely stable (i.e., no potential landslides). For the SMORPH model, this involved increasing the threshold gradients in each of the low, moderate, and high landslide-hazard potential categories until the model predicted that all watershed slopes would be fully stable. This was accomplished by shifting the slope hazard-potential classes calibrated from the landslide inventory along the horizontal plane of the slope matrix (Table 15). For the SHALSTAB model, the effective cohesion was increased' from c' = 0 kN/m^2 to c' = 8 kN/m^2, at which point all watershed slopes were predicted to be entirely stable.

Tables 14A (demonstrates for the SMORPH model that, as the gradient thresholds increase for each of the hazard-potential categories (i.e., low, moderate, high), the frequency of Type I model errors increases correspondingly (i.e., database intersections assigned two points in the ranking scheme). The percent change (A%) in assigned points between applying the calibrated slope matrix (i.e., Step 0) and adjusting the matrix so that all slopes are predicted to be stable (i.e., Step ∞) is A% = 0.04, when averaged over all eight basins. Likewise, Type I errors produced by the SHALSTAB model occur more frequently with increasing magnitudes of effective cohesion (Table 148). For the SHALSTAB model, the percent change averaged over eight test basins is Δ% = 0.09, when comparing model default options c' = 2 kN/m^2 and c' = 8 kN/m^2, where ϕ = 33" is held constant. The results for the default option of c' = 15 kN/m^2 and ϕ = 33" are not shown, given that: all watershed slopes were predicted to be fully stable at effective cohesions of c' > 5 kN/m^2.

The percent change with increasing values of the input variables for each model was compared graphically by scaling the y-axis of a SMORPH plot of gradient-threshold class boundaries (i.e., Steps 24, 47, 70, and 93) versus cumulative percent change, by the y-axis of a SHALSTAB plot of effective cohesions (i.e., c' = 2. 5, 8 kN/m^2) versus cumulative percent change (Figure 12), given regular increments of increasing gradient and cohesion along the respective x-axes. 'This permitted a visual comparison of the relative sensitivity of each model to changes in the magnitudes of input variables, as reflected in the incremental increases in the number of points assigned to correct (p = 0,1) and incorrect (p = 2) grid-ceil intersections. Figure 12 shows that the SHALSTAB model is somewhat more sensitive to increases in the value of effective cohesion than is SMORPH to increases in the gradient threshold at which grid cells are predicted to have a low hazard potential (i.e., Type I model errors). That is, for an incremental increase of c' = 3 kN/m^2 in the SHALSTAB model and gradient S = 23% in the SMORPH model, the former predicts a relatively greater percent change in the number of Type I errors than does the latter. Hence, we conclude that SHALSTAB is relatively more sensitive to changes in input variables than is SMORPH, although both models can produce erroneous results with inappropriately chosen values of the input variables.

Figure 13 shows, for the Morton test basin, the Q_c classes versus the cumulative percent area predicted by SHALSTAB to be unstable, for the a range of default effective-cohesion values (i.e., c' = 2. 5. 8 kN/m^2). The curve represented by star symbols corresponds to the default input values of c' = 0 kN/m^2 and ϕ = 45". This figure also depicts the significant variation in the number of predicted landslides with increasing effective cohesion. As summarized by Montgomery et al. (1998), existing literature regarding the influence of root strength on soil mobility suggests that c' = 2 kN/m^2 is appropriate for clearcut slopes with decaying tree stumps and c' = 8 kN/m^2 is more representative of mature, hardwood-dominated forests or younger conifer stands. We found from model tests in the Morton watershed, for example, that effective cohesions of c' $\geq$ 8 kN/m^2 led to model predictions of fully stable slopes for any critical rainfall of Q_c < 400 mm/day (16 in/day), which is twice the magnitude of a 100-year, 24-hour storm event.

This rainfall amount is greater than the probable maximum precipitation computed for the Morton area (N. Wolff, WDNR. pers. comm.), which suggests that cohesions of c' $\geq$ 8 kN/m^2, presumably characteristic of forested conditions, yield unrealistic model results when used as input values.

Hence, the value of c' for which the SHALSTAB model predicts roughly the same spatial distribution of existing landslides is that corresponding to a clearcut watershed. None of the eight tested watersheds is entirely clearcut. Using a value of c' more representative of partially or fully forested conditions, however, would have resulted in a significantly higher percentage of Type I model errors, including omission from model predictions of landslides known to have occurred in mature, previously unharvested stands (Le., in portions of the Middle Hoh test basin). This problem might be resolved by running the model for discrete forest-age-class units with c' chosen separately for each unit, We did not explore this possibility due to study time constraints.

The SHALSTAB model also appears to be quite sensitive to variation in the input value of the internal friction angle. We ran the model for the cases c' = 2 kN/m^2, ϕ = 33" and c' = 2kN/m^2, ϕ = 45". Increasing the phi angle by 12 degrees resulted in a decrease of 89% in the area predicted by the model to be highly unstable (e.g., Q_c classes 1 through 3 for the Upper East Fork Lewis basin).

Hence, we conclude that both SHALSTAB and SMORPH are relatively sensitive to the magnitudes of their respective input variables, and that SHALSTAB is measurably more sensitive than SMORPH. We suggest that SHALSTAB model users employ conservative estimates of ϕ and c', in the absence of reliable field measurements or proven methods for estimating appropriate values. Similar to Montgomery et al. (1998), we found that the combination of c' = 2 kN/m^2 and ϕ = 33" yielded predicted landslide spatial distributions most closely resembling measured landslide distributions in all watersheds tested by this study. In addition, SMORPH modelers should calibrate the slope matrix designations of landslide hazard for each gradient class using accurate landslide inventories, wherever possible, to reduce the potential for Type I model errors.

4.3 Geographic applicability

To investigate the ability of each model to correctly predict landslides in different western Washington terrains, we separated the eight test basins into six categories pertaining to major geomorphic provinces: continental glaciated terrain (Hazel), Cascades volcanic complex (Morton, Upper East Fork Lewis River), Northwestern Cascades system (Jordan-Boulder, North Fork Stilliguamish River), Olympic core rocks (Sol Duc River), Western Olympic Assemblage (Middle Hoh River), and Eocene volcaniclastics (Chehalis Headwaters) (see Figure 1). We rated each model in by the number of Type I errors it produced in each geomorphic province (Table 8, right-hand column).

As described in sections 4.1 and 4.2, the Soils screen performed least well overall because of the lack of soils-survey data in six of the eight test basins and the inability of the method to discriminate slope curvature (i.e., 32% Type I errors). Of the test basins with complete or nearly complete data, the screen yielded the greatest percent of Type I errors in the Eocene-volcaniclastics (32% of the test basins) and Cascades-volcanics provinces (48% of the test basins), both regions of which incorporate most of southwestern Washington. Likely reasons are that the method could not detect relatively steeper, convergent features (e.g., inner gorges) inside broader, gentle slope areas, particularly in areas of lower topographic relief like the Chehalis Headwaters basin in which a substantial fraction of the existing failures were found. As described in report section 4.1, the broad inclusion of slopes in soil hazard-potential polygons resulted in fewer Type I errors in some test basins with the SOILS screen than with SHALSTAB. although the results still indicated substantial predictive errors in certain terrains (i.e., the Hazel test basin).

The SHALSTAB model performed least well in the continental-glaciated terrain (e.g., 32% Type I errors in the Hazel test lbasin); Montgomery et al. (1998) also concluded in their test that the model predicted landslides least well in thick glacial deposits. The percent of Type I model errors increased by an average 69% over those computed for the other five provinces. The model also produced approximately 60% more Type I errors than the SMORPH model in the Olympic terrains, as discussed in report section 4.1.

The SMORPH model performed least well in the western Olympic test basin (7%) and, surprisingly, misidentified only 1% of the existing landslides in the continental-glaciated terrain. As discussed in report section 4.1, it appears that the Arc/Info™ GRID tool is capable of discerning variations in gradient and slope curvature on the order of one DEM grid cell, allowing the model to'detect $100m^2$ or larger shallow landslides superimposed on deep-seated failures. Thus, it appears that SMORPH might be more capable of identifying landslide features in glaciated terrain, although a larger sample of test basins would be required to properly evaluate this theory.

We conclude from this test that SHALSTAB and SMORPH could reasonably be employed in most western Washington terrains to predict shallow landslides. The SHALSTAB model appears to work least well in continental-glaciated terrain, while preliminary results suggest that SMORPH might perform substantially better than SHALSTAB in glaciated topography dominated by deep-seated failures. The SOILS screen runs a distant third in most terrains because of the incomplete nature of the GIS coverage and the relatively greater percent of Type I model errors.

4.4 Technical criteria

We asked five general questions with respect to technical aspects of each method: (1) How long does it take to run the model program?; (2) How much training is required?; (3) How much computer space is required by the model programs?; and, (4) How easy would it be for the user to modify the model? These questions would be important from a practical perspective and could influence how each model would be used in a management context.

Prior to running any GIS-model program, it is critical that someone familiar with computing systems and Arc/Info'" programming language verifies that the program is running correctly. The optimum method for double-checking program execution is to compare standard (i.e. default) model output with output obtained for the same geographic area from the model authors or from existing databases produced by the model on the WDNR-GIS system. None of the tested models has been refined sufficiently to document, internally or otherwise, all the known technical complexities of loading and running a program on a particular operating system, so it is important to test program execution.

The purpose of evaluating computer processing time was to provide users with an estimate of the average time necessary to create slope stability screens, particularly when working at a landscape or regional scale. The SMORPH model program runs about five times faster than that for SHALSTAB. On average, for 30-m. DEMs, the SMORPH program takes three minutes to run for a WAU (i.e., an area typically less than 200 km^2), while it takes 18 minutes to run the SHALSTAB program. Run time increases approximately three-fold when model programs are executed using 10-m. DEMs. If the user were to employ 10-m. DEMs in creating a slope stability screen of all western Washington WAUs, for example, the SMORPH program would require roughly 90 hours of computer time, while it would take well over 400 hours of computer time to process the SHALSTAB program. The SOILS screen exists already; therefore, computer use is limited to the time it takes to create a map.

A certain level of training is required to fully understand and use the model output, regardless of which model is being employed as a slope stability screen. Only very basic computer skills are necessary, however, to run model programs and create maps of the predicted slope-stability distributions, assuming that DEMs exist for the area of interest. The user should know how to obtain access (i.e., "log in"), navigate, execute basic tile commands, and run the model program on the computer system. However, a basic geotechnical understanding of landslide processes is necessary to calibrate the SHALSTAB and SMORPH models, and the SHALSTAB model additionally requires an ability to interpret and apply soil-property and hydrologic (e.g., precipitation) data.

Furthermore, the SOILS screen and SMORPH model give results of slope-stability analyses explicitly in terms of management criteria currently used in Washington (i.e., low, moderate, and high landslide potential), so that interpretation of output is straight-fotward if the user is familiar with their definitions. The default criteria used in SMORPH (Table 6) might need to be calibrated with landslide inventories from the basin of interest, or from an analogous watershed, and some training might be necessary in using the calibration algorithm. The current version of SHALSTAB provides no guidance for translating output to management criteria or for calibrating input variables to local area conditions. Consequently, more training and background knowledge are necessary for running the SHALSTAB program and interpreting model results.

Given that DEM data are the only absolute requirement for all three models, data input requirements can be relatively straightforward. SHALSTAB and SMORPH provide default values for soil and slope properties, respectively, allowing the user to run the computer programs without first having to calibrate the models. We strongly recommend, however, that input values be calibrated to achieve greater predictive accuracy.

High DEM resolution is lkey to producing reasonable results with SHALSTAB and SMORPH. The SOILS screen, in contrast, is unaffected by DEM resolution because it was derived from static data (i.e., fixed values for hillslope gradients and soil properties). DEMs with

10-m. resolution generate more accurate results than 30-m. data because they better represent the true topographic surface. Figure 14 compares the frequency of predicted landslides in the Jordan-Boulder, North Fork Stilliguamish, and Hazel test basins, using 10-m. versus 30-m. DEMs. Results of tests using the SMORPH and SHALSTAB programs indicate that, when 10-m. DEMs are used, both models predict relatively more failures in the unstable-slope classes (i.e., SMORPH slope-stability rating class 3 and SHALSTAB critical-rainfall classes 1, 2, and 3; see Figure 14). It should be noted, however, that this relatively greater number of landslides predicted using 10-m. DEMs is actually more representative of measured spatial landslide distributions in these basins. That is, employing 30-m. DEMs results in a higher percent of Type II model errors. On average, use of DEMs with 10-m. rather than 30-m. resolution leads to a 94% improvement in the predictive accuracy of the SMORPH model and 60% improvement in SHALSTAB results. Hence, it is recommended that 10-m. resolution data be used whenever possible.

The SMORPH model requires the least amount of storage space on a computer system. It produces grid data, which use less storage space than GIS coverages like the SOILS screen. The SHALSTAB model also generates grid data; however, it produces one grid for each of the default output options (i.e., one grid for each of the preset combinations of c' and ϕ). Additionally, SHALSTAB creates several other grids that typically are not used in a management context, although an experienced programmer can modify the code to circumvent creating these data layers. For geographic areas smaller than a typical WAU, data storage requirements do not pose problems for a computer with Arc/Info[tm] software, as a single grid or small set of grids does not take up much disk space. Data-storage problems are substantially greater for some systems (e.g., personal computers) when areas larger than the size of a typical WAU are used.

All tested models can be run on a personal computer with Arc/info™ software. The SHALSTAB model requires that the computer also have Fortran and C program executability for running subroutines that: (1) remove artificial topographic convergences occasionally created by the DEMs; and, (2) calculate upslope contributing areas to each grid cell for the hydrologic component of the model. Only the SOILS screen can be accessed on a personal computer with non-Arc/Info™ software and no additional programming. Both SHALSTAB and SMORPH would require additional programming to make them compatible with non-Arc/Info™ software.

User access to each of the three models would be improved by additional program or method documentation. SMORPH and SHALSTAB programs would benefit from more internal documentation, to assist future generations of programmers in adjusting the input variables. Some program documentation was developed, as part of this study, for both the SHALSTAB and SMORPH models. This on-line help consists of 'read.me' files (i.e., text files that assist with program executions) and internal documentation (i.e., comment lines within the program to assist the Arc/Info™ programmer in adjusting or calibrating the model). We also developed programs for viewing the model output and creating simple maps from the model data, and we created a menu-driven system for adjusting SMORPH slope criteria. A similar tool would enhance substantially the usability of the SHALSTAB program. A menu-driven system eliminates the need for a programmer to adjust the program input variables, and it serves. to remind the user that input variables generally need to be calibrated for the area of interest.

The SOILS screen, on the other hand, does not need internal documentation because it exists as a compiled cover, rather than an executable program. Metadata (i.e., data about the data) exist for the SOILS screen, but little documentation exists regarding the applicability of the WDNR-GIS SOILS layer to different management scenarios. The SOILS screen also lacks any accompanying written discussion of the rationale for assigning management criteria to certain. combinations of soil types and hillslope gradients. This information could assist the user in interpreting the accuracy of the slope-stability predictions, especially when the area of interest falls outside the specified soil-gradient class.

5.0 Discussion and Conclusions Regarding Model Applications

The primary purpose of this study was to evaluate in a management context the use of three currently available methods for predicting shallow landslides. In particular, our goal was to compare the current GIS slope-stability cover, used in Washington regulatory and management practices, with other, potentially more reliable, GIS-based models. Toward that end, we developed a rating scheme to measure the overall performance and applicability of the three tested methods with respect to the scientific and technical criteria discussed in this paper.

The rating scheme was formulated so that each model would be scored for each identified criterion based on either of the following: (1) statistical values summarized elsewhere in this paper; or, (2) assigned points representing qualitative answers to questions for which no quantitative measures could be found. The latter were expressed as "yes" (usually assigned zero points; p = 0) or "no" (p = 1) questions. The lowest cumulative score reflects the model that generates unstable-slope predictions most comparable with existing landslide databases (i.e., fewer Type I model errors) and would be the most readily applicable in a management context.

Table 16 shows the results of this rating exercise, and Table 17 lists, for each criterion, the rationale for the point assignment. The purpose of the numerical ranking is to describe relative performance; the magnitudes of the total scores, therefore, have no real significance for measuring how much better one model performs than the other. These results suggest that the

SMORPH model might offer more advantages in a management or regulatory context than the SOILS screen and the current version of SHALSTAB (i.e., with the values of soil-property and hydrologic input parameters held constant).

In general, the reasons for the relatively higher rating (Table 16) of the SMORPH model are:

(1) SMORPH generated spatial predictions of shallow landslides that most closely resembled the measured densities of known existing landslides (i.e., landslide inventory databases) and the field-derived maps of landslide hazard potential (Tables 8, 10, and 11). Specifically, the SMORPH model, on average, yielded fewer Type I and II model errors, even in continental-glaciated terrain;

(2) SMORPH contains fewer input variables than SHALSTAB; consequently, there is less potential for 'Type I and II model errors associated with using input values that are unrepresentative of the study area. In addition, the input variables in SMORPH (i.e., gradient and slope curvature) appeared in general to be less sensitive to variation than SHALSTAB input variables (i.e., effective cohesion and internal friction angles; see Figure 12). The predictive capability of the SOILS screen likely is limited by the absence of a slope-curvature parameter in the computational algorithm, and the GIS cover cannot be adjusted to reflect hillslope gradients and soil properties outside the specified general categories.

