Slope Stability Modeling and Landslide Hazard in Freshwater Creek and Ryan Slough Humboldt County, California

July 2008

Prepared for:
North Coast Regional Water Quality Control Board
5550 Skylane Boulevard Suite A
Santa Rosa, CA 95403

Prepared by:
Eileen Weppner, Professional Geologist #7587
Pacific Watershed Associates
Arcata, CA

Janet Hoyt
Sanborn
Portland, OR

William C Haneberg
Haneberg Geoscience
Seattle, WA

CONTENTS

I. Introduction 1
A. Purpose/Background 1
B. Scope of Work 2
C. Background Information/Reference Materials 3
II. Project Area 3
A. Geologic Setting of the Freshwater Creek and Ryan Slough Planning watershed 5
B. Slope stability 8
III. Methods 10
A. DEM Development 10
B. Shallow Landslide Models 11
1. SHALSTAB 11
2. PISA-m 12
C. Slope Stability Model Comparison 18
1. Model comparison with air photo identified landslide inventory 18
2. Model comparison between the three model outputs 20
IV. Results 20
A. SHALSTAB 21
B. PISA-m 25
C. SMORPH 29
D. Slope Stability Model Performance 33
1. Model performance based on correct prediction of air photo identified landslides 33
2. Model performance based on comparison of air photo identified landslides and randomly generated polygons 39
3. Model performance based comparison between the three models 49
a. Model predictive performance method 1 - model-specific landslide prediction 49
b. Model predictive performance method 2 -- cumulative percentage graphs 50
c. Model predictive performance method 3 - comparison by area of overlap 51
E. Supplemental PISA-m analysis to test the effects of clearcut timber harvest activities 56
V. Summary and Conclusions 60
VI. Limitations of analysis 63
VII. Future analyses and recommendations 65

List of Tables

Table 1. Lithology of Freshwater Creek and the Ryan Slough planning watershed 7
Table 2. Soil units in the Freshwater TMDL study area and their PISA-m parameter values 14
Table 3. PISA-m parameters for the vegetation layer 15
Table 4. SMORPH slope thresholds by lithologic unit. 17
Table 5. SMORPH Calculation Matrix 18
Table 6. SHALSTAB results by instability class and subwatershed 21
Table 7. PISA results by probability class and subwatershed 25
Table 8. SMORPH results by instability class and subwatershed 29
Table 9. Percent watershed area and percent frequency of air photo identified landslides by SHALSTAB instability class and subwatershed 38
Table 10. Percent watershed area and percent frequency of air photo identified landslides by PISA-m probability class and subwatershed 38

Table 11. Percent watershed area and percent frequency of air photo identified landslides by SMORPH instability class and subwatershed ... 39
Table 12. General statistics percent area of highest instability/probability class defining the classification of air photo identified landslides ... 44
Table 13. Air photo identified landslides and model predictive performance ... 50
Table 14. Model overlap of unstable areas by model type and subwatershed ... 54
Table 15. Model overlap of unstable and stable areas by subwatershed ... 55
Table 16. PISA-m model results for second model run assuming <15 year old vegetation age for all forested areas ... 58

List of Figures

Figure 1. Location Map of the Freshwater Creek TMDL Study Area ... 4
Figure 2. Geology Map of Freshwater Creek and the Ryan Slough planning watershed ... 6
Figure 3. Percent watershed area by slope gradient ... 9
Figure 4. Freshwater TMDL 4-meter Digital Elevation Model identifying extent of study area ... 11
Figure 5. Cumulative frequency plot of maximum percent slope within inventoried landslides by lithologic unit ... 17
Figure 6. SHALSTAB results for Freshwater Creek and the Ryan Slough planning watershed ... 22
Figure 7. SHALSTAB instability classes by geologic type, Freshwater Creek watershed ... 23
Figure 8. SHALSTAB instability classes by geologic type, Ryan Slough planning watershed ... 24
Figure 9. PISA-m results for Freshwater Creek and the Ryan Slough planning watershed ... 26
Figure 10. PISA-m probability classes by geologic type, Freshwater Creek ... 27
Figure 11. PISA-m probability classes by geologic type, Ryan Slough planning watershed ... 28
Figure 12. SMORPH results for Freshwater Creek and the Ryan Slough planning watershed ... 30
Figure 13. SMORPH instability classes by geologic type, Freshwater Creek ... 31
Figure 14. SMORPH instability classes by geologic type, Ryan Slough planning watershed ... 32
Figure 15. Air photo identified landslides and SHALSTAB results ... 35
Figure 16. Air photo identified landslides and PISA-m results ... 36
Figure 17 Air photo identified landslides and SMORPH results ... 37
Figure 18. Density of air photo identified landslides and random slides by SHALSTAB instability class, classification based on presence of highest instability class ... 41
Figure 19. Density of air photo identified landslides and random slides by PISA-m instability class, classification based on presence of highest instability class ... 42
Figure 20. Density of air photo identified landslides and randomly generated polygons by SMORPH instability class ... 43
Figure 21. Density of air photo identified landslides and randomly generated polygons by SHALSTAB instability class, classification based on minimum 10% area of highest probability class ... 46
Figure 22. Density of air photo identified landslides and randomly generated polygons by PISA-m probability class, classification based on minimum 12% area of highest probability class ... 47
Figure 23. Density of air photo identified landslides and randomly generated polygons by SMORPH instability class, classification based on minimum 33% area of highest instability class ... 48
Figure 24. Cumulative percentage plots comparing predictive success of the three models ... 52
Figure 25. Model output by instability classification ... 57
Figure 26. Comparison of percent watershed area by PISA-m probability class for 2 PISA-m model runs ... 59

I. Introduction
A. Purpose/Background

Under the Clean Water Act Section 303(d), the Freshwater Creek watershed was listed by the North Coast Regional Water Quality Control Board (Regional Water Board) and the U.S. Environmental Protection Agency (EPA) as a sediment impaired water body. Space Imaging was granted a CMAS contract to conduct the sediment source analysis for the Freshwater Creek Total Maximum Daily Load (TMDL) sediment study. Space Imaging partnered with Pacific Watershed Associates (PWA) and Dr. William C. Haneberg, an expert in landslide modeling, to complete the work elements necessary to conduct the project. Subsequently, Sanborn was spilt from Space Imaging and assumed continued management of the project.

The original work plan submitted by Sanborn for the Freshwater Creek TMDL sediment source assessment included five main tasks. Task 1 involved compiling the historic aerial photography and LiDAR imagery and developing a comprehensive GIS library with the necessary GIS coverages and shape files required to run the slope stability and road surface erosion models.

Task 2 involved conducting a comprehensive landslide inventory using the 1987, 1997 and 2003 historic air photo sets. In addition, Task 2 work elements included a field verification study of a sample of air photo identified landslides to verify landslide type, size, sediment delivery and land use association. Finally, Task 2 included an air photo identified landslide and field inventoried landslide comparison study designed to determine how many landslides were not detected in the air photo analysis due to obstruction by the forest canopy within old-growth, mature second-growth (>30 years old) and "young" forest (≤30 years old) sample plots. Two sample plots for each timber age class category were proposed for the air photo landslide and field identified landslide comparison study.

Task 3 included developing road surface erosion estimates using SEDMODL2. SEDMODL modeling was conducted by Kathy Dube in Freshwater Creek as part of the 2000 PALCO Watershed Analysis. As part of this project, SEDMODL2 was run on areas outside of the Freshwater Creek Watershed Analysis Area (i.e. Fay Slough and Ryan Slough planning watersheds). Road surface erosion estimates developed by Kathy Dube for the Freshwater Creek Watershed Analysis were combined with the road surface erosion estimates calculated as part of this study in order to provide an estimate of total road surface erosion for the entire Freshwater TMDL study area.

Task 4 involved conducting geologic slope stability modeling to identify varying landslide hazard throughout the Freshwater Creek TMDL study area. Slope stability models included: SHALSTAB (Dietrich and Montgomery 1998); PISA-m – Probabilistic Infinite Slope Analysis (Haneberg 2004; 2005); and SMORPH (Shaw and Vaugeois 1999).

Task 5 involved a bank erosion and in-stream wood survey using two published methodologies described in Reid and Dunne (1996) and Benda, et al. (2002). The bank erosion surveys were conducted along three 1,000 meter channel segments in each of three sub-watersheds in the Elk River watershed in both old growth and recently harvested areas.

The Freshwater Creek TMDL sediment source study was conducted in 2 phases. Phase I of the project consisted of Tasks 1, 2 and 3 outlined above. A final report detailing the methodologies and results for these tasks was provided to Sanborn and the North Coast Regional Water Quality Control Board in July 2006.

For Phase II, this report details the methodology and results for Task 4 involving the application of three slope stability models to determine landslide hazard in Freshwater Creek and the Ryan Slough planning watershed. The report outlining the results of the bank erosion and wood inventory conducted as part of Task 5 will be provided as a separate cover and is not included in this document.

B. Scope of Work

The scope of the work conducted as part of the slope stability modeling for Phase II of the Freshwater Creek TMDL Sediment Source Assessment included the following:

- The development of the LiDAR DEM used in the slope stability modeling applications.
- Derive the model variables from existing studies and published information specific to each of the applied models.
- The application of three shallow-rapid deterministic or probabilistic slope stability models including SHALSTAB and SMORPH (deterministic models) and PISA-m (probabilistic model) in order to assist in categorizing the watershed into landslide "hazard" or mass wasting susceptibility categories.
- Outputs from the three models are compared to the air photo identified landslide inventories, and to each other, to identify the degree of correspondence between each model and the air photo identified landslide inventories.
- Determine potential landslide hazard areas based on the model outputs.

The Freshwater Creek TMDL Sediment Source Assessment Phase II Slope Stability Modeling report is organized into 5 sections. The first section provides the introduction to the slope stability report including the purpose, scope of work, background information and reference materials used in the study. Section 2 of the summary report describes the geologic characteristics and tectonic setting of Freshwater Creek and the Ryan Slough planning watershed.

Section 3 describes the methodology used for the slope stability modeling and model verification. Section 4 describes the results for the slope stability modeling and is organized to: 1) describe the results for each of the three slope stability models; 2) describe the model performance in comparison to the air photo identified landslide inventories compiled as part of the Phase I of the Freshwater Creek TMDL sediment source assessment, and 3) to compare the three models to each other. Section 5 provides a summary of the model application results and conclusions and Section 6 describes the confidence in analysis. Finally Section 7 describes recommendations for future analyses and model applications in the Freshwater Creek and the Ryan Slough planning watersheds.

C. Background Information/Reference Materials

Source and reference information for the slope stability modeling included:

- Digital elevation model (DEM) for Freshwater Creek based on LiDAR imagery from Sanborn.
- SHALSTAB hillslope stability model
- PISA-m hillslope stability model
- SMORPH hillslope stability model
- Geology of the Cape Mendocino, Eureka, Garberville, and Southwestern Part of the Hayfork 30 x 60 Minute Quadrangles and Adjacent Offshore Area, Northern California (McLaughlin et al. 2000).
- PALCO data attributes for Freshwater Creek air photo identified landslides
- PALCO GIS data layers.
    - GIS shape file of Freshwater Creek 1987, 1997 and 2003 landslides.
    - GIS shape file of Freshwater Creek Watershed Analysis area sub-basins
    - GIS shape files of 1987, 1997 and 2003 road construction history
- GIS shape files and attribute data for the 1987, 1997 and 2003 air photo identified landslides identified as part of the Phase I Freshwater Creek TMDL sediment source assessment.
- "Landslide Hazard in the Elk River Basin, Humboldt County, California" (Stillwater Sciences 2007). Report referenced for SHALSTAB and PISA-m model variables used in the Elk River slope stability analysis.
- California Department of Forestry timber harvest GIS shape files from 1986-2003
- California Department of Forestry CALVEG (California Vegetation) layer

II. Project Area

The Freshwater Creek TMDL study area is located in Humboldt County in the vicinity of Eureka, California (Figure 1). The study area is comprised of three main watersheds including Freshwater Creek, the Ryan Slough planning watershed, and the Fay Slough planning watershed. The entire TMDL study area is approximately 149.7 km^2 in area. The Freshwater Creek watershed comprises approximately 53% (79.8 km^2) of the TMDL study area. The Ryan Slough planning watershed includes 26% (38.3 km^2) of the TMDL study area, and the Fay Slough planning watershed comprises 21% (31.6 km^2) of the TMDL study area.

Approximately 78% of the Freshwater Creek watershed and the Ryan Slough planning watershed are managed for commercial timber harvesting operations by the Pacific Lumber Company (Freshwater Creek) and Green Diamond Resource Company (Ryan Slough). Land use within the remaining 22% of the area within Freshwater Creek and the Ryan Slough planning watershed, as well as the Fay Slough planning watershed includes urban development, rural subdivision, recreation, and smaller private industry. Further detailed discussion of the land use practices and other watershed characteristics for the Freshwater Creek TMDL study area is provided in the Freshwater Creek TMDL sediment source assessment report submitted in July 2006 (PWA 2006).