(3) The GIS cover generated with SMORPH uses management criteria (i.e., low, moderate, and high landslide-potential ratings) to signify classes of slope instability, whereas the SHALSTAB model outputs values, in terms of critical rainfall required to initiate landsliding, that require geomorphic interpretation to be applied in a management capacity. The SHALSTAB model currently does not provide a mechanism for converting

from critical rainfall units to management criteria. Hence, the SMORPH model is more readily applicable in the current management decision-making framework in Washington.

(4) SMORPH runs approximately 80% faster than SHALSTAB on a computer workstation, which might be important to managers with limited computer resources and large data requirements (i.e., for creating regional screens of slope stability). SMORPH also requires about five times less data-storage volume than SHALSTAB and several times less storage volume than the SOILS screen. And;

(5) Relatively less training is necessary to instruct users on executing SMORPH programs and interpreting model results. The SMORPH model also requires comparatively less assistance from technical specialists in calibrating input variables (i.e., adjusting the slope matrix with landslide-inventory data) and interpreting model results. The SHALSTAB model requires more data collection (e.g., to properly characterize soil properties and calibrate the model for the precipitation regime in the area of interest) and interpretation of model predictions, which are accomplished more easily by users with some background in geomorphology, geoengineering, soil science, and/or hydrology.

Hence, the SMORPH model might fill the near-term needs of resource managers and regulators for a ready-to-use model that can create a landscape or regional shallow-landslide screen.

The SMORPH model potentially offers some disadvantages as well. Together with the SOILS screen and the current version of SHALSTAB, this model could lose some predictive capability in terrain where topographic controls on shallow landslide initiation are secondary to other destabilizing factors (e.g., snow avalanching, slumping along earthflow margins, ground subsidence, erosion of glacial deposits). The model would have to be modified substantially to include algorithms for explicitly treating variables other than slope gradient and form.

Alternatively, the SHALSTAB model, using the precipitation-rule method for converting model output to management criteria, yields results that are fairly comparable with those of the SMORPH model and existing landslide inventories. This model potentially could be more versatile than SMORPH or the SOILS screen, because it contains placeholders for algorithms that would address explicitly the spatial and/or temporal variability of soil and hydrologic factors. In addition, future comparisons of the SMORPH model and a more sophisticated SHALSTAB model might resolve whether explicit treatments of soil and hydrologic properties (e.g., SHALSTAB) yield substantially better predictions of slope-stability potential than do more sirnple models in which topographic parameters serve as proxies for these key variables (e.g., SMORPH). Test statistics from this study suggest that the current version of the SHALSTAB model performs no better than SMORPH, even though it includes several key variables (i.e., soil transmissivity. depth, cohesion, bulk density, and internal friction angle), albeit expressed as constants. This result could be attributed to a number of factors, including the possibility that soil properties are of secondary importance compared with topographic factors, and that including them explicitly in predictive models is less critical than accurately simulating fine-scale variations in slope topography.

The SOILS screen is relatively more "user-friendly" than the other two models because it is delivered to the user as a pre-compiled GIS cover that requires no calibration and gives results in terms of management criteria that can be incorporated readily in the existing regulatory and forest-management decision-making processes. Nonetheless, it received a comparatively less favorable score than SMORPH and SHALSTAB because it yielded significantly more Type I errors (i.e., incorrectly identified landslides). In addition, the SOILS screen contains large gaps in geographic coverage because digital soils-survey data are lacking on some portions of Washington, especially on federal lands. Furthermore, the digital soils layer maintained by the federal government (e.g., USDA Forest Service) can be incomplete, as was encountered in this study.

Incompleteness aside, the relative inaccuracy of the SOILS screen puts it at a disadvantage when compared with the more accurate SMORPH and SHALSTAB models. Study results imply that the SOILS GIS cover, maintained by the state for management and regulatory applications, should be replaced by one created with either predictive model. Given that the SHALSTAB and SMORPH models have been developed and tested in maritime climates of the Pacific Northwest, they should be similarly analyzed for precipitation regimes and terrains more typical of the continental interior, prior to their use east of the Cascades Range or elsewhere.

A number of interesting questions have arisen during this study regarding the technical merits of each GIS-based model, as well as the quality and applicability of landslide inventories and other databases used to calibrate the models. These include such issues as the relative need for including spatial variability of soil properties as elements of GIS-based models designed to be used in terrain where topographic controls dominate the spatial distribution of shallow landslides. Given that the SHALSTAB and SMORPH models, as currently configured, do not explicitly treat the stochastic nature of key variables, yet they predict relatively well the known distribution of landslide potential, attests to the real possibility that it might not be necessary to include spatial and ternporal variability in the model frameworks. Furthermore, the relative agreement between SMORPH model predictions and observed landslides suggests that including soil properties in the model equation might not even be necessary for producing a reliable, preliminary landslide-screening tool designed for management applications. The same argument could be made for SHALSTAB, since soilvariables are held constant and the model' in essence functions like SMORPH in discriminating landslide potential on the basis of topographic factors.

Our study also motivates the need for continued discussion of appropriate ways to parameterize model predictions of landslide potential in terms of management decision-making criteria. We have identified a number of alternatives for converting model predictions of landslide potential to decision criteria. All of them, however, rely on the current management formulation of what constitutes "hazard" and "risk", whereby hillslope processes are treated deterministically (e.g., the analysis of "hazard" does not necessarily take into account the history of landslide processes predating recent management activities). It may be that GIS-based topographic rnodels more accurately reflect the full spatial and temporal distribution of potential unstable slopes than do landslide databases generated during watershed analyses, because the former are measuring landslide potential based on landform characteristics that largely existed prior to 20th. century land management, while the latter are based heavily on aerial-photo interpretation and, hence, provide only a contemporary measure of landslide rates. GIS-based models, therefore, could be useful in helping to redefine the way in which hazard-zonation maps typically are generated.

Acknowledgments

This study was funded by the Washington Department of Natural Resources, the Washington Forest Protection Association, and the Cooperative Research and Effectiveness Monitoring Committee (CMER) of the Washington Timber/Fish/Wildlife program. We thank Weyerhaeuser Company, the Tulalip Tribe, Murray-Pacific Corporation, and WDNR for providing data and local knowledge of test watersheds. Landslide and soils data for the Lewis watershed were obtained, in part, from the Web site maintained by the U.S.D.A. Forest Service, Gifford Pinchot National Forest. We thank David Montgomery, Tien Wu, and David Johnson for valuable input and assistance in compiling and running their models. We also thank Matthew Brunengo, Kate Sullivan, Daniel Miller, Tien Wu, George Pess, and Paul Kennard for valuable input in the early phases of this project. In addition, we thank Harvey Greenberg for assistance in making the SHALSTAB model operable on the WDNR GIS system; Daniel Miller for sharing his model and many critical discussions; Andria Villines and Elizabeth Freeman for assistance with statistical analyses; and Venice Goetz for her field assistance in verifying landslide data.

References

Benda, L.E. and T.W. Cundy, 1990, Predicting deposition of debris flows in mountain channels. Canadian Geotechnical Journal 23: 409-417.

Coho, C., 1997, Mass-Wasting Assessment. *IN* Jordan-Boulder Watershed Analysis, Appendix A, Wash. Dept. Natural Resources, Olympia, WA. Appendix with maps.

Dietrich, W.E. and T. Dunne, 1978, Sediment budget for a small catchment in mountainous terrain. Zietschrift fur geomorphologie, suppl. bd. 29: 191-206.

Dragovich, J.D. and M.J. Brunengo, 1995, Landslide map and inventory, Tilton River - Mineral Creek area, Lewis County, Washington. Wash. Dept. Natural Resources, Div. of Geology and Earth Resources, Open-File Report 95-1, 165 pp.

Dietrich, WE, Wilson, C.J., and S.L. Reneau, 1986, Hollows, colluvium, and landslides in soil-mantled landscapes. IN Hillslope Processes, A.D. Abrahams (ed.), Allen and Unwin. Winchester, MA., p. 361-388.

Dietrich, W.E., Wilson, C.J., Montgomery, D.R., and J. McKean, 1993, Analysis of erosion thresholds, channel networks, and landscape morphology using a digital terrain model. Journal of Geology 101: 259-278.

Environmental Systems Research Institute, 1992, Arc/Info Users Guide. Chap. 6.1, Grid command references, 1 vol.

Gerstel, W.J., 1996, The upside of the landslides of February 1996 -Validating a stability analysis of the Capitol Bluffs, Olympia, Washington. Washington Geology 24(3): 3-16.

Hammond, C., Hall, D., Miller, S., and Swetik, P. 1992, Level 1 stability analysis (LISA). Documentation for version 2.0. USDA Forest Service, Gen. Techn. Report INT-285, 190 pp

Hoh Tribe and Washington Department of Natural Resources (WDNR), 1993, Forest agreement related to the Hoh River, Kalaloch Creek, and Nolan Creek drainages: Memorandum of understanding, and supporting documents. Wash. Dept. Natural Resources, Forest Resources Div., Olympia, WA.

MathSoft, Inc., 1998, S-Plus 5: Guide to Statistics. Chap. 3, Statistical inference for one and two sample problems, MathSoft Inc., Data Analysis Products Div., 702 pp.

Miller, J.F. et al., 1973. Precipitation-Frequency Atlas of the Western United States. U.S. Department of Commerce, National Oceanic and Atmospheric Admin. (NOAA), vol. IX.

Montgomery, D.R. and W.E. Dietrich, 1994, A physically based model for the topographic control on shallow landsliding. Water Resources Research 30(4): 1153-1171.

Montgomery, D.R, Dietrich, W.E., Torres, R., Anderson, S.P., Heffner, J.T., and Loague, K., 1997, Hydrologic response of a steep unchanneled·valley to natural and applied rainfall. Water Resources Research 33, 215230.

Montgomery, D.R., Sullivan, K., and H.M. Greenberg, 1998, Regional test of a model for shallow landsliding. Hydrological Processes 12: 943-955.

Murray Pacific Corporation, 1998, West Fork Tilton Watershed Analysis. Murray Pacific Corporation, Tacoma, WA, 1 volume with appendices.

Murray Pacific Corporation, 1998, Nineteen Creek Watershed Analysis. Murray Pacific Corporation, Tacoma, WA, 1 volume with appendices.

O'Connor, M. and T.W. Cundy, 1993, North Fork Calawah River watershed condition survey: Landslide inventory and geomorphic analysis of mainstem alluvial system. Part I: Landslide inventory and geomorphic analysis of mass erosion, Report prepared for USDA Forest Service, Olympic National Forest, Olympia, WA., 17 pp.

O'Loughlin, E.M.. 1986, Prediction of surface saturation zones in natural catchments by topographic analysis. Water Resources Research 22(5): 794-804.

Pack, R.T., Tarboton, D.G., and C.N. Goodwin, 1998, Terrain stability mapping with SINMAP. Technical description and users guide for version 1 .00. Terratech Consulting Ltd., Report No. 4114-0, Salmon Arm, B.C., Canada.

Perkins, S. And B.D. Collins, 1997, Landslide and channel response inventory for the Stilliguamish watershed, Snohomish and Skagit Counties, Washington. Report prepared for the Stilliguamish Tribe of Indians, Washington Dept. of Ecology, Snohomish County Dept. of Public Works, and the Tulalip Tribes Nat. Res. Dept. Tulalip Tribes, Natural Resources Dept., Marysville, WA., August 1997, 23 pp. with appendices.

Reddi, L.N. and T.H. Wu, 1991, Probabilistic analysis of ground-water levels in hillside slopes. Journal of Geotechnical Engineering 117(6): 872-890.

Schuster, J.E.. compiler, 1992, Geologic Map of Washington. Wash. Dept. Natural Resources, Div. Geology and Earth Resources, Olympia, WA.

Shaw, S.C. and D.H. Johnson, 1995, Slope morphology model derived from digital elevation data. Proceedings, Northwest Arc/Info Users Conference, Coeur d'Alene, ID., Oct. 23-25, 1995, 12 pp.

Sidle, R.C., Pearce, A.J., and C.L. O'Loughlin, 1985, Hillslope Stability and Land Use. Chapter 4, Natural factors affecting slope stability, pp. 31-72. Water Resources Monograph Series 11, American Geophysical Union, Washington, DC., 140 pp.

Swanson, F.J. and R.L. Fredricksen, 1982, Sediment routing and budgets: Implications for judging impacts of forest practices. IN Sediment Budgets and Routing in Forested Drainage

Basins, USDA Forest Service, Gen. Techn. Report PNW-141, Portland, OR., p. 129-137

Swanson, F.J., Swanson, M.M., and C. Woods, 1981. Analysis of debris-avalanche erosion in steep forested lands: An example from Mapleton, Oregon, USA. Intern? Assoc. Hydrol. Sciences, AISH Publ. 132: 67-75.

Thorsen, G.W., 1978, Landslide provinces in Washington, Engineering Geology in Washington, vol. 1, Wash. Dept. Natural Resources, Div. of Geology and Earth Res., Bull. 78, 90 pp.

Timber/Fish/Wildlife Agreement, 1987, The Timber Fish Wildlife Agreement: Introduction

Washington Department of Natural Resources, Forest Practices Division, Olympia, WA., 16 pp.

United States Dept. of Agriculture (USDA) Forest Service, 1997, Upper East Fork Lewis River Watershed Analysis. U.S. Forest Service, Gifford Pinchot Nat'l. Forest, Vancouver, WA., 1 volume with appenclices.

United States Dept. of Interior (USDI) Fish and Wildlife Service, U.S. Dept. Commerce National Oceanic and Atmospheric Admin., U.S. Environmental Protection Agency, Office of the Governor of the State of Washington, Wash. Dept. Natural Resources, Wash. Dept. Fish and Wildlife, Wash. Dept. Ecology, Washington State Association of Counties, Washington Forest Protection Association, Washington Farm Forestry Association, Colville Confederated Tribes, and other tribal governments, 1999, Forests and Fish Report. Report prepared for the Washington Forest Practices Board and the Governor's Salmon Recovery Office, Feb. 22, 1999. Wash. Dept. Natural Resources, 218 pp.

Walpole, R.E. 1974, Introduction to Statistics, 2nd ed., MacMillan Publishing Co, Inc. New York, 340 p.

Washington Department of Natural Resources, 1984, State Soil Survey. Wash. Dept. Natural Resources, Forest Land Mgmt. Div., Olympia, WA, 7 volumes.

Washington Department of Natural Resources, 1988, Criteria for proposed reclassification of slope stability ratings. Agency documentation supporting digitized slope-stability map maintained in the public domain by the Wash. Dept. Natural Resources, Info. Technology Div., Olympia, WA.

Washington Department of Watural Resources and US Forest Service, 1996, Sol Duc Watershed Analysis. Wash. Dept. Nat. Res., Olympia, WA, 1 volume with appendices

Washington Department of Watural Resources, 1997, Final Habitat Conservation Plan Washington Forest Practices Board, 1 volume with appendices.

Washington Department of Natural Resources, 1997, Jordan-Boulder Watershed Analysis Wash. Dept. Nat. Res., Olympia, WA, 1 volume with appendices.

Washington Department of Natural Resources, 1998, Hazel Watershed Analysis. Wash. Dept.

Nat. Res.. Olympia, WA, 1 volume with appendices:

Washington Department of Natural Resources, in preparation, Middle Hoh Watershed Analysis. Wash. Dept. Nat. Res., Olympia, WA, 1 volume with appendices,

Washington Forest Practices Board (WFPB), 1995, Forest Practices Act RCW 76.09, Title 222 WAC Rules and Board Manual. Washington Department of Natural Resources, Forest Practices Division, Olympia, WA. July 1995, 1 volume with appendices.

Washington Forest Practices Board (WFPB), 1997, Board Manual: Standard Methodology for Conducting Watershed Analysis Under Chapter 222-22 WAC. Washington Department of Natural Resources, Forest Practices Division, Olympia, WA. Version 4.0, Nov. 11997, 1 volume with appendices.

Weyerhaeuser Corporation, 1994, Chehalis Headwaters Watershed Analysis, Weyerhaeuser Corp., Federal Way, WA, 1 volume with appendices.

Wu, T.H. and M.A. Abdel-Latif, 1995, Prediction, mapping, and updating of landslide hazard. *IN* A.S. Balasubramanian et al. (eds.), Developments in Deep Foundations and Ground Improvement Schemes, Symposium Proceedings; Balkema, p. 155-159.

Wu, T.H. and M.A. Abdel-Latif, 1997, Prediction and mapping of landslide hazard. Final report, T/F/W CMER Publication, TFW-SH15-97-001, 46 pp.

Wu, T.H., Merry, C.J., Abdel-Latif, M., and C.B. Benosky, 1993, Landslide hazard mapping - 1. Estimating piezometric levels. Final report, T/F/W CMER Publication, TFW-SH-15-93-001, 100 pp.

Wu, W. and R.C. Sidle, 1995, A distributed slope stability model for steep forested basins. Water Resources Research 31(8): 2097-2110.

Zevenbergen, L.W. and C.R. Thorne, 1987. Quantitative analysis of land surface topography. Earth Surface Processes and Landforms 12: 47-56.

Table 1. Criteria for determining slope stability from the SOILS data

	Soils criteria for slope stability ratings	Mass Wasting Potential
Very Unstable	map units with slopes greater than 30%	
	map units with slopes up to 30%	high
Unstable	map units with slopes greater than 65%	very high
	map units with slopes up to 30%	medium
	map units with slopes from 30-65%	high
Stable	map units with slopes up to 30%, where the soil phase is rated as unstable	medium
	map units with slopes up to 30%	medium
	~~map units with slopes up to 30%. where the soil phase at 30.65% is also rated stable~~	low

Table 2. Matrix relating slope curvature and gradient to shallow landslide potential, as used in the SMORPH model. The number and distribution of slope gradient classes (i.e., A • E) are set for a specific geomorphic unit with the aid of landslide inventories or slope stability analyses.

Slope curvature	**Slope gradient (percent)**				
	A	**B**	**C**	**D**	**E**
Convex	low	low	low	low	moderate
Planar	low	low	low	moderate	high
Concave	low	moderate	high	high	high

Table 3. Critical rainfall classes (Q,) designated by the SHALSTAB model.

Q_c class	**Rainfall amount needed to induce failure**	**Q_c class**	Rainfall amount needed to induce failure
1	Unconditionally unstable at this cohesion	5	ZOO-100 millimeters par day
		6	greater than 400 millimeters per day
2	0-50 millimeters per day	7	Unconditionally stable
3	50-100 millimeters per day	8	Stable at this cohesion
4	100-200 millimeters per day		

Table 4. Physical and geologic characteristics of test basins

Test Basin	Physiographic Area	Geologic Province	Area (km* and acres)	Topographic Relief (m)	Number of Known. Landglides
Jordan-Boulder	North Cascades Range	Northwest Cascades Metamorphic Suite; includes meta-quartz diorite, low-grade schists and phyllites, and plutonics	133 km^2 32,987 ac.	1941	155
North Fork Stillaguamish River	North Cascades Range	Low-grade metamorphosed sediments, including phyllite and greenschist	130 km^2 32,144 ac.	1504	215
Haze!	western Rank of Cascades Range - Puget Lowlands	Continental glacial deposits overlying low-grade metamorphosed sediments	98 km^2 24,209 ac	1528	117
Sol Duc River	northern Olympic Peninsula	Crescent Basalt and Olympic Lithic Assemblage (metamorphosed marine sediments)	1 a5 km^2 45,674 ac.	915	101
Middle Hoh River	western Olympic Peninsula	Western Olympic Assemblage; extensively sheared and metamorphosed marine sediments	331 km^2 81 ,679 ac.	1575	733
Morton	Central Cascades Range	Eocene to Recent andesitic volcanics	88 km^2 21.686 ac	1127	980
Chehalis Headwaters	Coast Range (Willapa Hills)	Eocene to Miocene mafic volcanic assemblage	1 a2 km^2 45,000 ac.	818	134
Upper East Fork Lewis River	Central Cascades Range	Eocene to Recent andesitic volcanics with igneous intrusions	81 km^2 20,016 ac.	1022	a 9

[1] **Includes identified shallow and deep-seated landslides.**

Table 5. DEM resolution and sources of data for the eight test basins.

Test Basin	Source of Landslide inventory Data	Hazard-Zonation Map Available	DEM Resolution	Percent Basin with Soils Layer
Jordan-Boulder	WDNR, 1997	Yes	10m	63%
North Fork Stillaguamish River	Perkins and Collins (1997); inventories created for this study	No	10m	22%
Hazel	WDNR, 1998	Yes	10m	65%
Sol Duc River (4 WAUs)	WDNR and USDA Forest Service (1996)	Yes	30m	95%
Middle Hoh River	WDNR (in preparation)	No (not yet digitized)	30m	64%
Morton (Portions of 2 WAUs)	Murray Pacific Timber Corp. (1998)	No (not available in digital format)	10m	100%
Chehalis Headwaters	Weyerhaeuser co. (1994); updated for this study	No (errors in digital database)	10m	100%
Upper East Fork Lewis River	USDA Forest Service (1997) and inventories created for this study	Yes	10m	2 %

Table 6. Gradient threshold values (in percent) calculated from landslide databases for input to the SMORPH slope matrix (Table 2) for each test basin. See text for discussion.

Test Basins	Gradient threshold corresponding to "hazard" designations for each curvature class				
	Low for convex and planar, moderate for concave	Low for convex and planar, high for concave	Low for convex, moderate for planar, high for concave	Moderate for convex, high for planar, concave	High for all slope forms
Jordan-Boulder	15	45	50	70	∞
N.F. Stillaguamish River	15	40	47	70	∞
Hazel	15	24	47	70	∞
Sol Duc River	15	24	47	70	∞
Middle Hoh River	15	24	47	70	∞
Morton	25	55	65	70	∞
Chehalis Headwaters	15	65	70	80	∞
Jpper E.F. Lewis River	40	50	60	70	∞

Table 7. Precipitation "rules" used to create management criteria for the SHALSTAB model. See text for discussion.

Test Basin	Management Criteria			Area-Weighted Mean Precipitation	Area-Weighted Maximum Precipitation
	Low "Hazard"	Moderate "Hazard"	High "Hazard"		
Jordan-Boulder	6,7,8	5	1,2,3,4	108	127
Upper North Fork Stillaguamish	5,6,7,8	4	1,2,3	83	102
Hazel	5,6,7,8	4	1,2,3	80	102
Sol Duc	6,7,8	5	1,2,3,4	129	152
Middle Hoh	6,7,8	5	1,2,3,4	185	229
Morton	5,6,7,8	4	1,2,3	100	114
Chehalis Headwaters	6,7,8	5	1,2,3,4	116	140
E.F. Lewis	6,7,8	5	1,2,3,4	123	140

Table 8. Predictions of known, existing shallow landslides using the three models (SOILS screen, SMORPH, and SHALSTAB), given as the number of incorrectly identified landslides (no. missed) per total number of landslides in each basin (see text).

Test Basin	Number of Identified Land-slides (T)	SOILS		SMORPH		SHALSTAB (ϕ = 33°, C' = 2kN/m^2)	
		no. missed (N)	N/T	no. missed (N)	N/T	no. missed (N)	(N/T)
Jordan-Boulder	155	40	0.26	0	0.00	5	0.03
North Fork Stillaguamish River	215	202	0.94	1	0.00	20	0.09
Hazel	117	34	0.29	1	0.01	37	0.32
Sol Duc River	101	6	0.06	1	0.01	12	0.12
Middle Hoh River	733	67	0.09	53	0.07	84	0.11
Morton	134	64	0.48	5	0.04	14	0.10
Chehalis Headwaters	980	309	0.32	20	0.02	18	0.02
Upper East Fork Lewis River	89	89	1.00	2	0.02	1	0.01
Mean (Std. Dev.):	315.5	101.4	0.43 (± 0.36)	10.4	0.02 (± 0.02)	23.9	0.10 (± 0.10)
Total:	2524	811	0.32	83	0.03	191	0.08

Table 9. Wilcoxon rank-sum test for two populations, comparing means (μ) of error distributions generated by the SMORPH and SHALSTAB models (see Type I error estimates in Table 8 and 10).

Test Criterion	Test Variable	Comparison of SMORPH (1) and SHALSTAB (2)	Comparison of SOILS (1) and SHALSTAB (2)
Type I errors: Existing landslides	n_1, n_2	8, 8	8, 8
	a_1, a_2	15.5, 48.5	53.0, 11.0
	W test statistic	0.04	0.01
	significant at $\alpha = 0.05$?	Yes; $\mu_1 < \mu_2$	Yes; $\mu_1 > \mu_2$
	significant at $\alpha = 0.01$?	No; $\mu_1 = \mu_2$	Yes; $\mu_1 > \mu_2$
Type I errors: Hazard-zonation map units	n_1, n_2	4, 4	N/A (see text)
	a_1, a_2	7.0, 9.0	
	W test statistic	0.44	
	significant at $\alpha = 0.05$?	No; $\mu_1 = \mu_2$	
	significant at $\alpha = 0.01$?	No; $\mu_1 = \mu_2$	

Table 10. Type I model errors, in which each model predicts that shallow landslides likely do not occur, whereas field-derived maps of hazard zonation indicate that there is a moderate to high likelihood of landsliding

Test Basin	Mass-Wasting Map Unit Data		SMORPH Model				SHALSTAB Model			
	Basin Area with Moderate to High Hazard Rating (km^2) (A)	Total Basin Acres (%)	Map Unit No.[1]	Basin Area Predicted with Low Hazard Rating (km*) (M)	(A/M) = P	E = P(A/T)	Map Unit No.[2]	Basin Area Predicted with Low Hazard Rating (km^2) (M)	(A/M) = P	E = P(A/T)
lordan-Boulder	73.9	0.55	1	22.8	0.31	0.14	6, 7, 8	14.3	0.19	0.09
Hazel	78.9	0.81	1	12.3	0.16	0.08	5, 6, 7, 8	12.5	0.16	0.08
Sol Duc River	2.7	0.01	1	1.0	0.39	0.01	6, 7, 8	1.1	0.43	0.01
Jpper East Fork .ewis River	9.0	0.11	1	1.7	0.19	0.01	6, 7, a	2.3	0.26	0.02
Total:	164.5 (T)			37.8				30.2		
lean:	41.1	0.37		9.5	0.26	0.06		7.6	0.26	0.05

[1] Map unit corresponds to "high" hazard potential as defined by gradient-curvature class (see Table 2)
[2] Map unit corresponds to "high" hazard potential as defined by precipitation rules (see Table 7).

Table 11. Type II model errors, in which each model predicts that shallow landslides likely have a high probability of occurring, whereas field-derived maps of hazard zonation indicate that there is a low likelihood of landsliding

Test Basin	Mass-Wasting Map Unit Data		SMORPH Model				SHALSTAB Model			
	Basin Area with Low Hazard Rating (km²) (A)	Total Basin Acres (%)	Map Unit No.[1]	Basin Area Predicted with High Hazard Rating (km²) (M)	(A/M) = P	E = P(A/T)	Map Unit No.[2]	Basin Area Predicted with High Hazard Rating (km²) (M)	(A/M) = P	E = P(A/T)
Jordan-Boulder	59.6	0.45	3	9.1	0.15	0.03	1, 2, 3, 4	17.8	0.30	0.05
Hazel	18.4	0.19	3	6.3	0.34	0.02	1, 2, 3	14.3	0.78	0.04
Sol Duc River	182.1	0.99	3	19.0	0.10	0.05	1, 2, 3, 4	26.3	0.14	0.08
Upper East Fork Lewis River	72.0	0.89	3	11.7	0.16	0.03	1, 2, 3, 4	30.8	0.43	0.09
Total:	332.1 (T)			46.1				89.2		
Mean:	83.0	0.63		11.5	0.19	0.03		22.3	0.41	0.07

[1] Map unit corresponds to "low" hazard potential as defined by gradient-curvature class (see Table 2).
[2] Map unit corresponds to "low" hazard potential as defined by precipitation rules (see Table 7).

Table 12. Comparison of model performance in correctly and incorrectly predicting landslide potential. For each model, slope-stability ratings of each DEM grid cell were compared with the landslide-inventory database. A numerical value was assigned to each of three possible database-intersection outcomes, as described in the text.

	SMORPH Model			SHALSTAB Model		SOILS Screen	
Test Basins	**Number of slides**	**Calibrated model value**	**Normalized calibrated value**	**$c'=2$ kN/m^2 $\phi = 33°$**	**Normalized value**	**Modeled value**	**Normalized Value**
Jordan-Boulder	155	5	0.03	11	0.07	80	0.52
Upper N. F. Stillaguamish	215	15	0.07	50	0.23	404	1.88
Hazel	117	3	0.03	84	0.72	68	0.58
Sol Duc	101	11	0.11	26	0.26	12	0.12
Middle Hoh	733	155	0.21	177	0.24	134	0.18
Morton	134	28	0.21	44	0.33	128	0.96
Chehalis Headwaters	980	49	0.05	40	0.04	618	0.63
Lewis	89	9	0.10	2	0.02	178	2.00
Total:	2524	275	0.81	434	1.91	1622	6.87

Table 13. List of required input variables for each model, their default values as specified in the model, and the data sources available to calibrate model output for the watershed of interest. The variable is assigned a rating of 1 if the required data for a specific watershed are relatively easy to obtain without substantial field work, and 2 for the converse.

Method	input Variable	Description	Unite of Measure	Default Values Set in the Model	Data Source	Relative Ease of Collecting Watershed-Specific Data: Rating Points Assigned	Relative Ease of Collecting Watershed-Specific Data: Percent Total Points
SOILS screen	none	no changes can be made without recreating the OIO algorithm and cover layer			N/A	0	0
SMORPH model	hillslope gradient	threshold gradients for different hazard-rating classes	percent	gradient class boundaries set at 15%, 24%, 47%, 70% (Table 2)	class-boundary values and threshold values can be adjusted based on landslide inventories and extrapolated to adjacent watersheds	1 (via landslide inventories and DEMs)	0.1
SHALSTAB model	soil cohesion (c′)	tree root cohesion	kN/m²	0, 2, 5, 8, 15	forest-soils and experimental studies (see Montgomery et al., 1998. for references)	1 (can be estimated via forest-age-class maps)	0.9
	soil transmissivity (T)	depth-integrated, but soil depth (and, hence, wetness) held constant	m²/day	65	soil surveys and isolated site-specific studies	2 (spatially and temporally variable)	
	soil depth (h)	given as constant	m	1	soil surveys and isolated site-specific studies	2 (spatially and temporally variable)	
	phi (ϕ)	internal angle of soil friction	degrees	if c′ = 0, then ϕ = 45"; else ϕ = 33"	rock-mechanics literature	2 (varies spatially by rock type)	
	soil bulk density (ρ_s)	given as constant	kg/m³	2000	soil surveys and isolated site-specific studies	2 (spatially and temporally variable)	

Table 14. Results of a sensitivity test for the SMORPH and SHALSTAB models. For each model, the values of the input variables were changed and the predicted shallow-landslide distributions were compared with the existing landslide inventories according to the point scheme described in the text. A. Total scores for each test basin, shown as cumulative and normalized by the respective number of landslides, for the SMORPH model when the threshold gradients for low, moderate, and high landslide-hazard potential are increased (i.e., shifted laterally along the horizontal slope-matrix plane as in Table 15). B. Total scores, shown as cumulative and normalized, for the SHALSTAB model when the effective cohesion (c') and friction angle (ϕ) are changed.

A. SMORPH model:

Test Basin	Change in Gradient Thresholds for Low, Moderate, and High Landslide-Potential Designations (see Figure___) Number of Assigned Points (Cumulative and Percent Total)											
	No. of Slides	Step 0	Norm. Step 0	Step 24	Norm. Step 24	Step 47	Norm. Step 47	Step 70	Norm. Step 70	Step ∞	Norm. Step ∞	Percent Change
Jordan-Boulder	155	5	0.03	6	0.04	47	0.30	124	0.80	465	3.00	0.01
NF Stilla-guamish	215	15	0.07	27	0..13	147	0.68	307	1.43	645	3.00	0.02
Hazel	117	3	0.03	40	0.34	109	0.93	170	1.45	351	3.00	0.01
Sol Duc	101	11	0.11	41	0.41	111	1.10	179	1.77	303	3.00	0.04
Mid. Hoh	733	155	0.21	309	0.42	686	0.94	1174	1.60	1466	2.00	0.11
Morton	134	28	0.21	25	0.19	72	0.54	146	1.09	266	2.00	0.11
Chehalis	980	49	0.05	79	0.08	260	0.27	771	0.79	2940	3.00	0.02
EF Lewis	89	9	0.10	6	0.07	32	0.36	88	0.99	176	2.00	0.05
TOTAL:	2524	275	0.81	533	1.67	1464	5.12	2959	9.92	6616	21.00	0.04

[†] **Percent change from preferred values (i.e., those resulting in predictions most similar to the spatial distribution of existing landslides; Step 0) and maximum values at which the hillslopes are predicted to be fully stable (i.e., Step ∞).**

Table 14 cont'd. Results of a sensitivity test for the SMORPH and SHALSTAB models. For each model, the values of the input variables were changed and the predicted shallow-landslide distributions were compared with the existing landslide inventories according to the point scheme described in the text. B. Total scores, shown as cumulative and normalized, for the SHALSTAB model when the effective cohesion (c') and friction angle (ϕ) are changed as per the default values given by the model.

B. SHALSTAB model:

Test Basin	Change in Values of Effective Cohesion (c' = kN/m²) and Friction Angle (ϕ = degrees) Number of Assigned Points (Cumulative and Percent Total)									
	No. of Slides	c' = 2 ϕ = 33	Norm. value	c' = 0 ϕ = 45	Norm. value	c' = 5 ϕ = 33	Norm. value	c' = 8 ϕ = 33	Norm. value	Percent Change
Jordan-Boulder	155	11	0.07	15	0.10	90	0.58	465	3.00	0.02
NF Stilla-guamish	215	50	0.23	95	0.44	278	1.29	645	3.00	0.08
Hazel	117	84	0.72	116	0.99	182	1.56	351	3.00	0.24
Sol Duc	101	26	0.26	56	0.55	154	1.52	303	3.00	0.09
Mid. Hoh	733	177	0.24	281	0.38	851	1.16	1466	2.00	0.12
Morton	134	44	0.33	74	0.55	167	1.25	268	2.00	0.17
Chehalis	980	40	0.04	72	0.07	2940	3.00	2940	3.00	0.01
EF Lewis	89	2	0.02	9	0.10	60	0.67	178	2.00	0.01
TOTAL:	2524	434	1.91	718	3.19	4722	TT.04	6616	21.00	0.09

[1] Percent change from preferred values (i.e., those resulting in predictions most similar to the spatial distribution of existing landslides; c' = 2kN/m² and ϕ = 33") and maximum values at which the hillslopes are predicted to be fully stable (i.e., c' = 8kN/m² and ϕ = 33").