The slope stability modeling study was conducted for Freshwater Creek and the Ryan Slough planning watershed. The Fay Slough planning watershed was not included in the slope stability

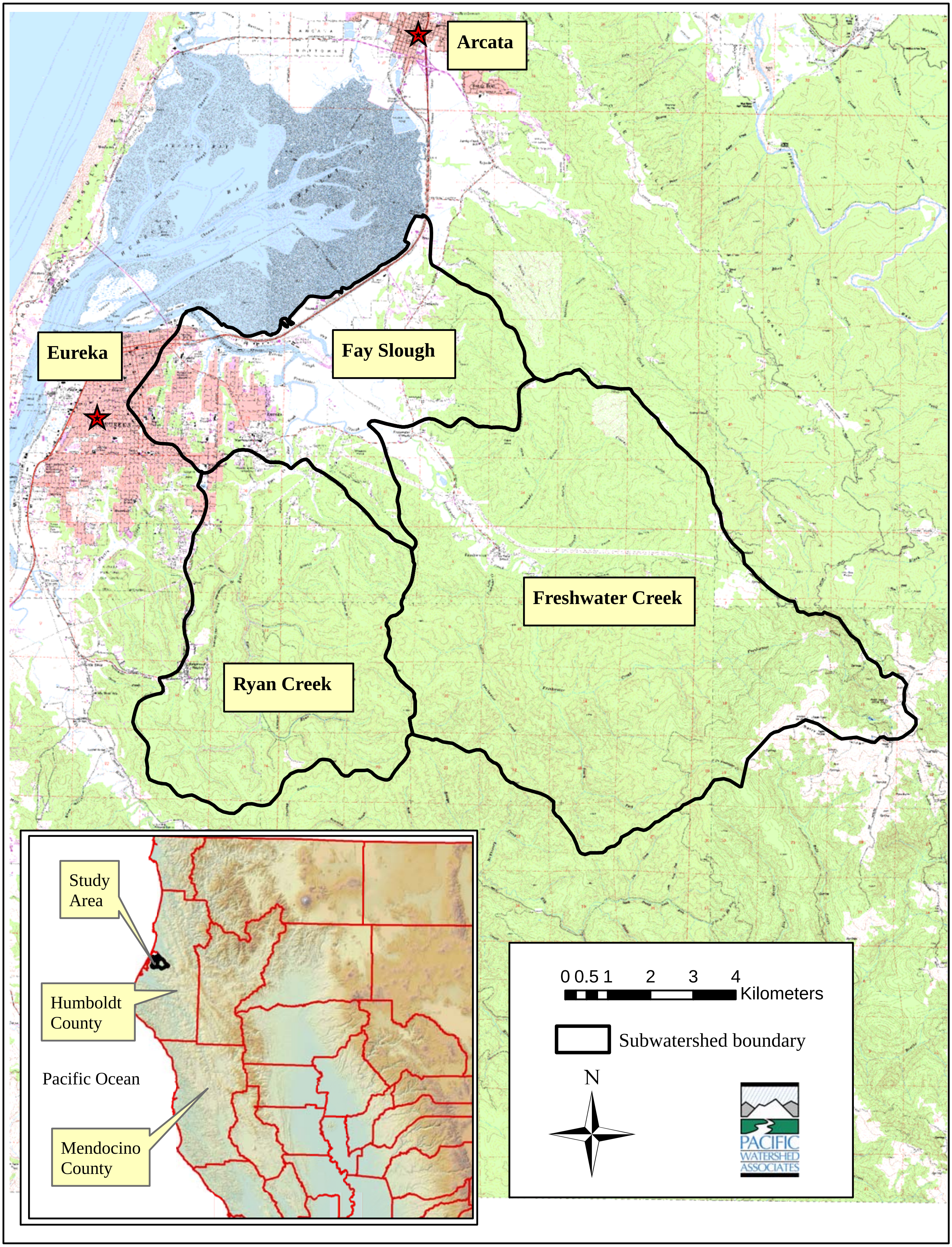


Figure 1. Freshwater Creek TMDL study area

analysis due to incomplete LIDAR coverage and the low occurrence of landsliding observed during the air photo analysis phase of the Freshwater Creek TMDL study.

A. Geologic Setting of the Freshwater Creek and Ryan Slough Planning watershed
Coastal California north of Cape Mendocino lies within the tectonically active convergent margin of the North American plate. Since the Mesozoic Era, the geologic development of Northern California has been dominated by plate convergence. During the last 140 million years, subduction and the resulting continental accretion have welded a broad complex of highly deformed oceanic rocks to the western margin of the North American plate. These accreted rocks now comprise the Franciscan Complex, which constitutes the basement of the North Coast region (Carver and Burke 1992). Throughout the latest geologic period, major uplift of the Coast Ranges and erosional stripping of the regionally extensive forearc sediments are postulated to have resulted from the combined effects of the eastward subduction of the Gorda plate and the northward migration of the Mendocino triple junction (Nilsen and Clarke 1987). Today, geologically youthful cover sediments are preserved in a series of structural settings such as those found within and around the greater Humboldt Bay region (Clarke et al. 1992; Nilsen et al. 1987; Carver 1987).

The distribution of lithologic units within Freshwater Creek and the Ryan Slough planning watershed is illustrated in Table 1 and Figure 2. The lithology for the study area was compiled from the Geology of the Cape Mendocino, Eureka, Garberville, and Southwestern Part of the Hayfork 30 x 60 Minute Quadrangles and Adjacent Offshore Area, Northern California developed by McLaughlin et al. (2000). Full discussion of the underlying lithology for the entire Freshwater Creek TMDL study area is provided in the Freshwater Creek TMDL Phase 1 report submitted in July 2006 (PWA 2006).

The Freshwater Creek and Ryan Slough planning watershed are comprised of the following lithologies:

Franciscan Complex –Central Belt and Coastal Belt units of the Franciscan Complex are found in the study area (Table 1 and Figure 2).

> **KJfm** – The Franciscan mélange consists of a matrix of clayey, penetratively sheared meta-argillite and blocks of metasandstone. Thirty percent (30%) of the slope stability study area is underlain by this unit, which exhibits rounded, poorly incised, lumpy and irregular topography (McLaughlin et al. 2000). The Franciscan mélange is found primarily in the eastern half of the Freshwater Creek watershed. The Ryan Slough planning watershed is not underlain by the Franciscan mélange.
>
> **KJbf –** The Broken Formation consists of bedded to massive, locally folded, rarely conglomeratic, meta-sandstone and meta-argillite. The Broken Formation typically exhibits sharp crested topography with regular well incised side hill drainages (McLaughlin et al. 2000). Less than 1% of the slope stability study area is underlain by the Broken Formation, and these rocks crop out along the eastern side of the Freshwater Creek Fault in the Freshwater Creek watershed.

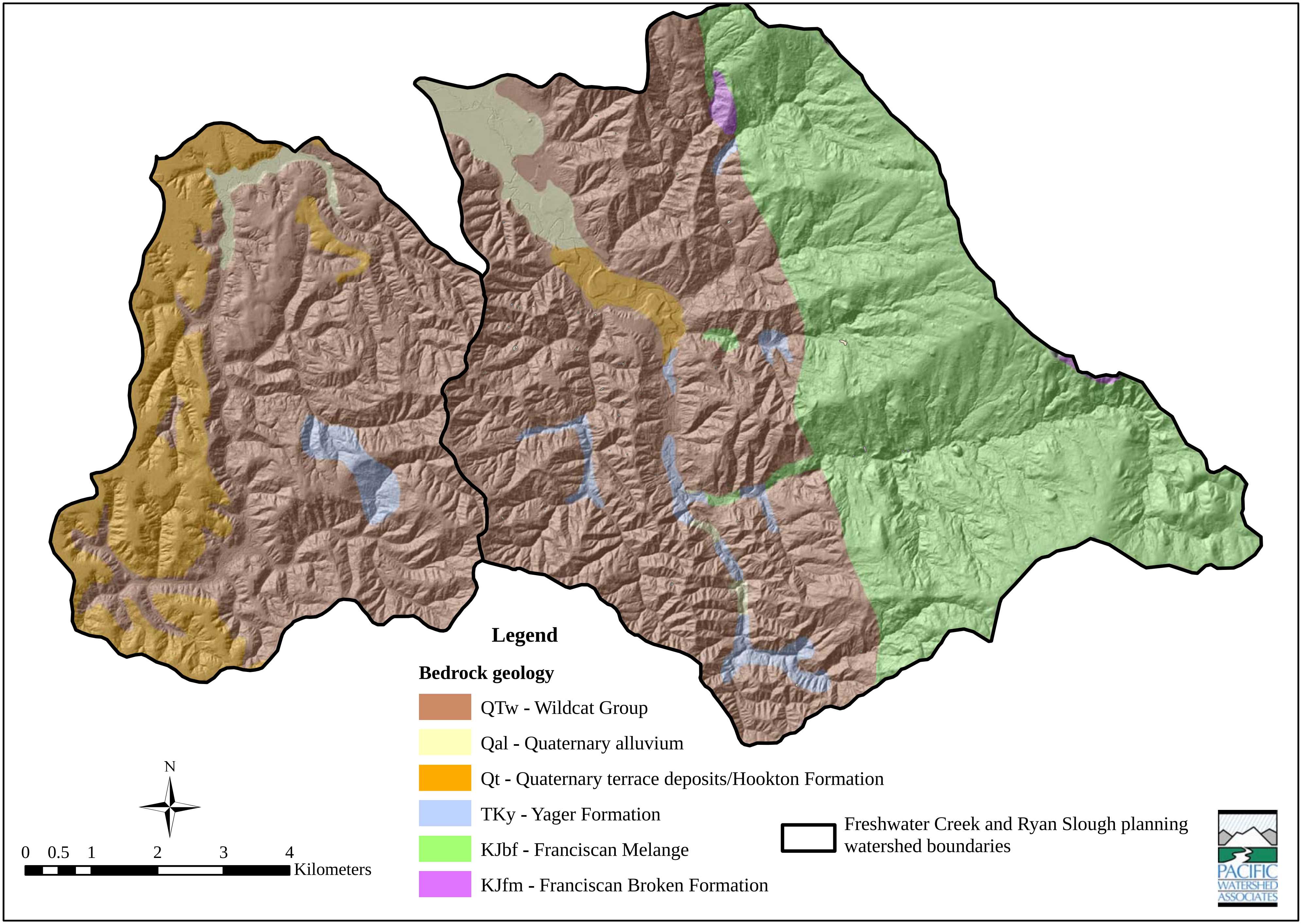


Figure 2. Bedrock geology in Freshwater Creek and the Ryan Slough planning watershed

Coastal belt – The Coastal Belt underlies the majority of the slope stability study area, comprising approximately fifty-seven percent (57%) of the study area (Table 1 and Figure 2). The Coastal Belt is divided into two structural and lithologic terranes.

TKy –The Yager terrane consists of thin to medium bedded argillite and arkosic sandstone and massive to thickly bedded arkosic sandstone with minor interbeds of argillite. The Yager terrane is divided into 3 subunits based principally on topographic expression: “y1” sheared and highly folded mudstone with irregular topography; “y2” highly folded broken mudstone, sandstone, and conglomeratic sandstone with sharp ridge crests and well incised drainages; and “y3” highly folded, broken sandstone, conglomerate, and mudstone with sharp ridge crests and well incised drainages. Only the “y1” subunit is located in the slope stability study area and represents approximately 3% of the study area. The “y1” subunit is found primarily in the incised stream channels in Freshwater Creek and locally in stream channel bottoms of the Ryan Slough planning watershed (Table 1 and Figure 2).

QTw – The undifferentiated Wildcat Group consists predominantly of weakly to moderately well lithified marine sandstone, siltstone, mudstone, and minor conglomerate. Fifty-four percent (54%) of the slope stability study area is underlain by this geologic unit (Table 1 and Figure 2. Approximately 67% of the Ryan Slough planning watershed and 48% of the Freshwater Creek are underlain by the undifferentiated Wildcat Group.

Table 1. Lithology of Freshwater Creek and the Ryan Slough planning watershed

Lithology	Ryan Slough PW		Freshwater Creek		Total	
	Area (km^2)	% area	Area (km^2)	% area	Area (km^2)	% area
Qal	0.78	2%	2.59	3%	3.37	3%
Qt	10.88	28%	1.29	2%	12.17	10%
QTw	25.64	67%	38.59	48%	64.23	54%
TKy	1.04	3%	2.07	3%	3.11	3%
KJfm	0.0	0%	34.96	44%	34.96	30%
KJbf	0.0	0%	0.26	<1%	0.26	<1%
Total	**38.34**	**100%**	**79.76**	**100%**	**118.1**	**100%**

Quaternary Terrace Deposits (Qt) – Terrace deposits consist of Holocene and Late Pleistocene undifferentiated non-marine terrace deposits. Specifically, Qt consists of dissected and uplifted gravel, sand, silt and clay deposited in fluvial settings. In addition, Qt consists of minor shallow intertongues and warped tilted beds of Late Pleistocene Hookton Formation (McLaughlin et al. 2000). Approximately 10% of the slope stability study area is underlain by Qt. In the Ryan Slough planning watershed, Qt primarily consists of the Hookton Formation and represents 28% of the planning watershed area (Table 1 and Figure 2).

Quaternary Alluvial Deposits (Qal) -Alluvial deposits are found along the low elevation areas of the main stem reaches of Freshwater Creek and the Ryan Slough planning watershed. They comprise approximately 3% of the slope stability study area (Table 1 and Figure 2). These deposits consist of Holocene clay, silt, sand, gravel and boulders deposited in stream beds, alluvial fans, terraces, floodplains and ponds (McLaughlin et al. 2000).