Table 15. Adjustment of the SMORPH slope matrix to test the sensitivity of increasing threshold gradients on model predictions of shallow-landslide potential. Step categories refer to step-wise shifls of the hazard-zonation criteria (i.e., low, moderate, high) with respect to the designated gradient-threshold classes, for the Olympic Peninsula test basins. L = predicted low shallow-landslide potential; M = moderate potential; H = high potential.

Step 0: (Calibrated using landslide inventories)

Slope Curvature	Gradient (%)				
	0	15	25	47	70 to ∞
convex	L	L	L	L	M
planar	L	L	L	M	H
concave	L	M	H	ii	H

step 47:

Slope Curvature	Gradient(%)						
	0	15	25	47	70	90	100 to ∞
convex	L	L	L	L	L	L	M
planar	L	L	L	L	L	M	H
concave	L	L	L	M	H	H	H

step ∞:

Slope Curvature	Gradient (%)				
	0	15	25	47	70 to ∞
convex	L	L	L	L	L
planar	L	L	L	L	L
concave	L	L	L	L	L

Step 24:

Slope Curvature	Gradient (%)					
	0	15	25	47	70	90 to ∞
convex	L	L	L	L	L	M
planar	L	L	L	L	M	H
concave	L	L	M	H	H	H

step 70:

Slope Curvature	Gradient (%)							
	0	15	25	47	70	90	100	110 to ∞
convex	L	L	L	L	L	L	L	M
planar	L	L	L	L	L	L	M	H
concave	L	L	L	L	M	H	H	H

Table 16. Rating scheme used to compare the management applicability of models using scientific and technical criteria discussed in the text.

TEST CRITERIA			RATING SCHEME	MODEL TESTED: SMORPH	SHALSTAB $c' = 2$ kN/m^2, $\phi = 33°$	WDNR SOILS SCREEN
SCIENTIFIC CRITERIA	Model performance	Comparison with landslide inventory-Type I model errors	See text and Tables 8, 10, 11, 12	0.03	0.08	0.32
		Comparison with Hazard-Potential Maps-Type I errors		0.06	0.05	1
		Comparison with Hazard-Potential Maps-Type II errors		0.03	0.07	1
		Comparison of overall predictive capability		0.8	1.9	7.0
	Method limitations	For greatest predictive accuracy. does the model need to be calibrated with field data?	Yes = 0 No = 1	0	0	1
		Input-variable data accessibility and adequacy	See Table 13	0.1	0.9	1
		Model accounts implicitly or explicitly for spatial variability of input variables	See Table 17	0.0	0.7	0.2
		Model sensitivity to changes in input variables	See text, Table 14	0.04	0.09	1
	Geographic applicability	Ability of model to correctly identify slides in each of the following terrain types:	See text and Table 8			
		Continental- glaciated terrain		0.01	0.32	0.29
		Cascade volcanics		0.03	0.06	0.74
		NW Cascades system		0.0	0.06	0.60
		Olympic core rocks		0.01	0.12	0.06
		western Olympic Assemblage		0.07	0.11	0.09
		Eocene volcaniclastics		0.02	0.02	0.32
	Management applications	Are management criteria (L,M,H hazard) built in to the model?	Yes = 0 No = 1	0	1	0
		Are models available to the general public?	Yes = 0 No = 1	0	0	0

		Can the following persons run model (assuming access to system that can run programs):				
SCIENTIFIC CRITERIA	Mgmt. Appl.	No GIS experience		1	1	0
		GIS experience		0	0	0
		Are model results interpretable by the following persons?:				
		No mass-wasting mapping experience	Yes = 0 No = 1	1	1 (existing model with no mgmt. criteria)	0
		Mass-wasting mapping experience		0	0 true for geomor-phologists and forest hydrologists	0
	Modification requirements	Can model be adjusted to work in all western WA. terrains?	Yes = 0 No = 1	0	0	1
		Is it essential that models include management criteria to be interpretable in the current foliowing arenas?:				
		regulatory application		1	1	1
		management application (e.g., harvest and road planning)	Yes = 1 No = 0	1	1 in most instances	1
		academic (e.g., for research and analysis)		0	0	0
		Can model be adjusted to include other key variables if topographic controls are not dominant in the watershed?	Yes =o No = 1	1	0	1
'ECHNICAL :RITERIA	Computer run time	10 m DEM	average time per basin	0.20	0.92	N/A
		30 m DEM		0.05	0.30	N/A
		10 m DEM	expected time to create w WA. coverage	(0.13	0.62	N/A
		30 m DEM		0.03	0.20	N/A

: 2 days)

.dequate = 0

	'raining equirements	How much training is needed to run model (assuming basic computer skills)?	None = 1 S... (<5 hrs.)	D 1 0	011	001
		How much training (i.e., office and field) is needed to interpret model output?	0 = some (± 1 day) 1 = more (≥ 2 days)			
'ECHNICAL :RITERIA	'raining equirements	What logistical documentation exists?	Adequate = o None = 1	Developed during this study		
	)ata equirements	Can model be run with DEMs and/or default values as the only required input?	Yes = 0 No = 1	0	0	1
		Does model accuracy improve with increasing DEM resolution?	Yes = 0 No = 1	0 (avg. 94% improve-ment)	0 (avg. 60% improvement)	1
	)ata storage & retrieval	Which model uses the biggest storage space?	Bigger = 1 Smaller = 0	0	1 (Needs ~ 5 X storage space of SMOKPH)	I
		Can model be run on a PC with ARC/INFO software?	Yes = 0 No = 1	0	0	0
		Can model be run on a PC with non-ARC/INFO GIS?	Yes, with additional program-ming = 1 Yes, w/o additional program-ming = 0	1	1	0
		Are there potential problems for PC users re: data storage requirements for areas larger than one WAU?	Yes = 1 No = 0	0	0	0
		Are there potential problems for PC users re: data storage requirements for areas larger than several WAUs?	Yes = 1 No = 0	0	1	1
	Modification 'equirements	Is model adequately documented internally (e.g., comment lines) for ease in adjusting input variables, or externally for interpreting results?	Yes = 0 No - I	1	1	N/A

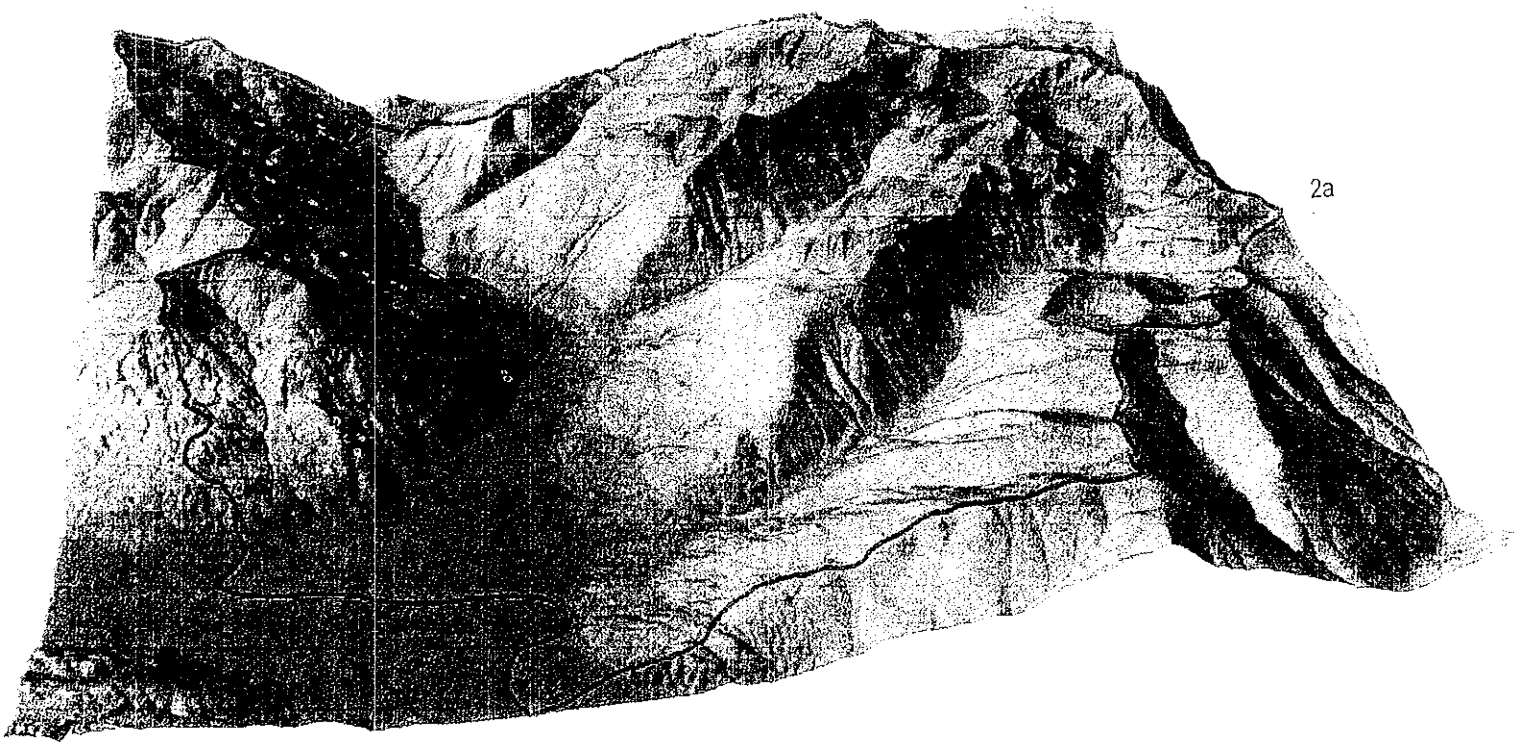
		Does model need more work and/or programming to adapt it for management use?	Yes = 1 No = 0	0	1	0
TOTAL SCORE:				9.6	16.5	23.6

Table 17. Criteria used for rating scheme (Table 16.)

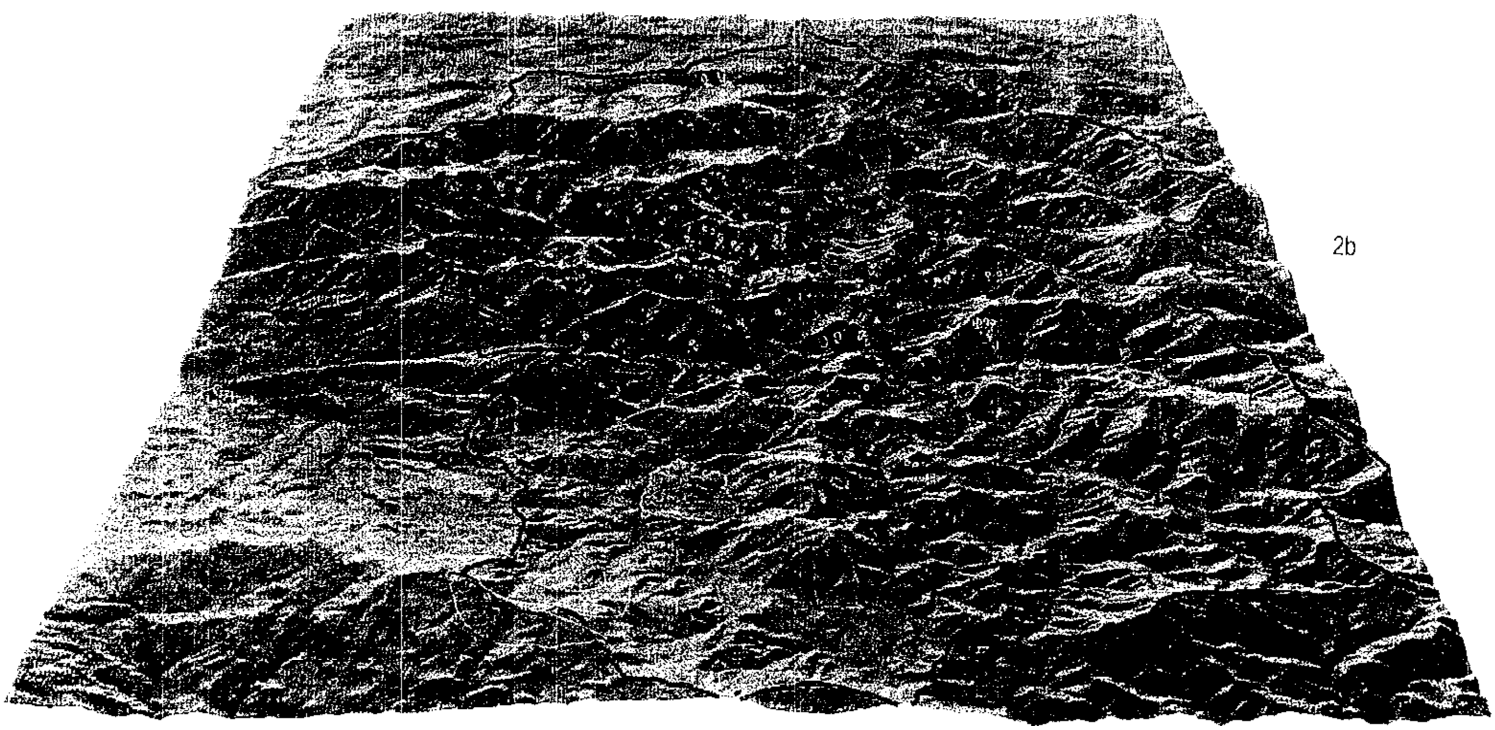
		TEST CRITERIA	Rationale Used for Point Assignment
SCIENTIFIC CRITERIA	Model performance	Comparison with landslide inventory - Type I model errors	See Table 8
		Comparison with Hazard-Potential Maps -Type I errors	See Table 10
		Comparison with Hazard-Potential Maps -Type II errors	See Table 11
		Comparison of overall predictive capability	See Table 12
	Method limitations	For greatest predictive accuracy, does the model need to be calibrated with field data?	Both of the models should be calibrated. The SOILS data cannot be calibrated.
		Input-variable data accessibility and adequacy	See Table 13. The SOILS data cannot be updated.
		Model accounts implicitly or explicitly for spatial variability of input variables	Topographic variables explicit for SHALSTAB and SMORPH (0 pts.), not for SOILS (lpt.). SHALSTAB cohesion explicit (0 pt.). transmissivity, depth, phi, and bulk density set as constants (4 pts.); if not held constant, assign 0 pts. Soil properties implicit in SMORPH and SOILS (0 pts.). Sum total and divide by total number of points possible.
		Model sensitivity to changes in input variables	See Table 14. The SOILS data cannot be updated and thus is insensitive.
	Geographic applicability	Ability of model to correctly identify slides in each of the following terrain types:	
		Continental- glaciated terrain	See Table 8.
		Cascade volcanics	
		NW Cascades system	
		Olympic core rocks	
		Western Olympic Assemblage	
		Eocene volcaniclastics	
	Management applications	Are management criteria (L,M,H hazard) built in to the model?	SHALSTAS does not have management criteria set.
		Are models available to the general public?	All tested models are available to the public.
		Can the following persons run model (assuming access to system that can run programs):	

SC::NTIFIC CRITERIA	Mgmt. Appl.	No GIS experience	The SOILS data is easy to access with no GIS experience. Both models require some GIS experience to access and run.
		GIS experience	
		Are model results interpretable by the following persons?:	
		No mass-wasting mapping experience	Experience with mass-wasting concepts is a necessary ingredient in understanding both models outputs. The SOILS data does not require this experience.
		Mass-wasting mapping experience	
	Modification requirements	Can model be adjusted to work in all western WA. terrains?	The SOILS data cannot be adjusted.
		Is it essential that models include management criteria to be interpretable in the current following arenas?:	
		regulatory application	To be useful as a regulatory tool, any model must have management criteria.
		management application (e.g., harvest and road planning)	Most management applications would benefit from having criteria set.
		academic (e.g., for research and analysis)	It is not necessary for criteria to be set for strictly academic uses of any model.
		Can model be adjusted to include other key variables if topographic controls are not dominant in the watershed?	The SMORPH model assumes topography control: landslide behavior. If this is not the case, model output suffers. The SOILS data cannot be calibrated.
TECHNICAL CaTERIA	Computer run time	10 m DEM-time per basin	Divide average number of minutes to complete a model run by 60. The SOILS data is a static layer and as such, it requires no time to run.
		30 m DEM-time per basin	
		10 m DEM-time for western VVA	Divide average number of hours by 672 (number of hours in a month). The SOILS data is a static layer and as such, it requires no time to run.
		30 m DEM-time for western WA	
	Training requirements	How much training is needed to run model (assuming basic computer skills)?	Training would consist of how to access the models, determine whether the model is appropriate for the intended use, how to calibrate the models, and how to interpret the model results
		How much training (i.e.. office and field) is needed to interpret model output?	Because SHALSTAB does not have management criteria, it is important to include extra training to understand how to use that models output in the area of interest. Some knowledge of hydrology is useful.
TECHNICAL CaTEaA	Training requirements	What logistical documentation exists?	No documentation exists for the SOILS layer regarding its use as a slope stability screen.
	Data requirements	Can model be run with DEMs and/or default values as the only required input?	Both models can be run using a DEM and the default values. The SOILS layer is a static coverage and therefore does not require a DEM.
		Does model accuracy improve with increasing DEM resolution?	See Figure 14. The SOILS layer is a static coverage and therefore does not require a DEM.