B. Slope stability

Many of the Coast Range watersheds of Northern California comprise some of the most unstable terrain in the Pacific Northwest. The aim of a slope stability analysis is to identify areas of unstable terrain and potential slope instability. General factors that may influence slope stability include:

1) Geologic structure and discontinuities: dip of beds (dipping down slope nearly parallel to or less than the slope inclination); fractures; weak beds inter-bedded with competent beds; faults; shear zones; and surfaces of weakness are all contributing factors present on the North Coast and each lends to decreased slope stability. Geologic structure, bedding orientation and weathering profiles all appear to influence the location of slope failures in the Wildcat Group of rocks.
2) Material strength: decomposition of materials over time reduces strength; chemical weathering; diminished rooting strength of vegetation following vegetation removal (by fire or harvesting) is important in shallow soils overlying bedrock; deep colluvium deposits located on steep slopes (especially in steep swales) are factors lending to inherently unstable conditions. Likewise, road and landing construction can change the distribution of mass in steep swales and on potentially unstable hillslopes.
3) Increased pore water pressure: elevated pore pressures can reduce resisting forces and increase driving forces to further reduce slope stability on steep colluvial slopes. A variety of factors affecting precipitation input (intensity duration, and antecedent soil moisture content), infiltration, changes in interception rates, soil water movement rates, and seepage rates can have an influence on slope stability.
4) Weather: intense rainfall events are common on the North Coast, and intense and prolonged precipitation events are known to be associated with increased rates of landsliding, especially when high antecedent pore water pressures or soil moisture content exist (Harden et al. 1995; Kelsey et al. 1995).
5) Slope angle:
    i) Geologic materials have characteristic slopes at which they are stable and just barely stable: residual soils - 30 to 40 degrees; colluvium - 20 to 30 degrees (Hunt 1994). This slope gradient is often termed the "angle of repose" and is typically applied to low density, non-cohesive soils. Such an analysis produces overly conservative values for natural slopes that contain denser soils with cohesive strength and root reinforcement.
    ii) Slope inclination is locally increased by road cuts and sidecasting, as well as by processes associated with stream bank erosion, and the redistribution of soils during landsliding events.
    iii) Tectonic uplift rates can cause both subtle and rapid increases in slope over time.
    iv) Streams incise in response to lowering base levels, producing steep slopes and steep inner gorges along major drainages.

6) Seismic activity: seismicity, especially the proximity to major earthquake sources, can be an important factor affecting slope stability and the occurrence of landslides (Wieczorek 1996, Keefer 1984). Earthquakes may also have an effect on the movement of subsurface water through deep colluvial soils, such as those found locally in Freshwater Creek and the Ryan Slough planning watershed, by cutting off macro-pores and increasing pore water pressures in potentially unstable, steep hillslope areas. The effects of large earthquakes, such as those of 1991-1992, are most likely to express themselves during subsequent high magnitude storms following severe seismic shaking (e.g. the winters of 1995, 1996 and 1997).

Steep slopes are the most common locations for landslides in Freshwater Creek and the Ryan Slough planning watershed. Ten percent (10%) of the Ryan Slough and 12% of Freshwater Creek planning watersheds are composed of slopes exceeding slope gradients of 65% (33°), while approximately 25% of the Ryan Slough and 28% of Freshwater Creek planning watersheds exhibit slope gradients exceeding 50% (26.6°) (Figure 3). Slope gradients over 50% (26.6°), and especially those over 65% (33°), are most common along portions of the inner gorge of the Freshwater Creek and Ryan Slough planning watersheds, and their major tributaries. Steep inner gorge slopes are generally not continuously present along the major stream channels. Elsewhere, steep slopes are most commonly found in the headwall areas of steep swales developed in the Wildcat Group (e.g. Little Freshwater Creek). Importantly, 46% of the Ryan Slough and 41% of the Freshwater Creek planning watersheds exhibit slope gradients of less than 30% (16.7°), and at a generally low risk of landsliding. In addition, approximately 74% of the Ryan Slough and 72% of Freshwater Creek planning watersheds exhibit of slope gradients less than 50% (26.6°) (Figure 3). Refer to the Phase I Freshwater Creek sediment source assessment report for a detailed analysis of the magnitude and frequency of landsliding according to slope location (PWA 2006).

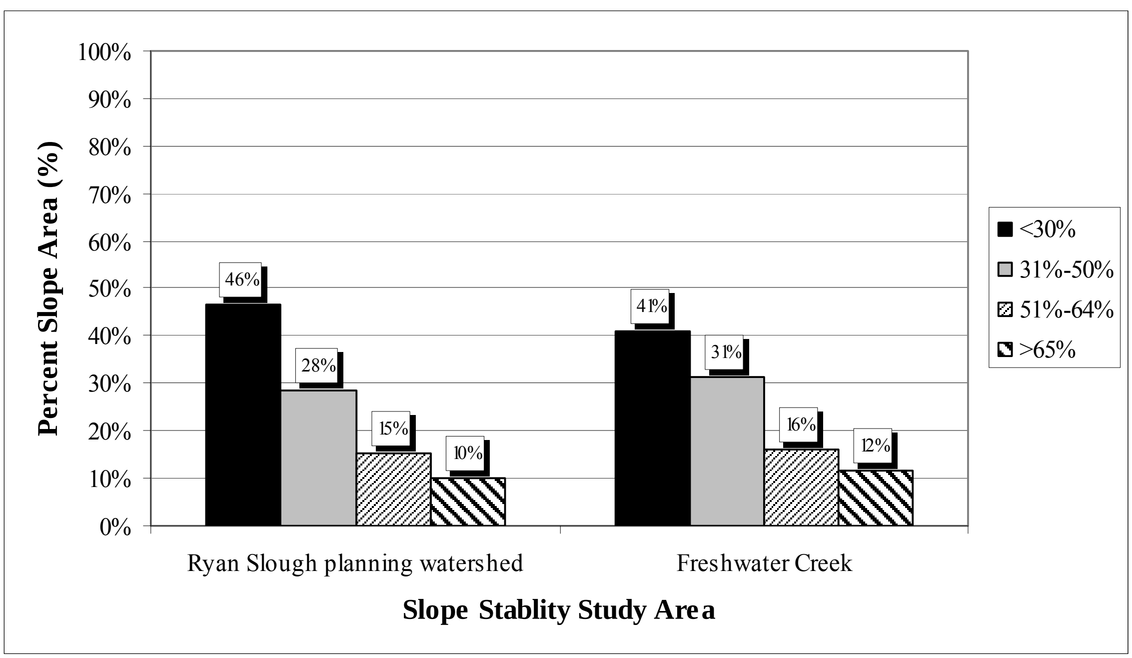


Figure 3. Percent watershed area by slope gradient, Freshwater Creek TMDL slope stability study area

III. Methods

As part of the development of a comprehensive landslide hazard map for the Freshwater Creek and Ryan Slough planning watershed, three shallow slope stability models (SHALSTAB, PISA-m, and SMORPH) were used to identify the potential for landslide hazard in the two subwatersheds. The slope stability models used in this project were determined during a kickoff meeting via phone with the North Coast Regional Water Quality Control Board (NCRWQCB), Dr. Haneberg, Sanborn, and PWA. Lessons learned from the application of slope stability models as part of Elk River TMDL sediment source assessment played an important role in the model selection. The PISA-m and SHALSTAB models were determined to be appropriate for the Freshwater Creek TMDL sediment source assessment in part due to their simplicity to run. The SMORPH model was included because of its geometric and parametric simplicity.

Each slope stability model required specific input layers with data files describing various geotechnical parameters related to the input layers. The common base data layer for all models was the digital elevation model (DEM) for the area.

A. DEM Development

The DEM used for this project was produced from a LiDAR (Light Detection and Ranging) dataset collected in March 2005 for the North Coast Regional Water Quality Control Board, in order to support the analyses conducted as part of the Elk River and Freshwater Creek total maximum daily load studies. The LiDAR dataset, acquired by Space Imaging, was processed to obtain first and last return point data. The last return data were further filtered to yield a LiDAR surface representing the bare earth. The filtered bare earth last return data were used to make a point shape file of elevation values. This in turn served as the input to the interpolation process that created a regularly spaced grid of elevation values, the 1-meter DEM[1].

Based on findings from the previous Elk River TMDL landslide hazard study, the detailed 1-m DEM was known to contain real short-wavelength topography, such as fallen logs, stumps, or brushy areas, as well as some topographic “noise” or random errors. In addition, the 1-m DEM caused the SHALSTAB modeling program to crash (Stillwater Sciences 2007). That study determined that a 4-m grid was optimal for modeling hillslope stability using the SHALSTAB model within the Elk River Watershed. In order to be consistent with this methodology so as to allow for future comparisons between the two studies, a 4-m DEM was also used in the Freshwater TMDL project.

For the Freshwater Creek TMDL slope stability study, the 1-meter DEM was re-sampled to a 4-meter DEM using the bilinear interpolation re-sampling routine available in ArcMap. This technique is commonly used to smooth continuous data and provides a weighted average of the four nearest input cells, adjusted to account for their distance from the center of the output cell. The resulting 4-meter DEM is shown in Figure 4.

---

[1] For more detailed information on the LiDAR collect and processing, refer to: Freshwater Creek Watershed and Elk River Watershed Tributaries of Humboldt Bay, California, LIDAR Campaign Final Report, Sanborn, March 2005.

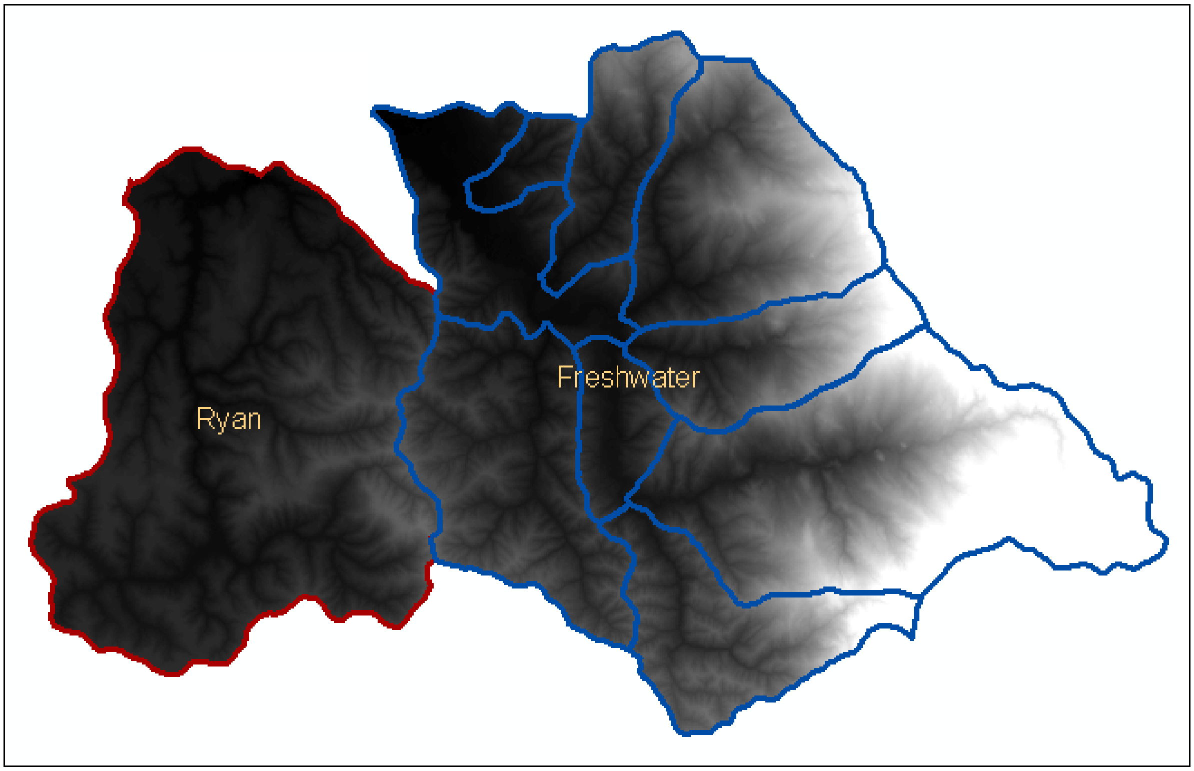


Figure 4: Freshwater TMDL 4-meter Digital Elevation Model identifying extent of study area

B. Shallow Landslide Models

1. SHALSTAB

SHALSTAB is a GIS-based deterministic model that runs in the ArcView environment (Dietrich et al. 1998). The SHALSTAB method is based on a coupled hydrologic-slope stability equation that solves for the hydrologic ratio (q/T):

$$\frac{q}{T} = \frac{\rho_s}{\rho_w}\left(1 - \frac{\tan\theta}{\tan\phi}\right)\frac{b}{a}\sin\theta$$

where

q = effective precipitation
T = vertical integral of saturated conductivity
ρ_s = soil bulk density
ρ_w = water bulk density
$\sin\theta$ = head gradient
$\tan\phi$ = angle of internal friction of the soil mass at the failure plane
a = drainage area
b = width of outflow boundary

The model assumes that slope angle and slope curvature are the main factors driving slope stability. The model also incorporates soil physical and hydrological properties as mathematical constants by including friction angle and cohesion value, but not tree root strength or seismicity.

The only input layer the SHALSTAB model requires is a depressionless DEM. In addition, it requires parameters for the angle of internal friction and the soil bulk density, which should be the approximate saturated bulk density. The SHALSTAB manual strongly recommends using the defaults of 45 degrees and 1700 kg/m^3 unless there are data to suggest otherwise. For the Freshwater TMDL, we retained the defaults.