	Data storage & retrieval	Which model uses the biggest storage space?	Both models produce grid type data, that requires less storage space than coverage type data, like SOILS. However. SHALSTAB produces volumes o' extraneous data.
		Can model be run on a PC with ARC/INFO software'?	In addition to ARC/INFO, the PC user must also have Fortran and C to run SHALSTAB.
		Can model be run on a PC with non-ARC/INFO GIS?	The SOILS layer is an existing coverage and does not necessarily need ARC/INFO software to create a map.
		Are there potential problems for PC users re: data storage requirements for areas larger than one WAU?	For a small area (a WAU or two), there should be no data storage problems.
		Are there potential problems for PC users re: data storage requirements for areas larger than several WAUs?	Because both SOILS and SHALSTAB require more disk space, over large areas (e.g.. WRIAs), there may be data storage problems.
	Modification requirements	Is model adequately documented internally (e.g., comment lines) for ease in adjusting input variables, or externally for interpreting results?	None of the models tested were more than skeletally documented internally. The SOILS data has no need for internal documentation. as it is not a program.
		Does model need more work and/or programming to adapt it for management use?	The SHALSTAB model does not currently have management criteria.



2a



2b

Figure 2. Shaded relief maps of the basins with the most and least amount of topographic relief. Red lines indicate basin boundaries, white circles are landslide locations. Figure 2a is the Jordan-Boulder basin, with a view to the east, up the Cascade River. The valleys that contain the Jordan, Boulder, and Irene creeks are on the right, Monogram Peak is on the left Figure 2b is the Chehalis Headwaters basin, which has the least amount of topographic relief. The view is to the north, towards the town of Pe Ell.

Figure 1. Index Map of western Washington. The inset map shows the coterminous United States with Washington state shaded. In the main map, the test basins used in this study are highlighted. County boundaries and some of the major cities of western Washington are shown for orientation. Heavy lines describe the approximate boundaries of the major geomorphic terranes.

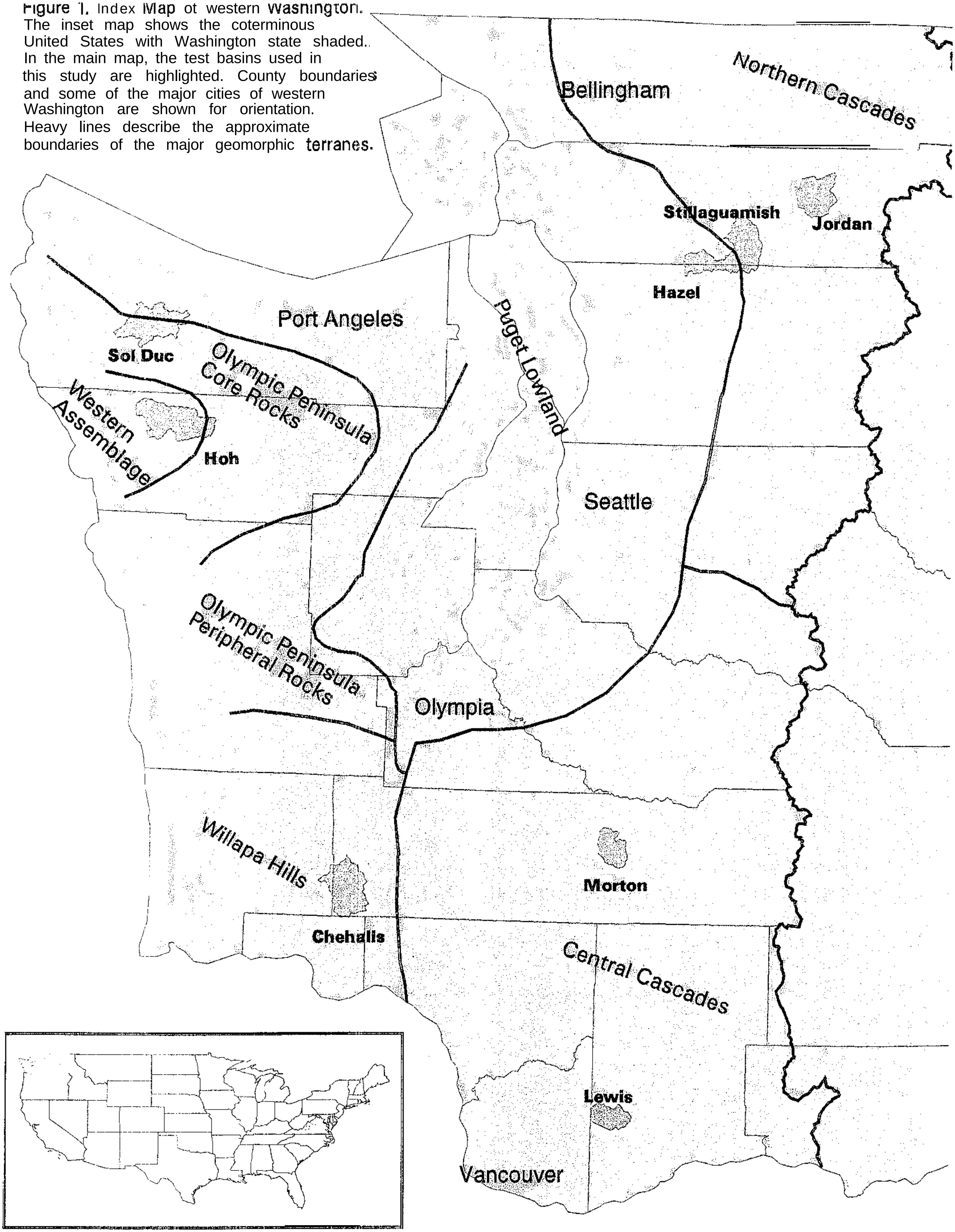
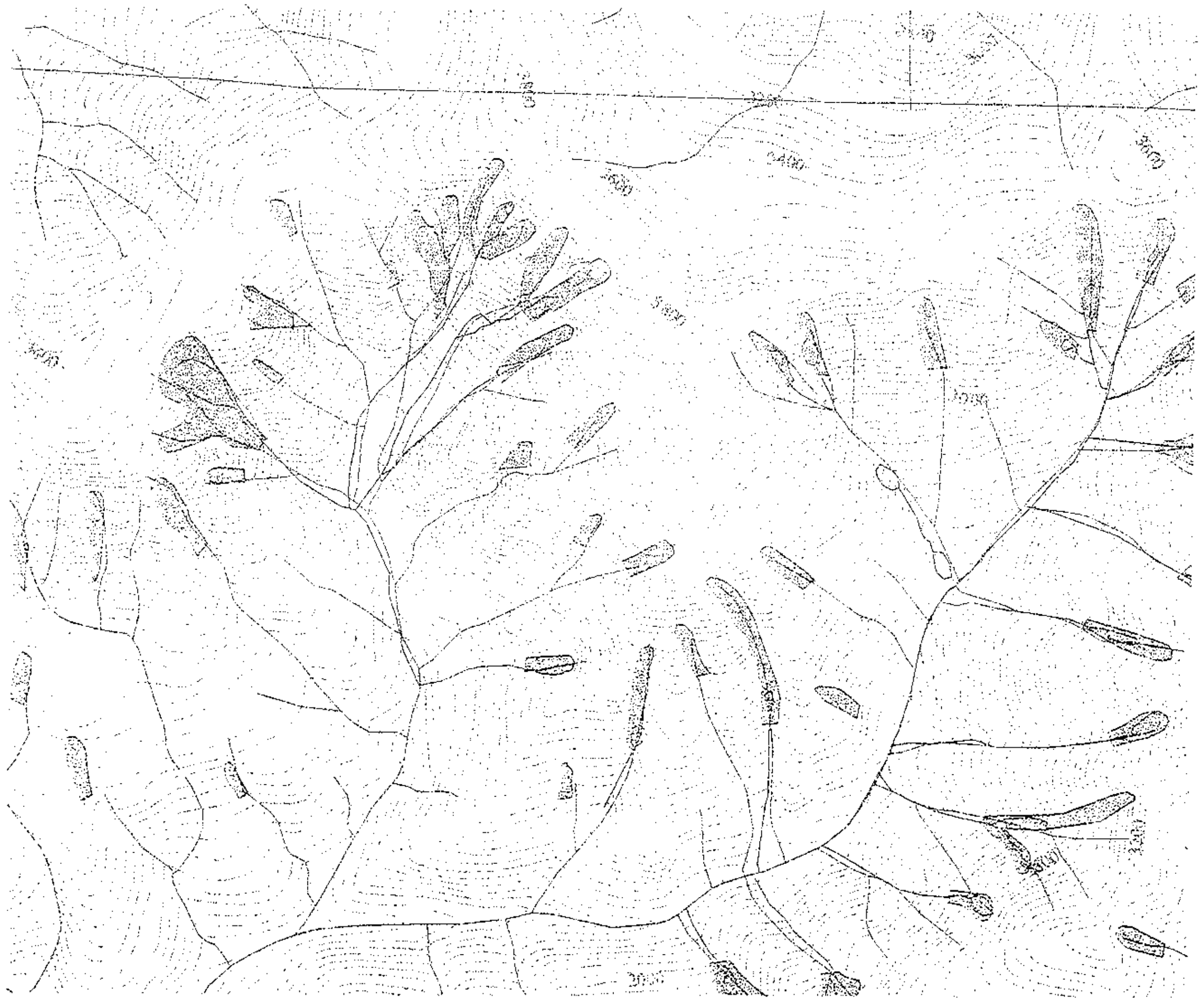


Figure 3. Comparison maps of the SHALSTAB, MWMU, SMORPH, and SOILS data for a portion of the Jordan-Boulder and Lewis basins. The hydrography is drawn in blue, the landslide locations are depicted as black polygons and circles. The lack of SOILS data in the Jordan-Boulder basin (3b, upper right) led to large numbers of slides having an inventoried 'no data' value. The broad-brush approach to MWMU mapping in the Jordan-Boulder basin led to large inclusions of the landbase into a high hazard category compared with the modeled output. In the Lewis basin (3a), no soils information on stability exists. The fine-scale approach to MWMU mapping in the Lewis basin more closely approximates the modeled output. In the Chehalis Headwaters basin (3c), the SOILS data is drawn with the hydrography, topography, and landslide inventory on the left. Much of the basin is in the unstable or very unstable SOILS categories irrespective of whether the ground is at the ridgetop, on a sideslope, or in the valley. The SHALSTAB and SMORPH datasets (on the right) both show significantly less of the landbase in an unstable category, and the areas that are in an unstable category are all topographically based.

3a



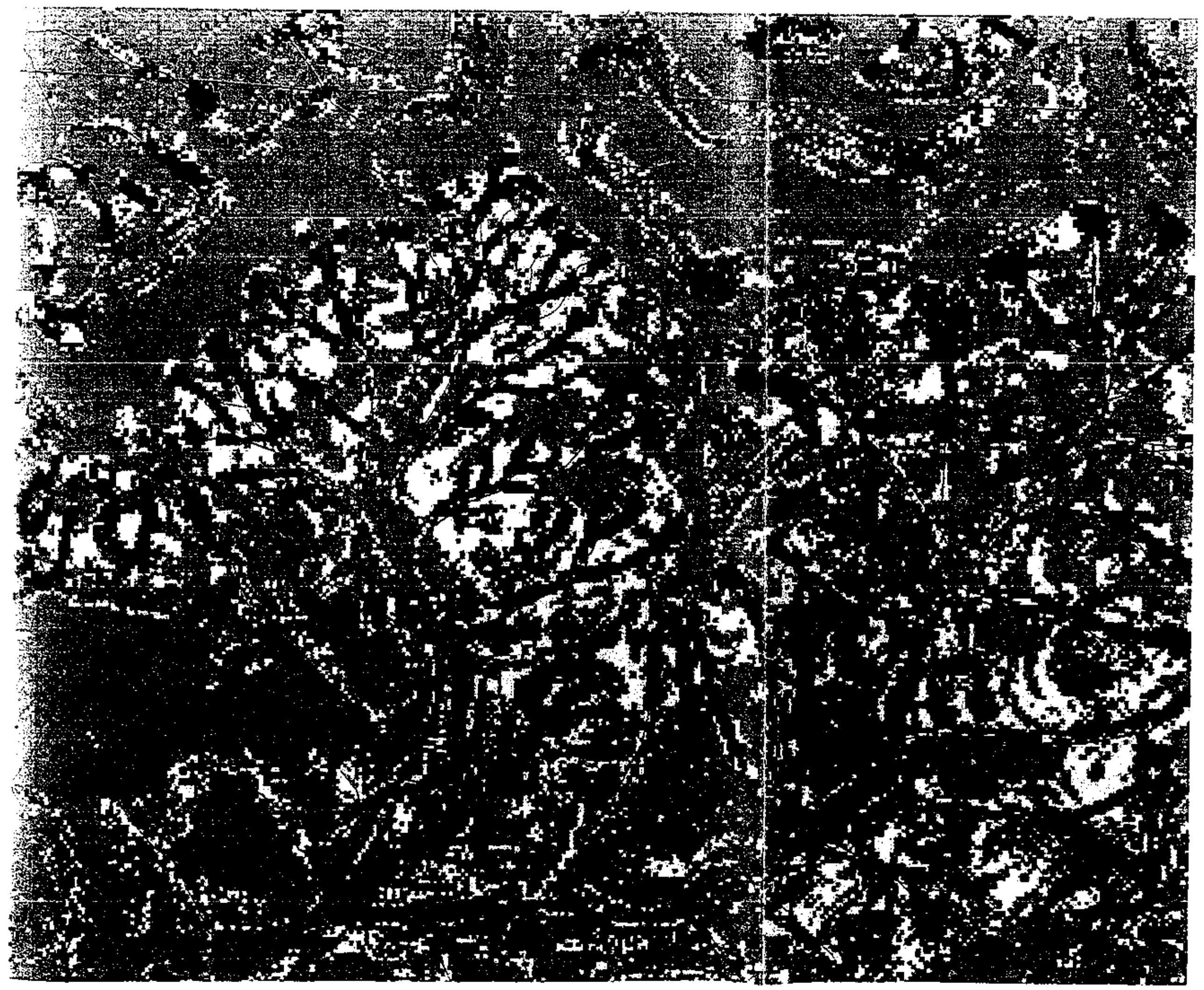
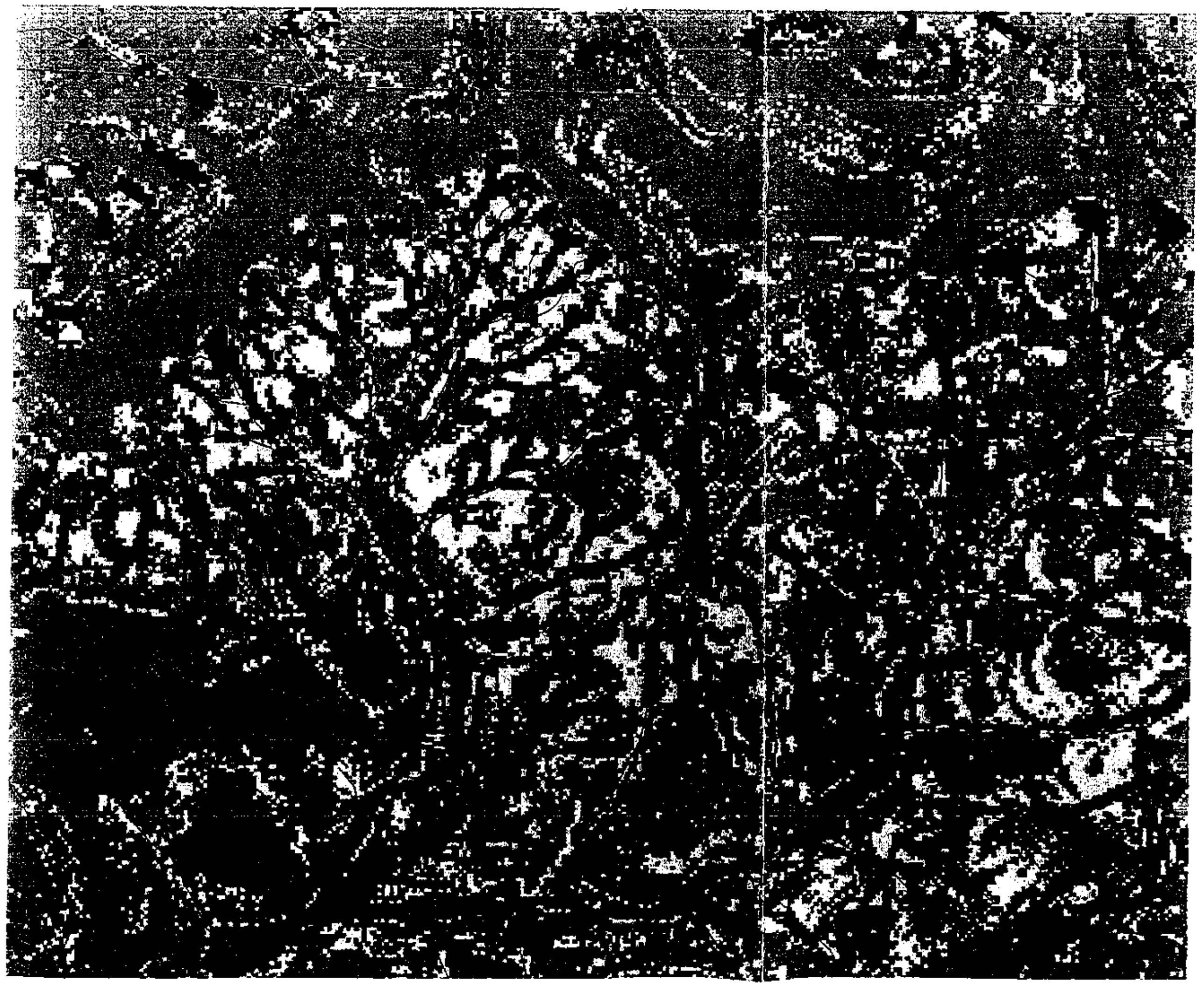


Figure 3. Comparison maps of the SHALSTAB, MWMU, SMORPH, and SOILS data for a portion of the Jordan-Boulder and Lewis basins. The hydrography is drawn in blue, the landslide locations are depicted as black polygons and circles. The lack of SOILS data in the Jordan-Boulder basin (3b, upper right) led to large numbers of slides having an inventoried 'no data' value. The broad-brush approach to MWMU mapping in the Jordan-Boulder basin led to large inclusions of the landbase into a high hazard category compared with the modeled output. In the Lewis basin (3a), no soils information on stability exists. The fine-scale approach to MWMU mapping in the Lewis basin more closely approximates the modeled output. In the Chehalis Headwaters basin (3c), the SOILS data is drawn with the hydrography, topography, and landslide inventory on the left. Much of the basin is in the unstable or very unstable SOILS categories irrespective of whether the ground is at the ridgetop, on a sideslope, or in the valley. The SHALSTAB and SMORPH datasets (on the right) both show significantly less of the landbase in an unstable category, and the areas that are in an unstable category are all topographically based.