The SHALSTAB program itself provided some technical difficulties during the project. The Fill Sinks routine, one of the first tasks in creating the depressionless DEM, did not function on the input DEM, and the Freshwater TMDL study area was too large to be processed as a whole. Fortunately, solutions for both of these issues were developed: the former by manually filling sinks using a series of steps in ArcMap; the latter by processing the data in three sub-sections with each subsection containing an overlap area to ensure no edge effects in the final output dataset. In addition, the final analysis steps within SHALSTAB would not open the landslide inventory shape files. Again, this issue was not critical as the SHALSTAB model results could be output to GRID and assessed manually. Sanborn would like to acknowledge the assistance of the author of the SHALSTAB user interface, Rafael Real de Asua, in working to resolve these issues.

The output from the SHALSTAB model identifies the ratio of steady-state effective precipitation (rain minus evapotranspiration) to transmissivity (the ground's subsurface ability to convey the water downslope), or the q/T ratio (Dietrich et al. 1998). For the Freshwater TMDL, this ratio was calculated using the slope and area grids calculated from the depressionless DEM within SHALSTAB. The lower the q/T ratio the more unstable the land. It should be noted that the transmissivity used in SHALSTAB assumes a completely saturated soil, which is not the case in most places. As such, it will tend to over predict the capability of the slope to discharge water that would otherwise increase pore water pressure and decrease slope stability.

Because the SHALSTAB model works with mathematical constants for the geographic area analyzed, the level of accuracy or resolution of the slope stability results may be reduced. This is one of the reasons for the inclusion of the PISA-m and SMORPH models in this study.

2. PISA-m

PISA-m is a physically based probabilistic slope stability model that utilizes a first-order, second-moment (FOSM) formulation of the same infinite slope stability equation used by the US Forest Service LISA and DLISA programs. (Haneberg 2005; 2004):

$$FS = \frac{c_r + c_s + \left[q_1 + \gamma_m D + \left(\gamma_{sat} - \gamma_w - \lambda_m\right) H_w D\right] \cos^2 \beta \tan \phi}{\left[q_t + \gamma_m D + \left(\gamma_{sat} - \gamma_m\right) H_w D\right] \sin \beta \cos \beta}$$

where

FS = factor of safety
c_r = cohesive strength contributed by tree roots
c_s = cohesive strength of soil
q_t = uniform surcharge due to weight of vegetation
γ_m = unit weight of moist soil above phreatic surface
γ_w = unit weight of water (9810 N/m^3 of 62.4 lb/ft^2)
D = thickness of soil above slip surface
H_w = height of phreatic surface above slip surface, normalized relative to soil thickness
β = slope angle
ϕ = angle of internal friction

As discussed by Haneberg (2000), physically based (rational) probabilistic models are based on the underlying physics of geologic processes but incorporate elements of probability theory to account for the natural variability or uncertainty of input variables like soil shear strength. PISA-m utilizes topographic attributes and a variety of geotechnical random variables including pore water pressures and tree root shear strength. These random variables are treated as continuous probability distributions so the results can include the mean and variance of the factor of safety against sliding, probability of sliding, or non-parametric slope reliability. The inclusion of tree root strength parameters allows PISA-m to be used for evaluation of land management options such as logging. Although it was not used in this study, PISA-m also includes seismic capabilities based on a probabilistic implementation of Jibson's simplified Newmark method. The PISA-m model requires two input layers in addition to the DEM: a soils unit layer and a forest cover type map. In addition, a parameter file identifying the geotechnical characteristics of each map unit in the soils and vegetation layers is required.

The soils unit layer was derived from the six geology types that were delineated in the Freshwater TMDL study area[2]. The geotechnical parameters used in the PISA-m model for each of these geology types are identified in Table 2. Because the Quaternary alluvium layer is generally flat and in the valley bottoms, the only type of landslides one might expect to occur in it are bank failures or slumps along incised streams. Because PISA-m was not designed to predict this type of landslide, the geotechnical parameters for this layer were set to zero. All coefficients used β-PERT distributions except for the dimensionless pore water pressure coefficient, for which an extreme value distribution was used to simulate peak annual values. The probabilities calculated by PISA-m therefore have a temporal significance because of this extreme value distribution. β-PERT distributions— which were originally developed to analyze complicated processes in which random variables could be quantified in terms of optimistic, most likely, and pessimistic values— are well suited to geologic applications in which numerical estimates of variables are tempered by qualitative professional experience. They are three parameter versions of standard four-parameter β distributions.

[2] The geology was derived from McLaughlin, R.J., S.D. Ellen, M.C. Blake, Jr., A.S. Jayko, W.P. Irwin, K.R. Aalto, G.A. Carver, and S.H. Clarke, Jr. 2000. Geology of the Cape Mendocino, Eureka, Garberville, and Southwestern part of the Hayfork 30 x 60 Minute Quadrangles and Adjacent Offshore Area, Northern California. Miscellaneous Field Studies MF-2336, Version 1.0.

Table 2. Soil units in the Freshwater TMDL study area and their PISA-m parameter values.

Geology Type	Symbol		b-PERT distributions					Extreme value distribution
			Unit Weight of soil, dry gm (lbs/ft3)	Unit Weight of Soil Saturated gsat (lbs/ft3)	Angle of Internal Friction (degrees)	Soil Cohesion Cs (lbs/ft^2)	Soil Thickness (ft)	Pore Pressure Coefficient
Quaternary Alluvium	Qal	Min	0	0	0	0	0	0
		Likely	0	0	0	0	0	0
		Max	0	0	0	0	0	0
Wildcat Group	QTw	Min	85	115	30	100	1	0.5
		Likely	100	125	32	160	5	0.1
		Max	125	142	34	300	13	0
Quaternary Non-Marine Terrace Deposits (Hookton)	Qt	Min	100	125	31	0	1	0.5
		Likely	115	135	33	50	5	0.1
		Max	130	145	35	230	13	0
Formation Franciscan mélange	KJfm	Min	75	110	18	100	1	0.5
		Likely	85	115	25	200	5	0.1
		Max	100	125	32	460	13	0
Franciscan Broken Formation	KJbf	Min	75	110	18	100	1	0.5
		Likely	85	115	25	200	5	0.1
		Max	100	125	32	460	13	0
Yager	TKy	Min	75	110	28	0	1	0.5
		Likely	85	115	31	140	5	0.1
		Max	100	125	35	280	13	0

The vegetation layer for this project was developed using the vegetation classes from the CALVEG[3] (CDF 2001) dataset coupled with stand age attributes compiled by PWA using the Fire and Resource Assessment Program's (FRAP) Timber Harvest Plans (THP). The CALVEG dataset was first recoded to Forested or Non-forested and then divided into two age groups, less than or equal to 15 years and greater than 15 years, with the intent of capturing the varying root strengths between younger and older forests. The 15 year break was chosen because past studies have shown that reduced root reinforcement from harvesting occurs within the first 10 - 15 years for Douglas-fir (Pseudo-tsuga menziesii) and pine forests (Roering et al. 2003; Abe et al. 1991; Krogstad 1995; Sidle 1992). Redwoods and hardwoods tend to have less reduced root strength because they re-sprout after cutting (Reid 1998). The first 15 years is thought to be a critical time for landslides triggered by harvesting activity.

Two special cases should be noted. First, there were a few instances where the THP data and the CALVEG data did not agree. That is, the CALVEG data indicated an area as non-forested while the THP identified the area as having a stand age. In these cases, the THP was assumed to be the most accurate record, and the area was relabeled as forested and given a stand age. Second, water was treated as "NODATA" in the PISA-m model because any area covered by a large amount of water is almost certainly flat and therefore has essentially no potential for shallow landsliding. PISA-m will not calculate a result for any raster in which any of the three input maps (DEM, soils, or forest cover) has a "NODATA" value. The geotechnical parameters used in the PISA-m model for the vegetation layer are shown in Table 3.

Table 3. PISA-m parameters for the vegetation layer.

Vegetation Type		**Root Cohesive Strength Cr (psf)**	**Tree Surcharge Q (psf)**
Non-Forested	Min	0	0
	Likely	0	0
	Max	0	0
Trees > 15 years	Min	90	5
	Likely	100	25
	Max	525	40
Trees < 15 years	Min	0	0
	Likely	50	5
	Max	100	10

[3] **CALVEG** is a hierarchical classification system of actual vegetation designed to assess vegetation-related resources throughout California. It was developed by the U.S.F.S Region 5 Remote Sensing Lab.

At the request of the NCRWQCB, the PISA-m model was run a second time in order to determine the effects of clear cut timber harvesting in the Freshwater Creek and Ryan Slough planning watersheds. This was achieved by using one vegetation age of "<15 years" and the corresponding PISA-m parameters for root cohesive strength and tree surcharge for all forested areas in the 2 subwatersheds. A vegetation age of "<15 years" is assumed to represent recent clear cut timber harvest conditions. The results of the second analysis are provided in Section IV-E of this report.

3. SMORPH

SMORPH is a semi-empirical method developed by the Washington Department of Natural Resources to identify steep and concave portions of watersheds susceptible to shallow landslides that may mobilize into debris flows. The method is computationally simple and therefore well suited to raster GIS implementation. More importantly, SMORPH has been shown to produce results that are as good as, if not better than, SHALSTAB (Shaw and Vaugeois 1999), and it is less computationally intensive.

Along with the DEM, SMORPH requires slope and curvature thresholds as input to define the surfaces with regard to landslide potential. These thresholds must be characterized for each lithologic unit within the study area that responds differently to mass wasting. For the Freshwater TMDL, four unique units were identified the Wildcat Group (QTw), the Hookton Formation (Qt), the Franciscan Central Belt (KJfm/KJbf) and the Yager Formation (TKy).

For SMORPH, slope thresholds needed to be calculated for five slope ranges: Relatively Flat, Low Steepness, Moderate Steepness, Very Steep and Extremely Steep. To calibrate slope thresholds, a slope layer created from the 4-meter DEM was intersected with the landslide inventory data to determine the maximum slope at which each landslide occurred. This was computed using the ZONALMAJORITY feature in ArcMap. For each lithologic unit, these maximum slopes were ordered and then plotted against the cumulative frequency of these maximum slopes (Figure 5). As per the SMORPH literature (Shaw and Vaugeois 1999), the lower range for the Moderate Steepness Slope was identified as the slope at which 15% of the landslides occurred. The lower values for the next two slope thresholds, Very Steep and Extremely Steep, are then computed by stepping up at 10% intervals.

Again per the SMORPH calibration recommendations, the minimum threshold for the Low Steepness slope range is identified as the minimum slope at which a landslide occurs. This rule was used for both the Hookton Formation and Wildcat Group units. However, for the Franciscan and Yager units, the minimum slope was equal to the slope at 15% cumulative frequency. For these units, the slope threshold for Low Steepness was reduced an additional 5 percent. The final slope threshold ranges used for the Freshwater TMDL are identified in Table 4.

Once the slope thresholds were determined, values for curvature were tested for all lithologic units. For each unit, the model was run using both the lowest SMORPH default curvature range ($\pm$0.01) and the highest SMORPH default curvature range ($\pm$ 0.5). SMORPH uses a non-conventional definition of curvature, so it is difficult to rigorously assess the topographic significance of the numbers. However, little to no difference was visualized in the model outputs. Using conventional definitions, a curvature value of ± 0.5 would be a very broad definition of "planar". As the model outputs indicate, the results are relatively insensitive to curvature so we opted to use the more narrow definition of "planar" as ± 0.01. The final SMORPH results were obtained by running the model separately for each lithologic unit and then combining the results. The SMORPH results are presented in three classes: Stable, Caution, and Unstable with the SMORPH Calculation Method identified in the matrix below (Table 5).

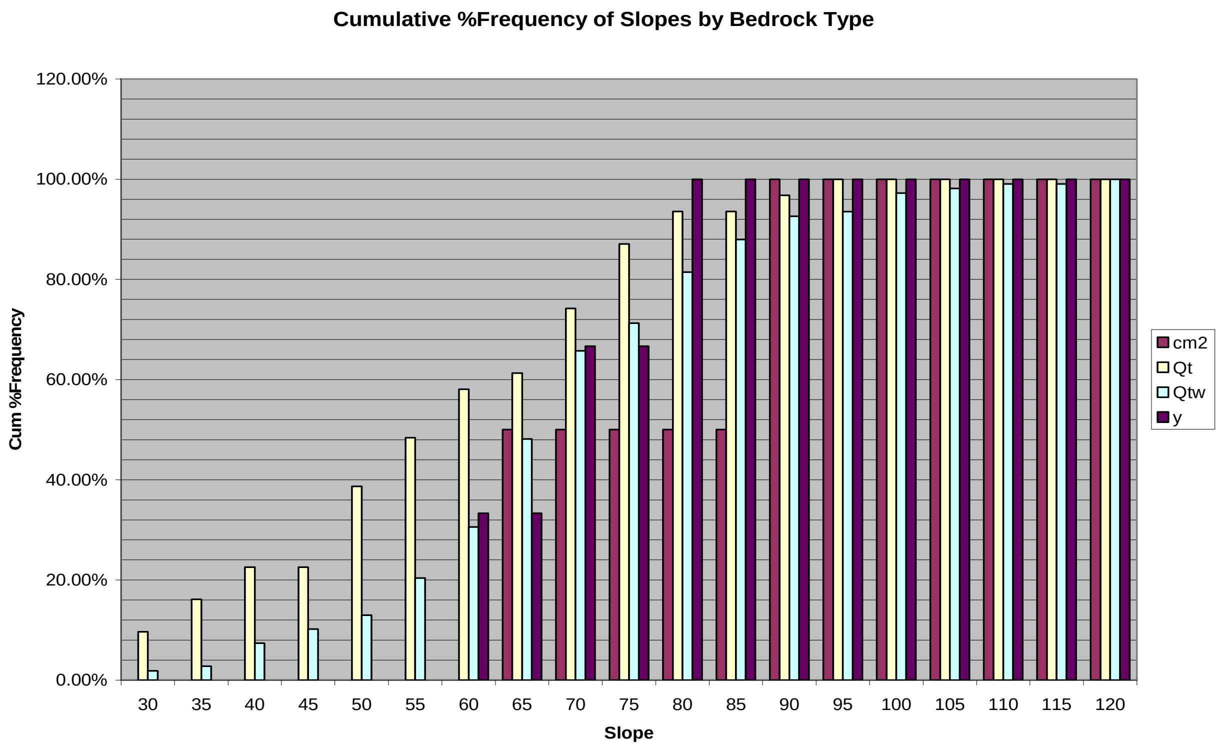


Figure 5. Cumulative frequency plot of maximum percent slope within inventoried landslides by lithologic unit.

Table 4: SMORPH slope thresholds by lithologic unit.[1]

Lithology	Relatively Flat	Low Steepness	Moderate Steepness	Very Steep	Extremely Steep
Wildcat Group (QTw)	$0 < m \leq 30$	$30 < m \leq 55$	$55 < m \leq 65$	$65 < m \leq 75$	$75 < m$
Hookton Formation (Qt)	$0 < m \leq 30$	$30 < m \leq 35$	$35 < m \leq 45$	$45 < m \leq 55$	$55 < m$
Franciscan Central Belt KJfm/KJbf)	$0 < m \leq 32$	$32 < m \leq 65$	$65 < m \leq 75$	$75 < m \leq 85$	$85 < m$
Yager Formation (TKy)	$0 < m \leq 30$	$30 < m \leq 60$	$60 < m \leq 70$	$70 < m \leq 80$	$80 < m$

[1] *m* – minimum threshold

Table 5. SMORPH Calculation Matrix

Curvature Class	Slope Class				
	Relatively Flat	Low Steepness	Moderate Steepness	Very Steep	Extremely Steep
Concave (CV < -0.01)	Stable	Caution	Unstable	Unstable	Unstable
Planar (-0.01 < CV < 0.01)	Stable	Stable	Stable	Caution	Unstable
Convex (CV > 0.01)	Stable	Stable	Stable	Stable	Caution

C. Slope Stability Model Comparison

Slope stability model performance was assessed in two ways. First, the model outputs were compared with the air photo identified landslides compiled as part of the Phase I Freshwater Creek TMDL sediment source assessment. Specifically, model performance was determined by identifying the degree of spatial correspondence between each model and the air photo identified landslides. This was accomplished from more of a heuristic approach than a rigorous statistical procedure since the air photo identified landslide analysis was inherently incomplete, focused within a short time period (air photo identified landslides were identified from the 1987, 1997 and 2003 aerial photography), and limited to the conditions that existed shortly before they were collected. Thus, a model is not necessarily wrong if it predicts a landslide in a place where none has been mapped. Landslides can be missed in the air photo inventory process; for example an old landslide may be difficult to identify on the historic aerial photos, or can be obscured by vegetation. Alternatively, the model may correctly predict a landslide that has not yet occurred and therefore not shown on an inventory. Second, the three models were compared to each other in order to determine whether they indicate potential for landsliding in similar rating classes.

1. Model comparison with air photo identified landslide inventory

A historic air photo analysis of shallow landsliding was conducted as part of the Phase I Freshwater Creek TMDL sediment source assessment (PWA 2006). The air photo identified landslide inventory included shallow landslides with sediment delivery to streams that were observed on the 1987, 1997, and 2003 aerial photography. The air photo landslide inventories originated from two sources: (1) air photo identified landslides observed in Freshwater Creek were compiled from existing studies including the Mass Wasting Component of the 2001 Freshwater Creek Watershed Analysis (1987 and 1997 air photo identified landslides) (PALCO 2001), and the PALCO 2003 forensic landslide study (2003 air photo identified landslides) (PALCO 2003); and (2) air photo identified landslides observed in the Ryan Slough planning watershed were inventoried by PWA on the 1987, 1997 and 2003 historic aerial photography as part of the Phase I Freshwater Creek sediment source assessment (PWA 2006). Both air photo landslide inventories employed the same general methods for landslide identification and attribution. Detailed description of the air photo landslide analysis methods and specific landslide attributes collected is provided in the Phase I Freshwater Creek sediment source assessment report (PWA 2006).

Landslide polygons for air photo identified landslides in the Freshwater Creek and the Ryan Slough planning watersheds were “heads-up”[4] digitized in ArcMAP using the air photo mylars, landslide width and length attributes, and the 1-m LiDAR imagery. This method differs from that used in Elk River where landslide polygons were not available (Stillwater Sciences 2007). Because landslide polygons are available for Freshwater, and there is concern over the ability to accurately locate the true point of initiation for each landslide, the polygon comparison was determined to be a better approach. This will, however, only allow indirect comparisons with the Elk River slope stability modeling results.

In order to examine the performance of the three models, a random sample of circular polygons was developed for Freshwater Creek (N=383) and the Ryan Slough planning watershed (N=381). The random sample population for each planning watershed was developed from a random sample generator using a 95% confidence level and 5% confidence interval. Random sample points were generated using the GIS “Hawth’s Analysis Tools” extension in ArcMap. Using the same GIS extension, circular polygons were generated from the randomly generated points with a 15 m buffer. A 15 m buffer was chosen because it corresponded with the average area of the air photo identified landslides observed in the Freshwater Creek and Ryan Slough planning watershed. Randomly generated polygons were then overlain on the model outputs to determine whether the models predict higher instability at air photo identified landslide locations as compared to random locations on the landscape.

The air photo identified landslide polygons and randomly generated polygons were then compared to the model outputs by overlaying each polygon with the outputs from the models to determine whether the models predicted unstable slopes in the polygon area. The landslide stability/probability class for the air photo landslides and randomly generated polygons was determined by the presence of the highest instability/probability class within the landslide/randomly generated polygon. The presence of the highest instability/probability class regardless of the area is a common approach used in previous SHALSTAB and SMORPH analyses, and was employed to remain consistent with the methodology employed in the Elk River slope stability study (Stillwater Sciences 2007).

To examine each model’s performance at correctly identifying air photo identified landslide occurrence in unstable areas, landslide density graphs were developed comparing densities of air photo identified landslides and randomly generated polygons per total area of subwatershed by each stability/probability class. Each air photo identified landslide and each randomly generated landslide-sized polygon was evaluated as to which instability class (SMORPH; SHALSTAB) or probability class (PISA) the polygon was assigned, according to the rules of assignment.

A slope stability model is considered to perform relatively poorly if the landslide density graphs show little variability in instability/probability classification between air photo identified landslide polygons and randomly generated polygons. That would suggest that the slope stability model cannot differentiate between the actual landslides and the randomly generated polygons. A model is considered to perform well (or better) if the landslide density graph exhibits a

[4] “Heads-up” digitizing refers to the manual transference of mapped data from air photo mylars or base maps directly to the computer screen. This method does not use a spatially referenced digitizing board.

divergence in the density plots, indicating increased differentiation (identification) of the air photo identified landslides compared to the density of randomly generated polygons. That is, where the plots show little variance between the distribution of air photo identified landslides and randomly generated polygons for the instability/probability classes, the models are judged not good at predicting landslide locations and where the plots visibly diverge, the model is considered to be a relatively better predictor of landslide location. The greater the divergence of the plots, the better the predictive performance is judged to be.

To further evaluate model performance, air photo identified landslides and randomly generated polygons were classified by instability/probability class using a more rigorous classification method. The alternative classification method involved establishing a minimum instability/probability classification threshold based on a larger percent area of instability/probability class within the air photo landslides and randomly generated polygons. That is, the polygon or landslide must contain a designated percentage of a certain hazard class for it to be classified in that instability class. This might provide more realistic accuracy than a classification based only on the presence of a single “pixel” of the highest probability class. This additional analysis was conducted to determine whether instability/probability classification based on a larger percent area of instability/ probability class within the sample populations would result in better model performance, as compared to classification based on strictly presence/absence of the highest instability/probability class. Landslide density graphs were re-plotted for each of the three models, and analyzed to evaluate model performance.

2. Model comparison between the three model outputs

In addition to comparing the model outputs to actual air photo landslide inventories, the results of the three models were also compared to each other. First, the three models were analyzed to determine the accuracy of correct air photo identified landslide prediction by comparing the frequency of correctly predicted landslides with the frequency of landslides that were not predicted by the slope stability models. Second, cumulative percentage graphs plotting the percentage of the study subwatershed occupied by model hazard class and the percentage of landslide cells captured within the model hazard class. Cumulative percentage plots were developed to measure the predictive success of the three models. Third, the watershed areas were analyzed to determine the degree of overlap between the three models; that is, whether zero, one, two or all three of the models indicate an elevated potential for landsliding.

IV. Results

Three physically based slope stability models were employed in the analysis of slope stability in the Freshwater Creek and Ryan Slough planning watershed: the deterministic SHALSTAB and SMORPH slope stability models and the probabilistic slope stability model PISA-m. The following results discuss the spatial distribution instability/probability classes generated by each of the three models individually for Freshwater Creek and the Ryan Slough planning watershed, as well as the evaluation of model performance. The two subwatersheds were analyzed separately due to the differences in the type and distribution of the underlying lithologic types. In addition, the data sources for each watershed were compiled from different sources (PWA analysis of Ryan Creek, and PALCO analysis of Freshwater Creek). It was determined that separate analysis of the subwatersheds would be consistent with the previous analyses.

A. SHALSTAB

Figure 6 illustrates the results from the SHALSTAB analysis. The SHALSTAB results are provided as range instability classes in units of - log (*q/T*). The SHALSTAB instability classes were broken down by the following ranges: (1) -10 (chronic or unstable), (2) -9.99 to -3.1 (high), (3) -3.1 to -2.8 (moderate high), (4) -2.8 to -2.5 (moderate), (5)-2.5 to -2.2 (moderate low), (6) >-2.2 (low), and (7) 10 (stable). These ranges are based on results derived from previous studies (Dietrich et al. 2001; Montgomery et al. 1998). Approximately 92% of the entire slope stability modeling study area (Freshwater Creek and Ryan Slough planning watershed) were classified as stable (10 –log(*q/T*)), low (>2 –log(*q/T*)), and moderate low (-2.5 to -2.2 –log(*q/T*)) (Table 6).

Table 6. SHALSTAB results by instability class and subwatershed.

Subwatershed		SHALSTAB instability class (-log(*q/T*))							Total
		-10 (Unstable)	<-3.1	-3.1 - -2.8	-2.8 - -2.5	-2.5 - -2.2	>-2.2	10 (Stable)	
Freshwater Creek	Area (km^2)	0.3	1.2	1.5	3.2	6.4	19.7	47.0	79.4
	%	0.4%	1.6%	1.9%	4.1%	8.1%	24.8%	59.2%	100.0%
Ryan Slough planning watershed	Area (km^2)	0.1	0.5	0.6	1.3	2.9	9.0	23.7	38.1
	%	0.3%	1.2%	1.6%	3.5%	7.5%	23.7%	62.2%	100.0%
Total	Area (km^2)	0.4	1.7	2.1	4.6	9.3	28.7	70.6	117.4
	%	0.4%	1.5%	1.8%	3.9%	7.9%	24.4%	60.1%	100.0%

Approximately 60% of the entire slope stability modeling area study area was classified as stable. Overall, Freshwater Creek and Ryan Creek displayed nearly equivalent percent areas in each of the SHALSTAB instability classes (Table 6). Conversely, areas classified as chronically unstable (-10 –log(*q/T*)), high instability (<3.1 –log(*q/T*)), and moderate high instability (-3.1 to -2.8 –log(*q/T*)) only represent approximately 4% of the entire slope stability modeling study area. Figure 7 and Figure 8 illustrate the distribution of SHALSTAB instability classes by each subwatershed (Freshwater Creek and Ryan Slough planning watershed) and underlying lithology. In Freshwater Creek, the Wildcat Group (QTw) and Franciscan mélange (KJfm) make up nearly 92% of the underlying lithology (48% and 43.9%, respectively) (Figure 7). The majority of the area within the Wildcat Group and Franciscan mélange was classified as stable or low instability (81% and 87%, respectively). Only 4% of the Wildcat Group and Franciscan mélange was classified as chronically unstable to moderate high instability.