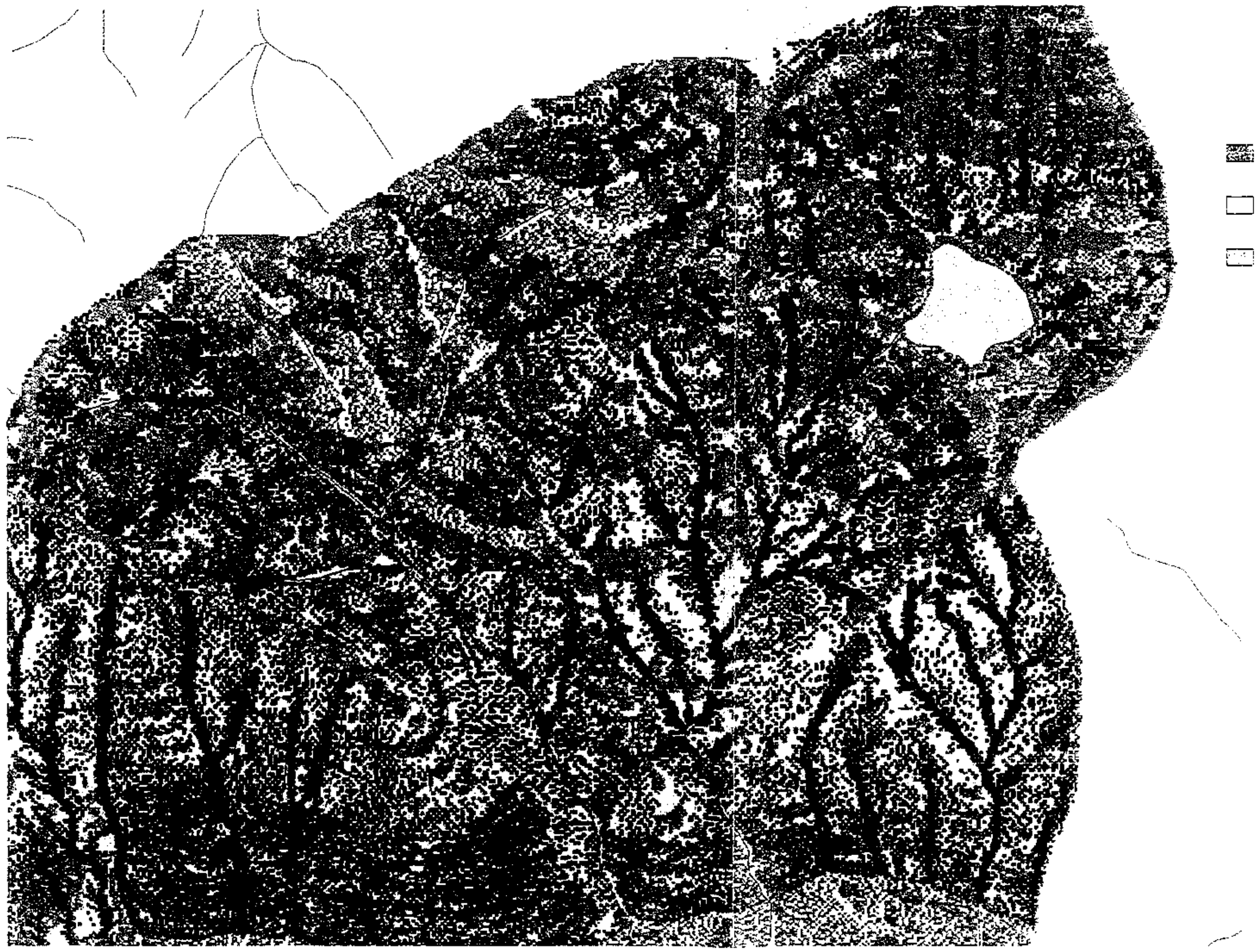
3a




Unconditionally unstable at this cohesion
0 - 50 mm/day
50 - 100 mm/day
100 - 200 mm/day
200 - 400 mm/day
> 400 mm/day
Unconditionally stable
Stable at this cohesion


DNR Soils Layer Stability Descriptions
No Data
Unstable
Very Unstable


Mass Wasting Map Unit Description of Stability Hazards
Low Hazard
High Hazard


Smorph calls
Most Sensitive to Mass Wasting.
Moderately Sensitiv to Mass Wasting.
Least Sensitive to Mass Wasting.

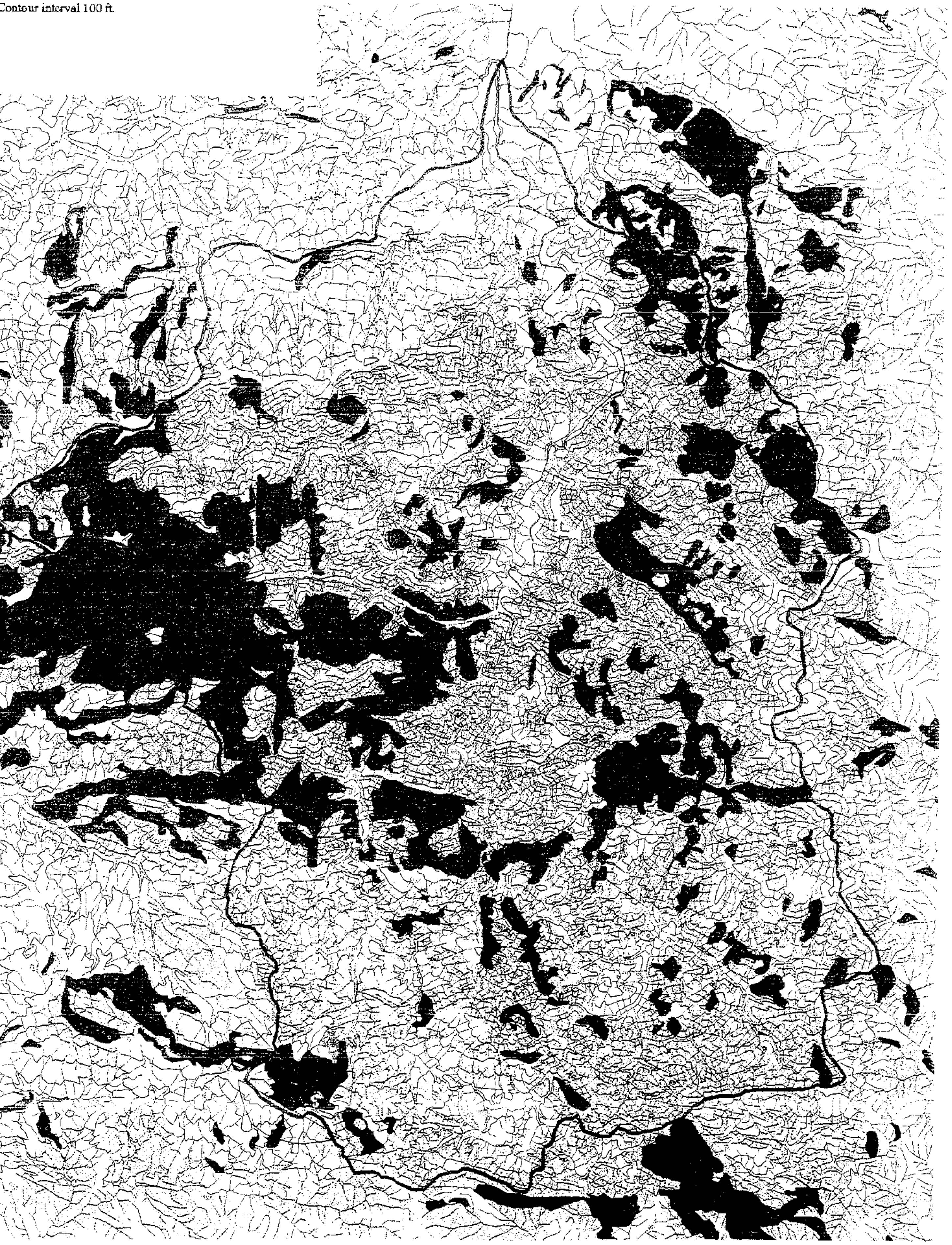
DNR Soils Layer
Stability Descriptions

Stable or No Data

Unstable

Very Unstable

Contour interval 100 ft.



Unconditionally unst
at this cohesion

0 - 50 mm/day

50 - 100 mm/day

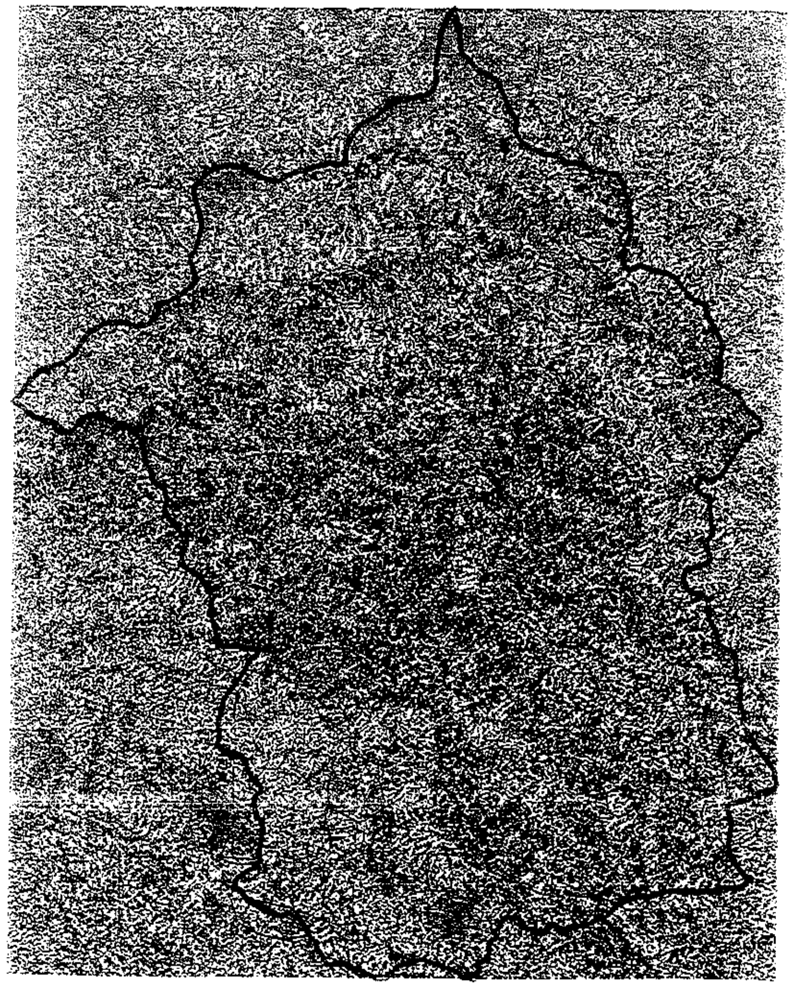
100 - 200 mm/day

200 - 400 mm/day

> 400 mm/day

Unconditionally stabl

Stable at
this cohesion

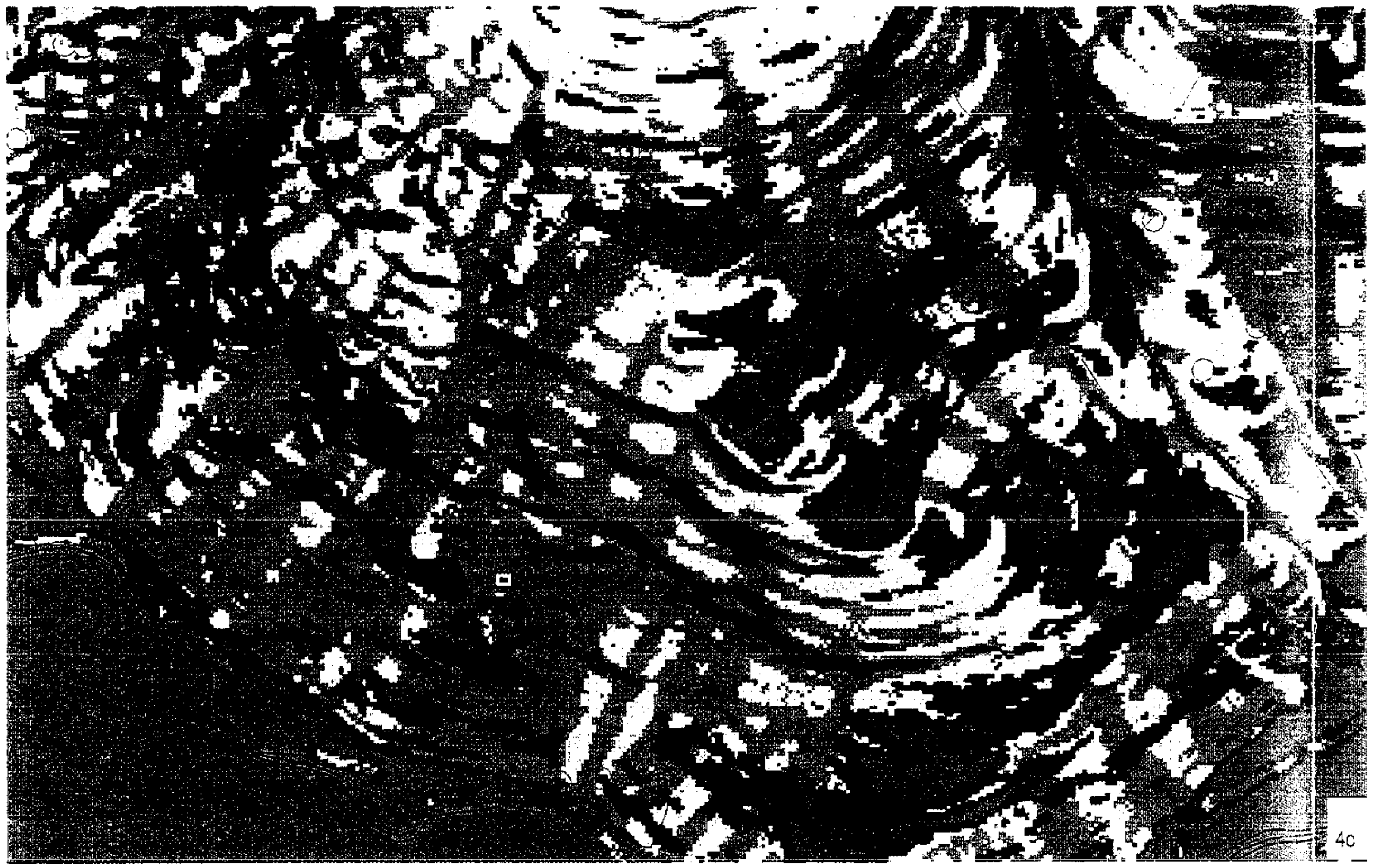


Smorph cells

Most Sensitive to
Mass Wasting.

Moderately Sensitive
to Mass Wasting.

Least Sensitive to
Mass Wasting.



Smorph calls

Most Sensitive to Mass Wasting.

Moderately Sensitive to Mass Wasting.

Least Sensitive to Mass Wasting.