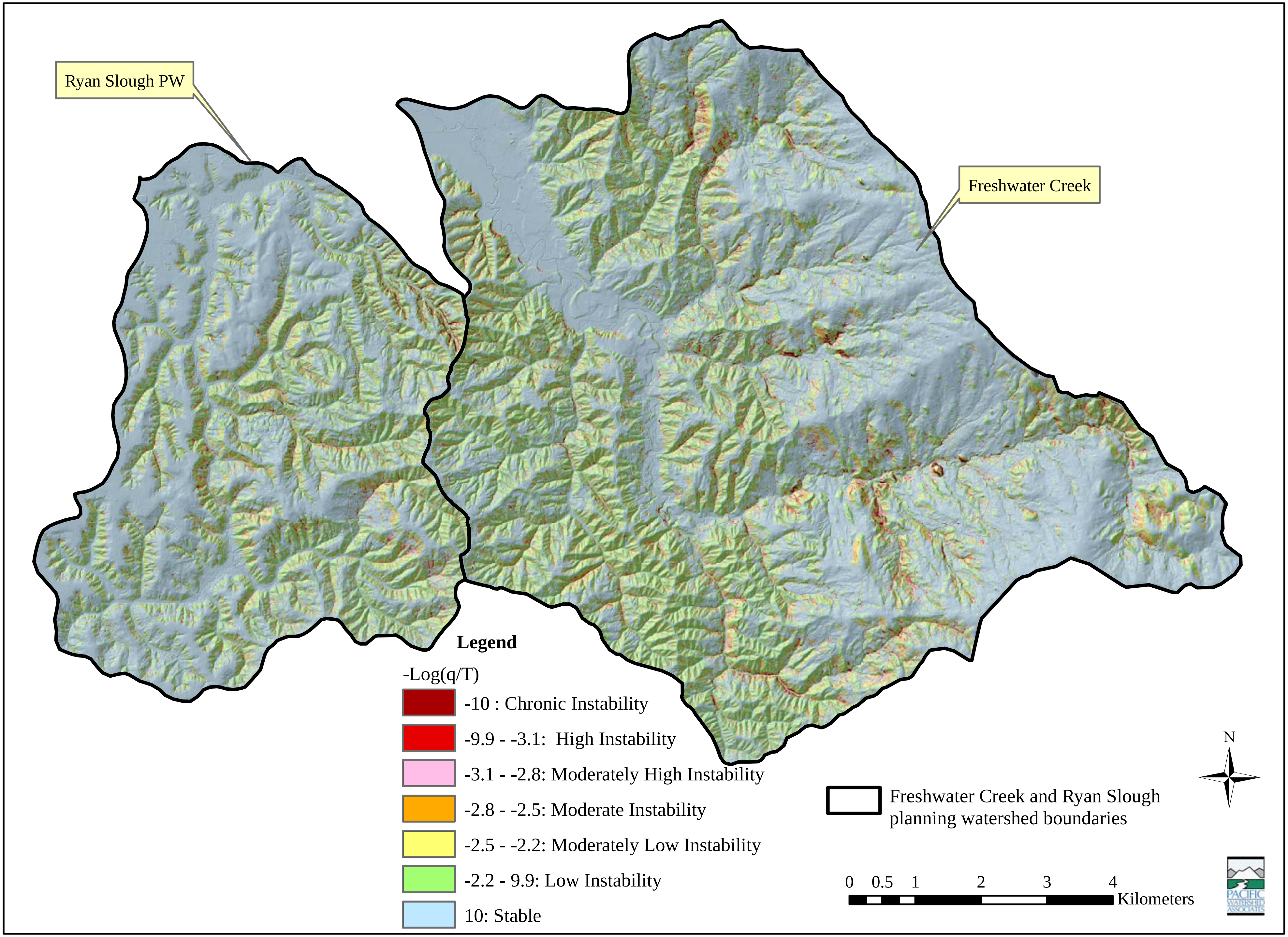


Figure 6. SHASTAB results for Freshwater Creek and the Ryan Slough planning watershed

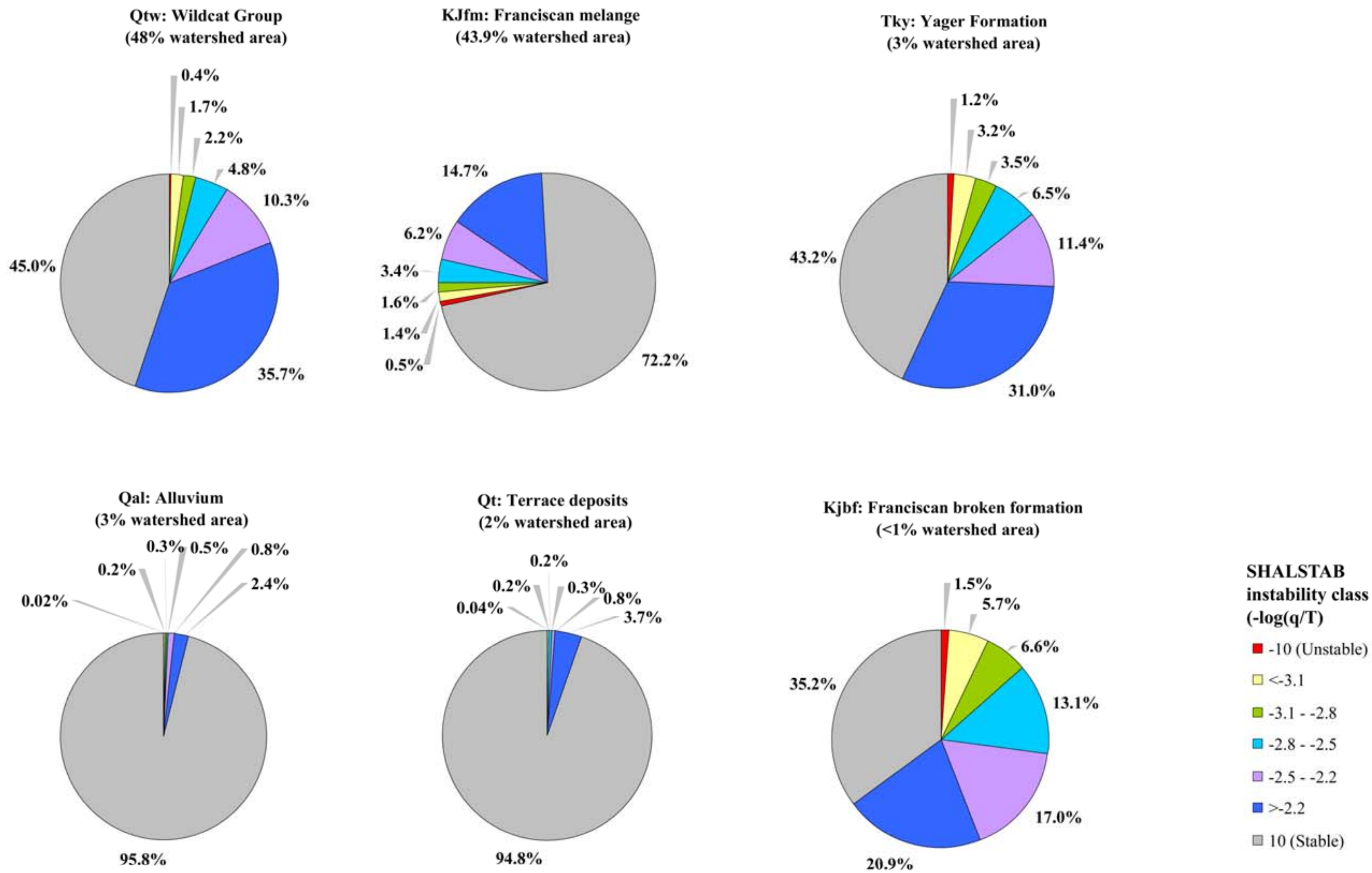


Figure 7. SHALSTAB instability classes by geologic type, Freshwater Creek watershed

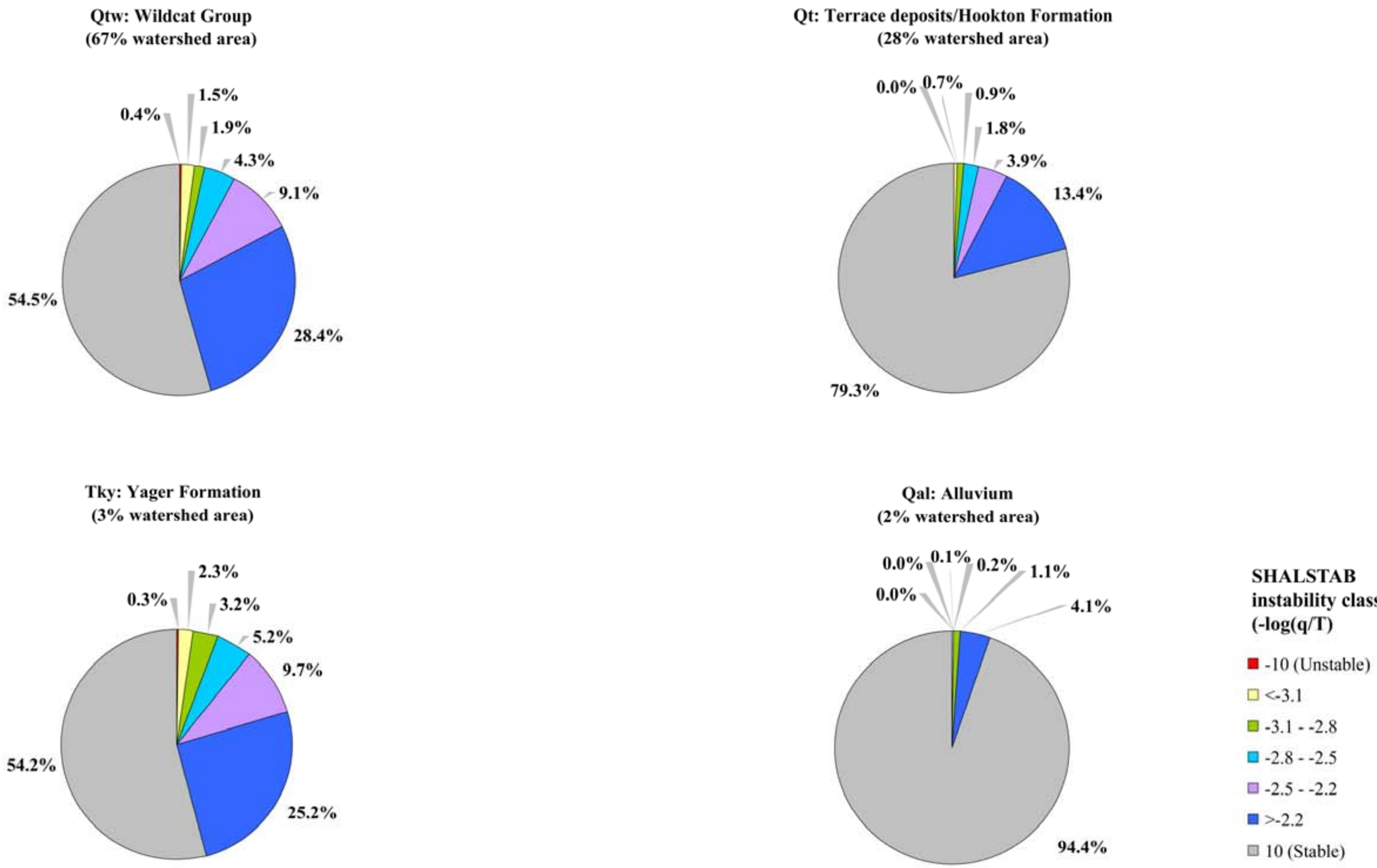


Figure 8. SHALSTAB instability classes by geologic type, Ryan Slough planning watershed

The Ryan Slough planning watershed is underlain primarily by the Wildcat Group (67%) and Qt (28%: mainly the Hookton Formation) (Figure 8). According to the SHALSTAB model outputs, approximately 83% of the Wildcat Group and 93% of the Hookton Formation underlying the Ryan Slough planning watershed was classified as stable or low instability. Nearly 4% of the subwatershed area underlain by the Wildcat Group and Hookton Formation was classified as chronically unstable to a moderate high instability (Figure 8).

B. PISA-m

Figure 9 and Table 7 illustrate the PISA-m model outputs as six slope instability probability classes. The PISA-m slope probability classes range from 0 to 0.02, 0.02 to 0.05, 0.05 to 0.065, 0.066 to 0.1, 0.1 to 0.2, and >0.2. The PISA-m ranges were delineated to best illustrate the distribution of instability probability values and landslide recurrence intervals. According to the PISA-m model outputs, approximately 82% of the entire slope stability study area was classified in the lowest probability class (0 to 0.02 or greater than 50 year landslide recurrence interval). Furthermore, approximately 90% of the slope stability modeling study area was classified with less than a 5% chance of landsliding (p= 0 to 0.05, or greater than 20 year landslide recurrence interval) (Table 7).

Table 7. PISA results by probability class and subwatershed.