Unconditionally unstable at this cohesion

0 - 50 mm/day

50 - 100 mm/day

100 - 200 mm/day

200 - 400 mm/day

> 400 mm/day

Unconditionally stable

Stable at this cohesion

Figure 4. Digital Elevation Model (DEM) artifacts. Tiling artifacts (4a, lower left) can arise when DEM data from different quadrangles are appended together. They express themselves as false cliffs oriented north-south and east-west along quadrangle boundaries. Tiling artifacts are common in the 30m resolution data, but rare in the 10m resolution data, as great care was used to remove tile artifacts in the finer-scale data. Edge effects (4b, lower right) occur along the outermost rind of pixels of a DEM, where the model is not able to correctly identify the slope characteristics of the edge pixels relative to its eight nearest neighbors. Elevation banding (4c, upper left) occurs only in the 10m resolution data, and is most noticeable in basins with high relief.

Smorph cells

Most Sensitive to Mass Wasting.

Moderately Sensitive to Mass Wasting.

Least Sensitive to Mass Wasting.

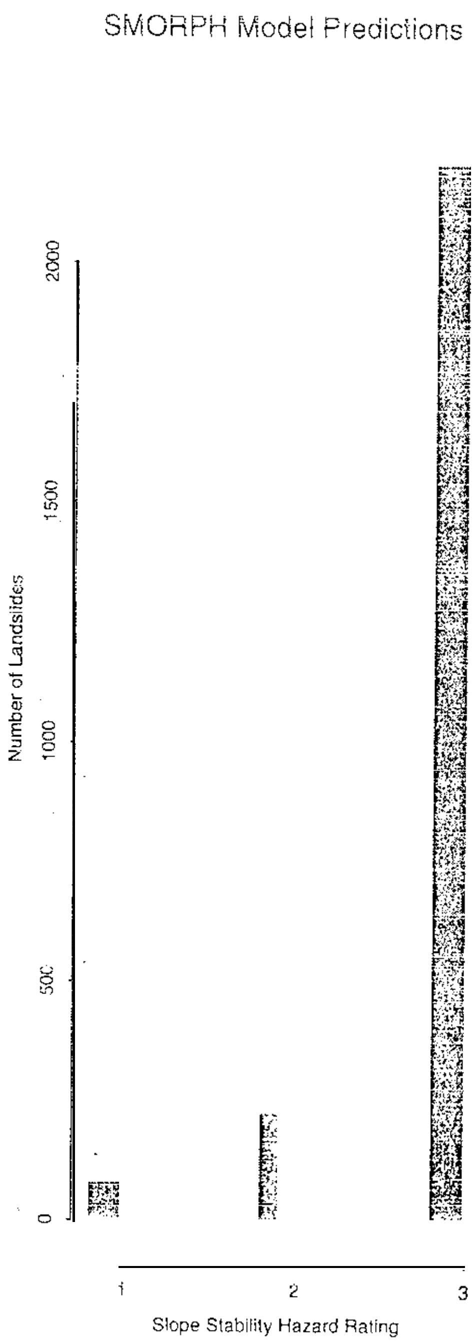
Figure 10. Map of the Jordan-Boulder basin with SMORPH plotted against SHALSTAB using management criteria. In order to directly compare models, the SHALSTAB model output was categorized into high, moderate, and low hazard. The areas within these categories were then compared with those of SMORPH, which outputs management criteria automatically.

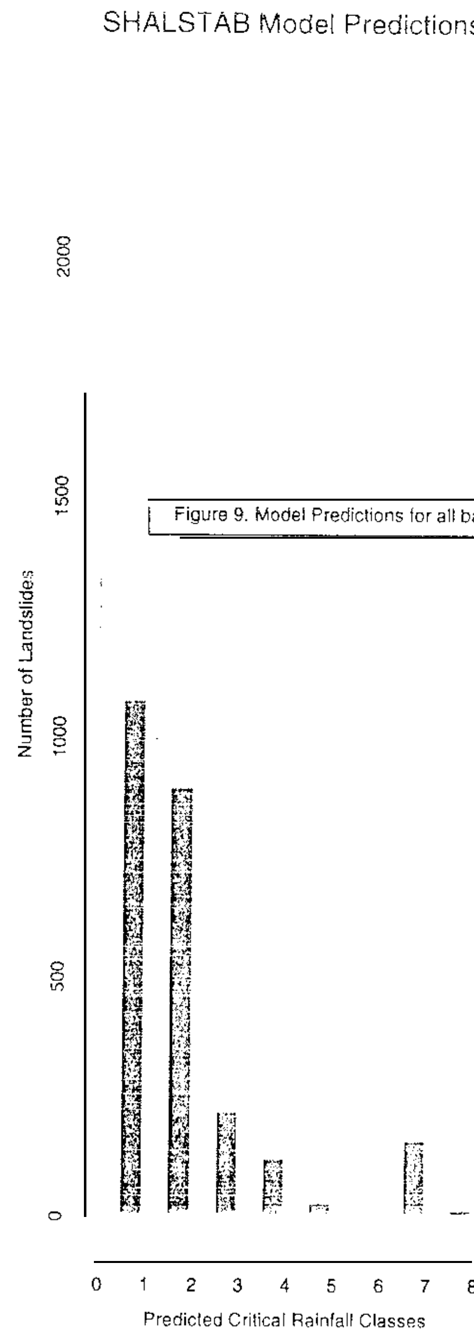
Shalstab using management criteria

Unstable

Moderately Unstable

Stable





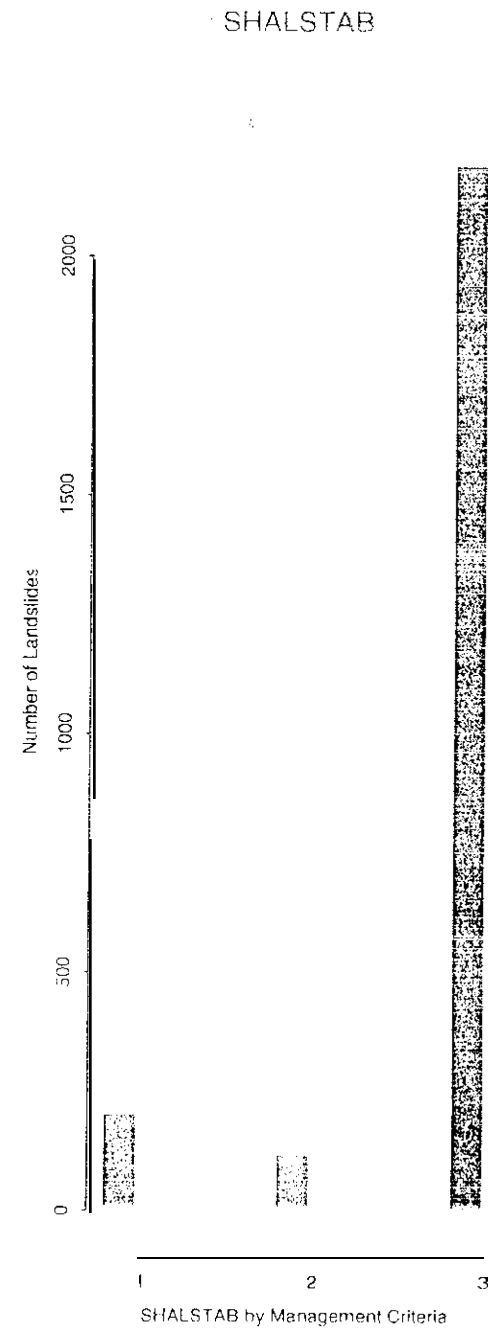


Figure 9. Model Predictions for all basins

Figure 8. Map of original SHALSTAB and SHALSTAB shaded by management criteria.

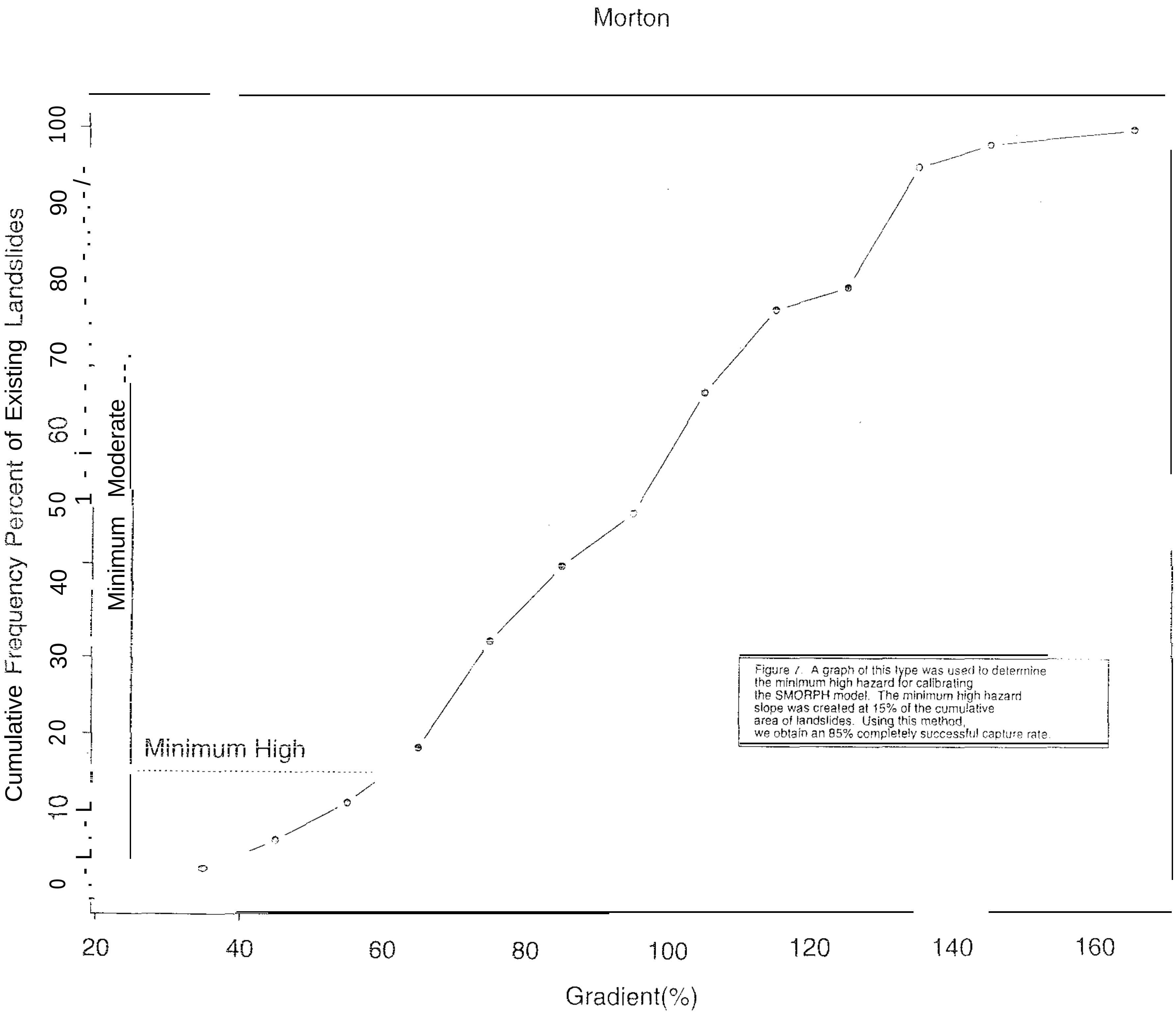


Figure 7. A graph of this type was used to determine the minimum high hazard for calibrating the SMORPH model. The minimum high hazard slope was created at 15% of the cumulative area of landslides. Using this method, we obtain an 85% completely successful capture rate.

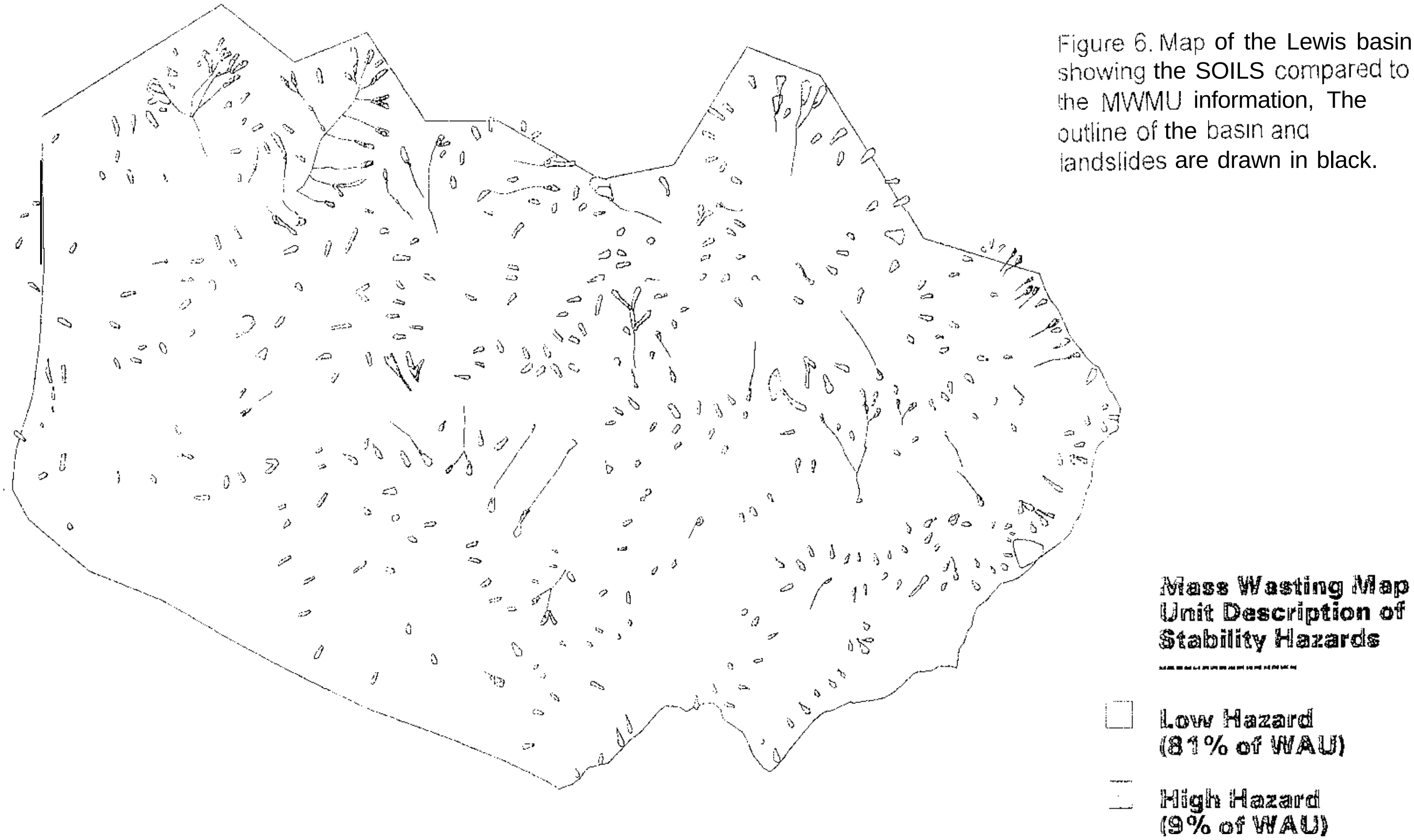


Figure 6. Map of the Lewis basin showing the SOILS compared to the MWMU information, The outline of the basin and landslides are drawn in black.

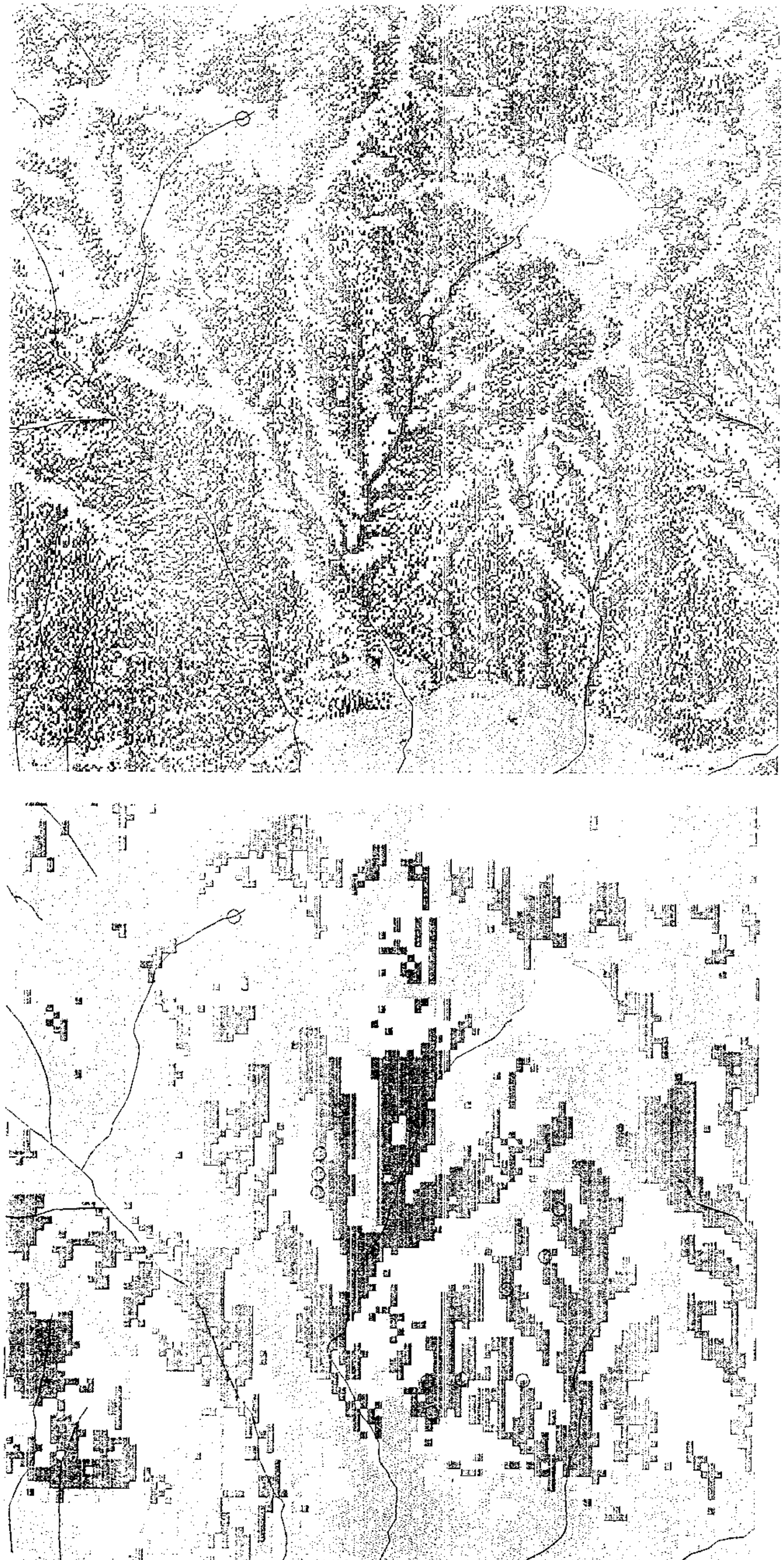


Figure 5. Comparison maps of 10m vs 30 resolution DEM data for a portion of the Jordan-Boulder basin, with streams in blue and landslide locations in black. The resolution of the DEM greatly influences the ability of the model to predict landslide prone terrain. The upper map is derived from the 10m DEM. Note the greater ability of the 10m data to resolve small bedrock hollows and stream channels, where shallow landslide commonly occur.

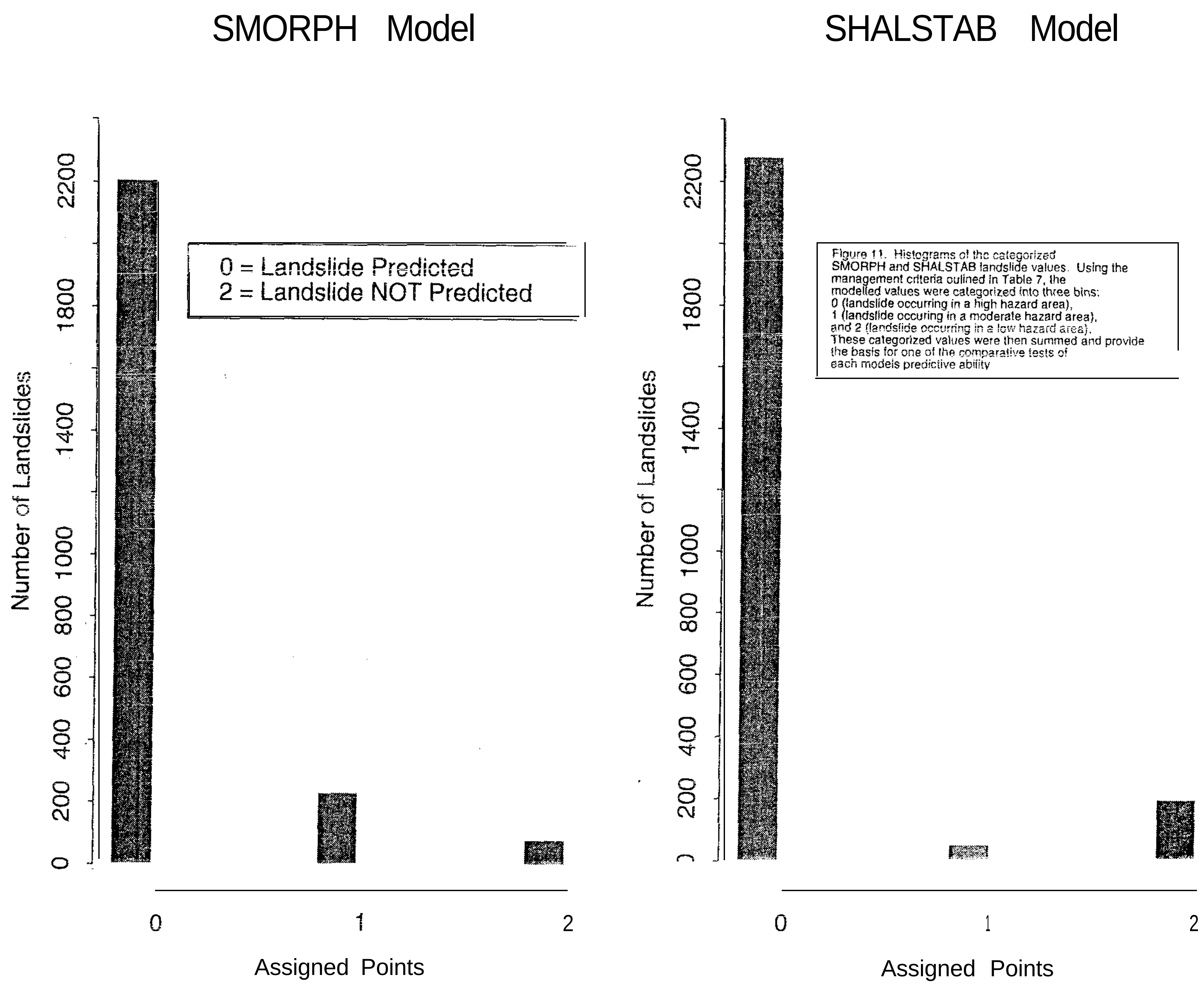
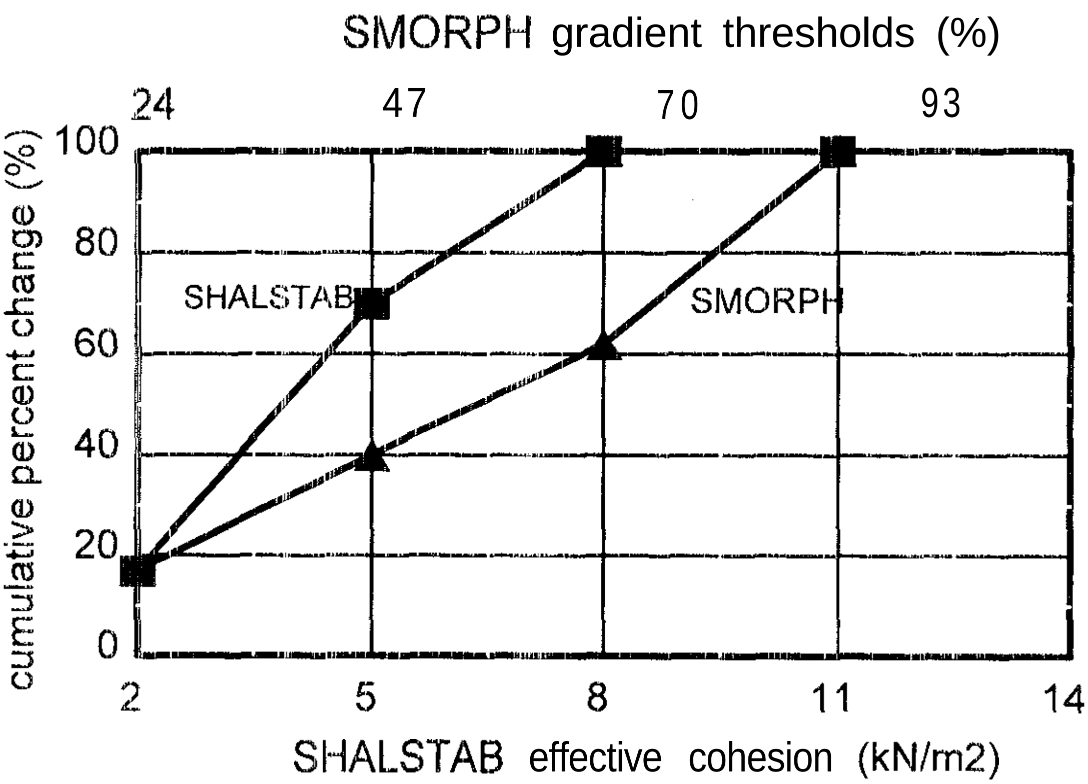


Figure 11. Histograms of thc categorized SMORPH and SHALSTAB landslide values. Using the management criteria oulined in Table 7, the modelled values were categorized into three bins: 0 (landslide occurring in a high hazard area), 1 (landslide occuring in a moderate hazard area), and 2 (landslide occurring in a low hazard area). These categorized values were then summed and provide the basis for one of the comparative tests of each models predictive ability

Figure 12. Cumulative percent change in the number of correctly and incorrectly predicted landslides for: (1) increasing effective cohesions (c' = kN/m^2) input to the SHALSTAB model; and, (2) increasing gradient-threshold values (S = %) input to the SMORPH model. This graph permits visual comparison of the relative sensitivities of the models when the value of input variables is changed.



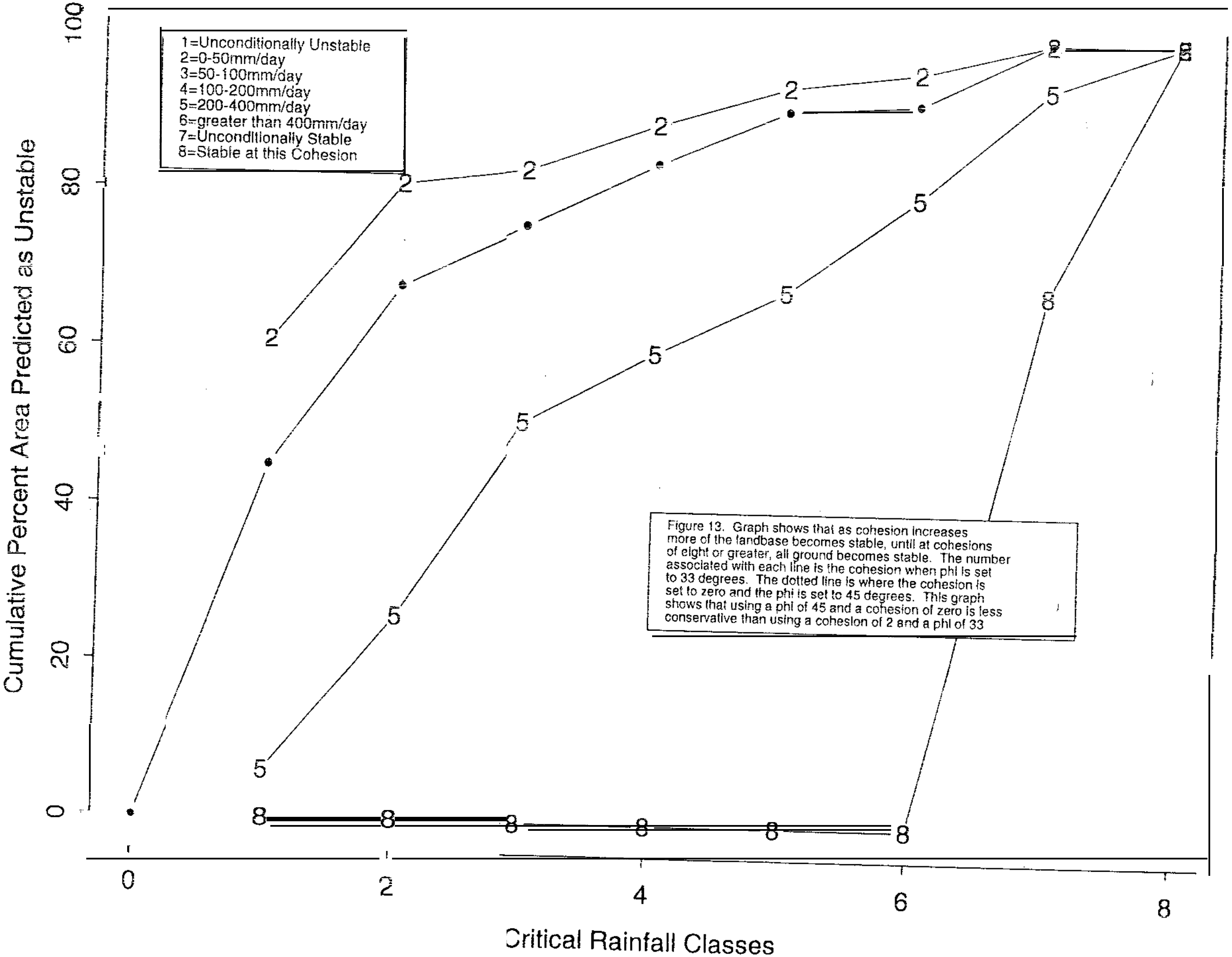


Figure 13. Graph shows that as cohesion increases more of the landbase becomes stable, until at cohesions of eight or greater, all ground becomes stable. The number associated with each line is the cohesion when phi is set to 33 degrees. The dotted line is where the cohesion is set to zero and the phi is set to 45 degrees. This graph shows that using a phi of 45 and a cohesion of zero is less conservative than using a cohesion of 2 and a phi of 33

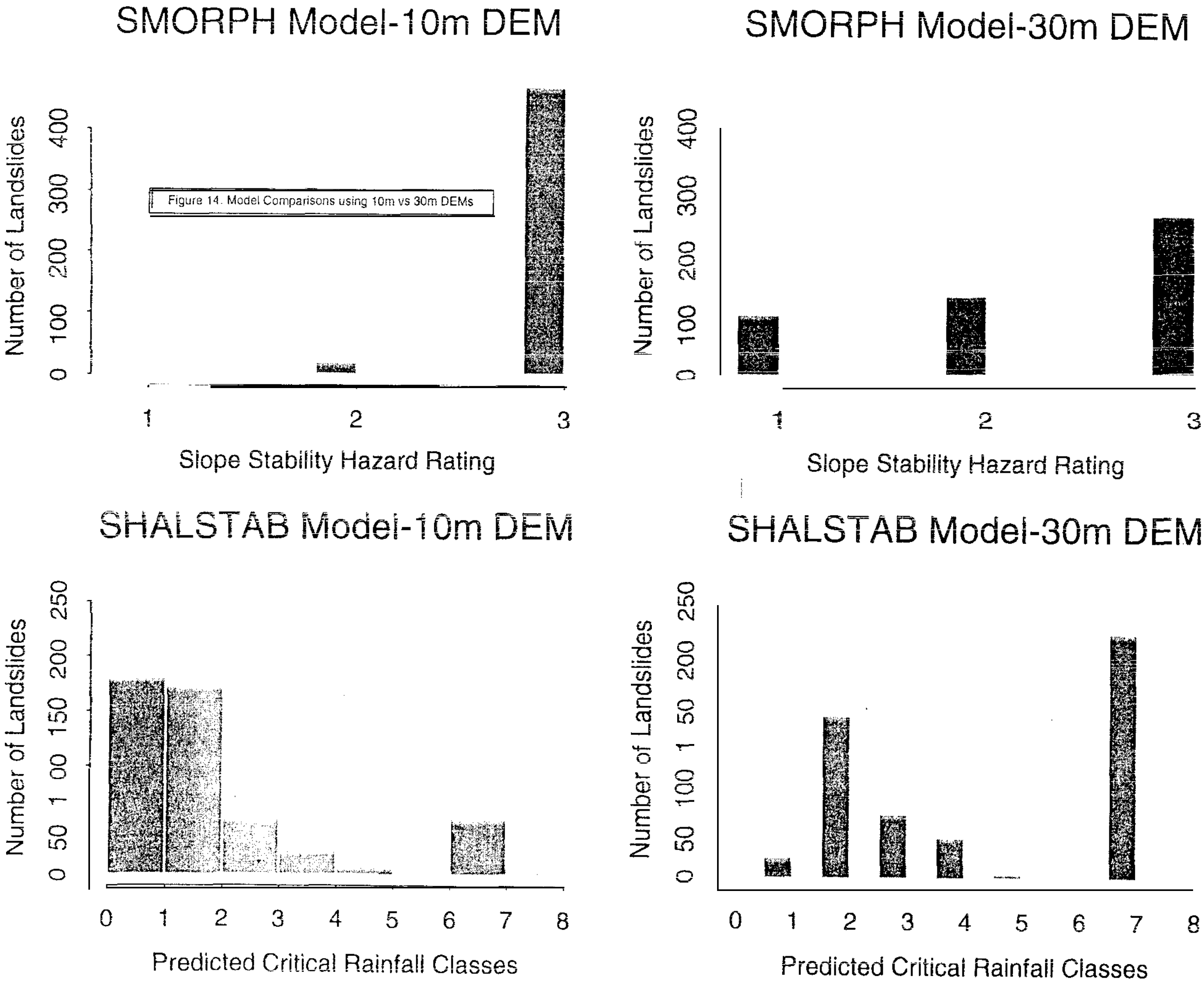


Figure 14. Model Comparisons using 10m vs 30m DEMs