Sub watershed		PISA probability class (p)						Total
		0 – 0.02 (Low probability >50 yr recur.)	0.02 – 0.05 (20-50 yr recur.)	0.05 – 0.065 (15-20 year recur.)	0.066-0.1 (10-15 year recur.)	0.1-0.2 (5-10 year recur.)	>0.2 (High probability <5 year recur.)	
Freshwater Creek	Area (km^2)	59.2	6.0	1.6	2.4	2.7	1.3	73.2
	%	81%	8%	2%	3%	4%	2%	100.0%
Ryan Slough planning watershed	Area (km^2)	27.2	2.1	0.5284	0.7	0.8	0.5	32.0
	%	85%	6%	2%	2%	3%	2%	100.0%
Total	**Area (km^2)**	**86.4**	**8.1**	**2.2**	**3.1**	**3.5**	**1.9**	**105.1**
	%	**82%**	**8%**	**2%**	**3%**	**3%**	**2%**	**100.0%**

Figure 10 and Figure 11 illustrate the distribution of PISA-m instability probability classes by each subwatershed (Freshwater Creek and Ryan Slough planning watershed) and underlying lithology. In Freshwater Creek, the majority (78%) of the area underlain by the Wildcat Group was classified with the lowest instability probability rating (p≤0.02 or greater than 50 year landslide recurrence interval) (Figure 10). Eighty-five percent (85%) of the area in Freshwater Creek underlain by the Franciscan mélange was similarly classified with the lowest instability

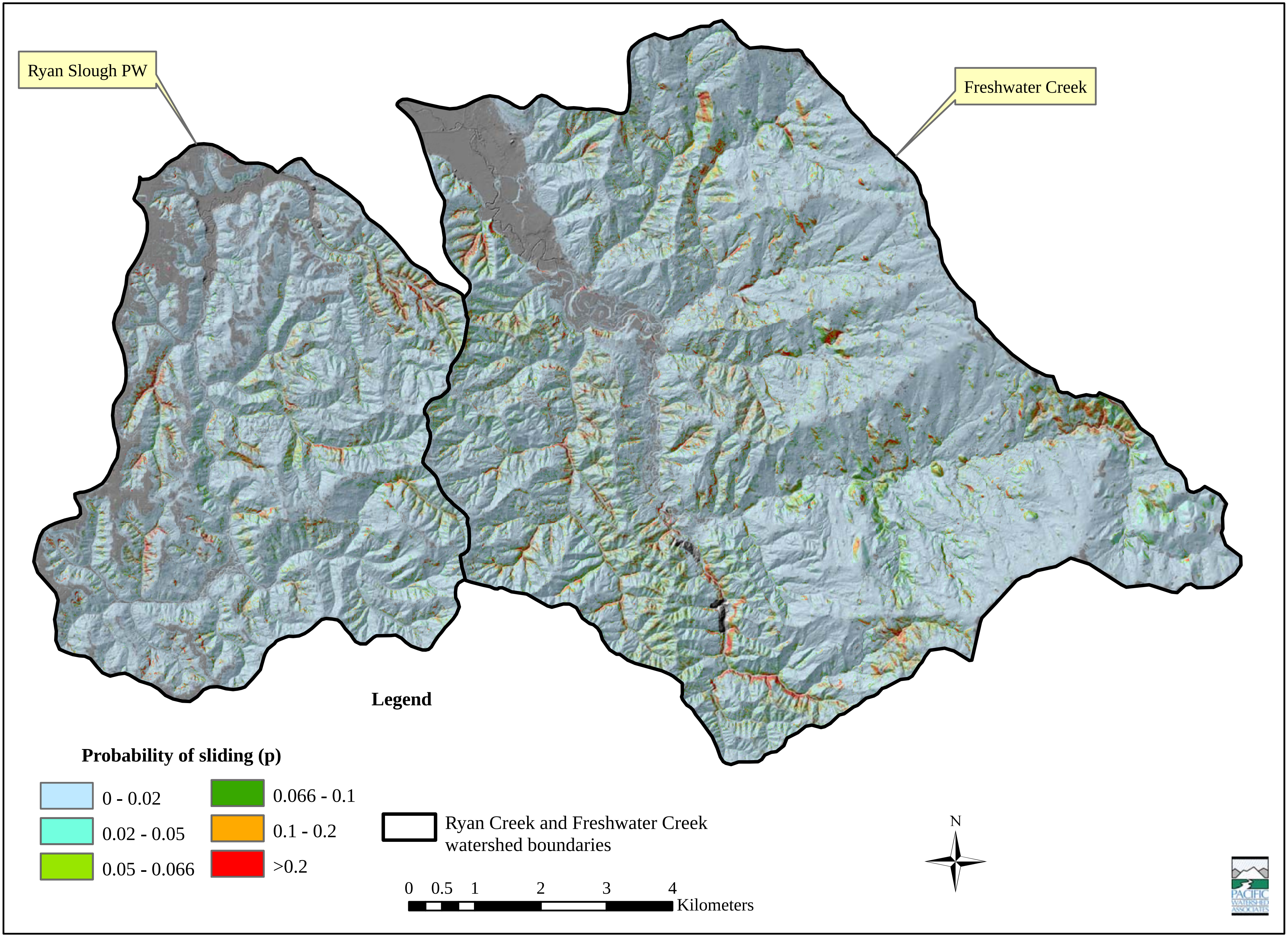


Figure 9. PISA-m results for Freshwater Creek and the Ryan Slough planning watershed

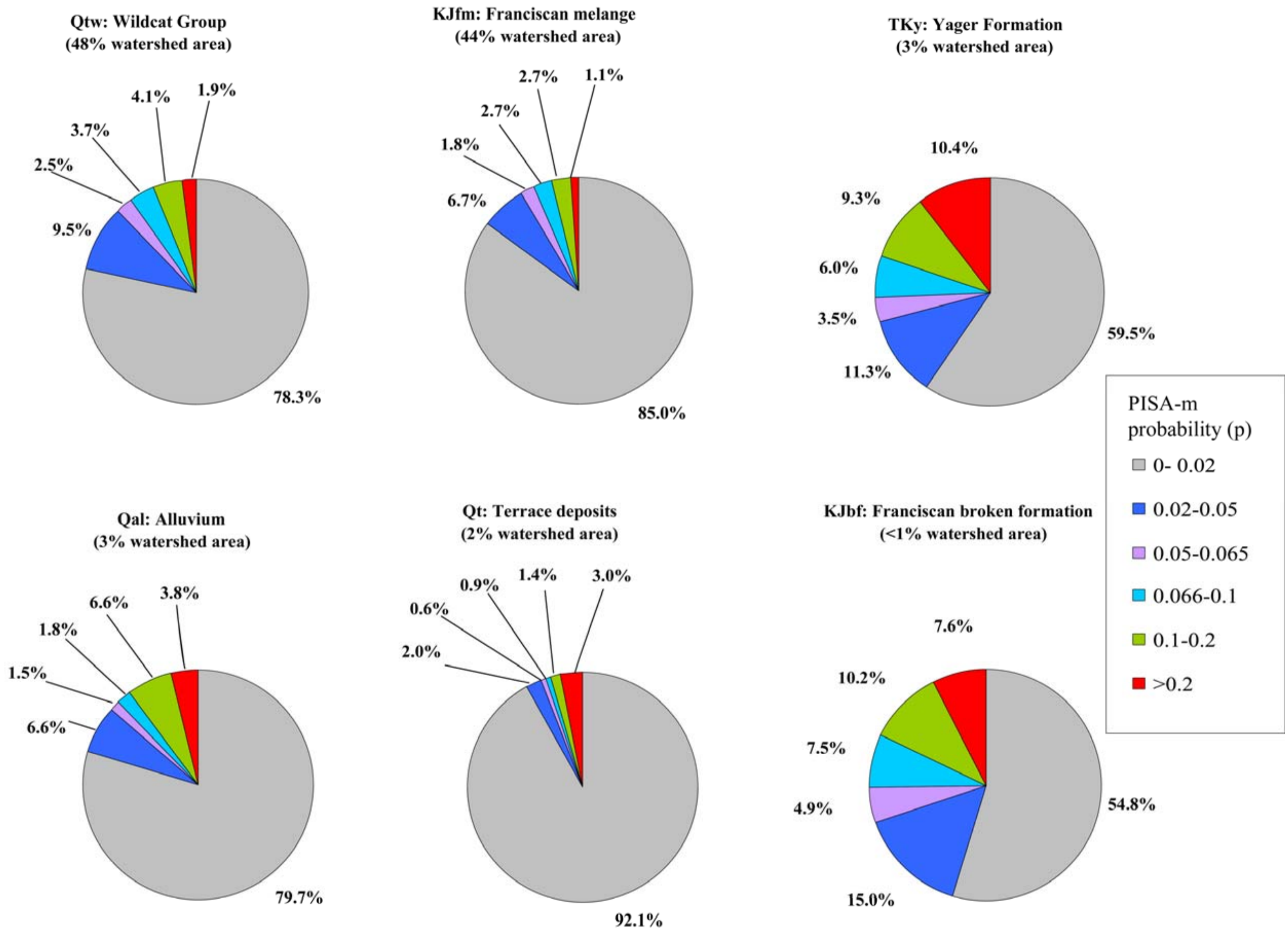


Figure 10. PISA-m probability classes by geologic type, Freshwater Creek

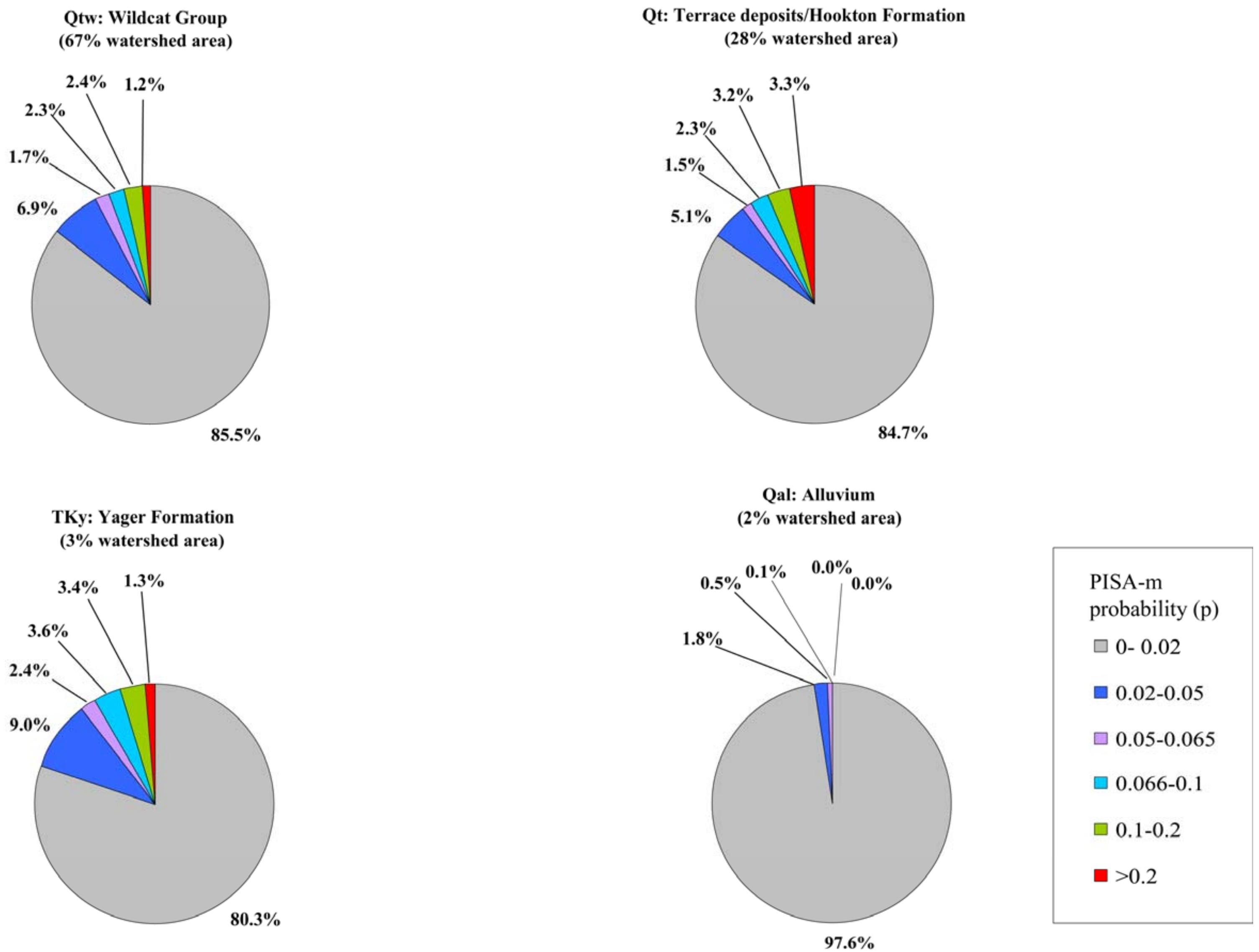


Figure 11. PISA-m probability classes by geologic type, Ryan Slough planning watershed

probability (p≤0.02 or greater than 50 year landslide recurrence interval). Approximately 6% of the Wildcat Group and 4% of the Franciscan mélange slopes in the Freshwater Creek watershed were classified in the least stable category with probability values greater than 0.1 or less than a 10 year landslide recurrence interval (Figure 10).

According to the PISA-m model outputs for the Ryan Slough planning watershed, approximately 86% of the Wildcat Group and 85% of the Hookton Formation was classified with the lowest instability probabilities (p≤0.02). Similar to Freshwater Creek, less than 8% of the Ryan Slough planning watershed area underlain by the Wildcat Group and Hookton Formation was classified in the least stable category probability values greater than 0.1 or less than a 10 year landslide recurrence interval (Figure 11).

C. SMORPH

Figure 12 and Table 8 illustrate the distribution of SMORPH instability classes for the slope stability study area, including Freshwater Creek and the Ryan Slough planning watershed. SMORPH instability classes are delineated according to model design (Shaw and Vaugeois 1999) as “Stable”, “Caution”, and “Unstable” (e.g. low, moderate, and high rating). According to the SMORPH model outputs for the Freshwater Creek slope stability modeling study, approximately 76% of the study area was classified with a low instability rating, 15% was classified with a moderate instability rating, and 9% was classified with a high instability rating (Table 8).

Table 8. SMORPH results by instability class and subwatershed.

Sub watershed		SMORPH instability class			Total
		Stable	Caution	Unstable	
Freshwater Creek	Area (km^2)	61.6	11.3	6.6	79.5
	%	77.5%	14.2%	8.3%	100.0%
Ryan Slough planning watershed	Area (km^2)	27.2	6.9	4.1	38.2
	%	71.3%	18.0%	10.7%	100.0%
Total	Area (km^2)	88.8	18.2	10.7	117.7
	%	75.5%	15.4%	9.1%	100.0%

Figure 13 and 14 describe the distribution of SMORPH instability classes by subwatershed and underlying geology types. In Freshwater Creek, approximately 62% of the Wildcat Group was classified as stable, nearly 25% was classified with a moderate instability rating, and approximately 13% was classified with a low instability rating (Figure 13). In comparison, 93% of the Franciscan mélange was classified as stable, and only 7% was classified with a moderate or high instability rating (3.6% and 3.3%, respectively). The influence of geology on model

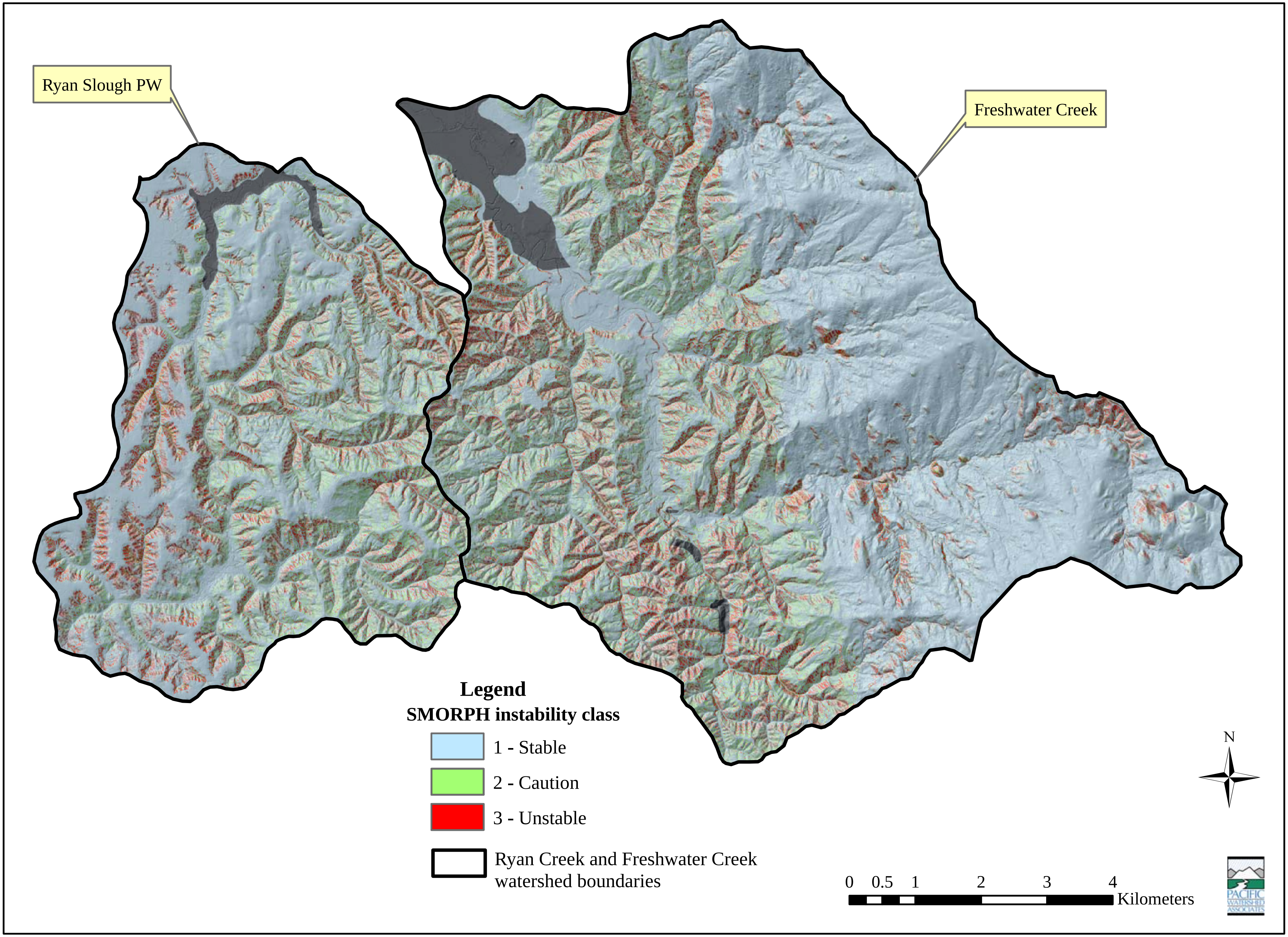


Figure 12. SMORPH results for Freshwater Creek and Ryan Slough planning watershed

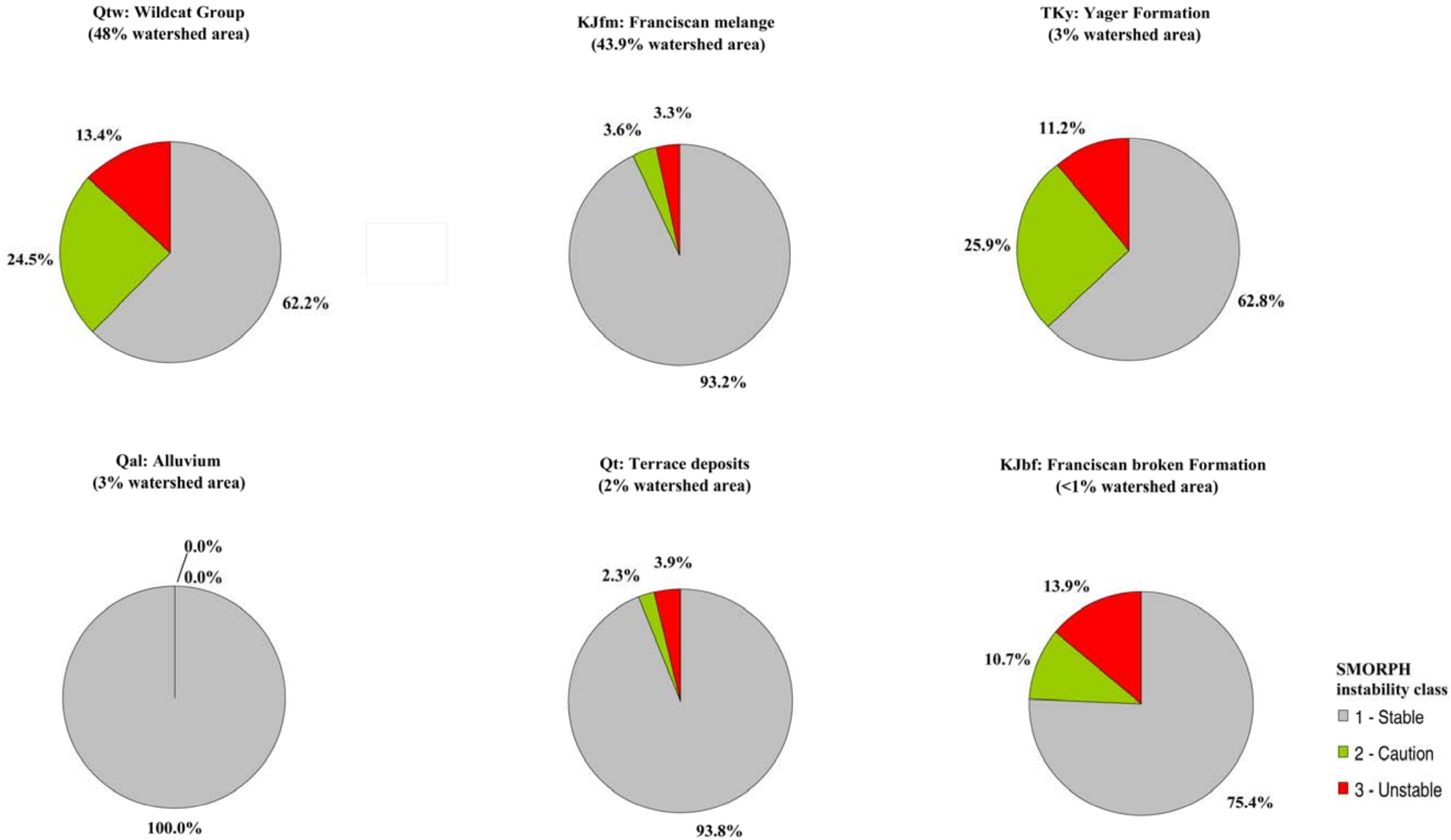


Figure 13. SMORPH instability classes by geologic type, Freshwater Creek

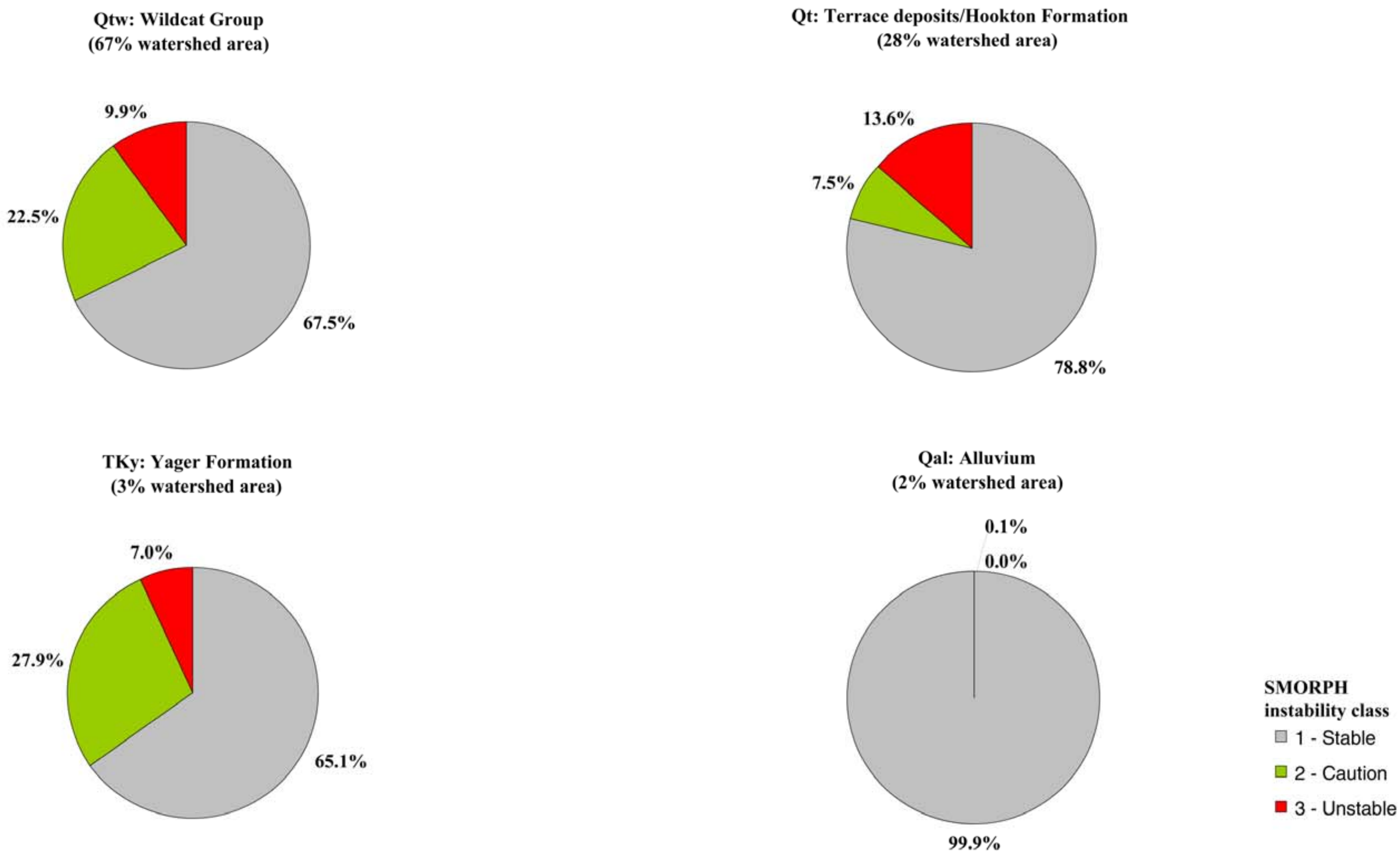


Figure 14. SMORPH instability classes by geologic type, Ryan Slough planning watershed

output is primarily manifested by its effect on hillslope steepness. For example, slopes underlain by the Franciscan mélange typically exhibit slope gradients less than 50% with fewer localized areas of steep streamside and inner gorge slopes. In comparison, slopes underlain by the Wildcat Group exhibit steeper and more incised topography with the majority of slope gradients between 50% and 65% and a comparatively high occurrence of steep streamside and inner gorge slopes >65%.

In the Ryan Slough planning watershed, the distribution of SMORPH instability classes in the Wildcat Group was very similar to the distribution observed in the Freshwater Creek subwatershed, where 68% of the Wildcat Group was classified as stable, nearly 23% was classified with a moderate instability rating, and approximately 10% was classified with a low instability rating (Figure 14). According to the SMORPH model outputs, approximately 79% of the Ryan Slough planning watershed was classified as stable, nearly 8% was classified with a moderate instability rating, and approximately 14% was classified with a high instability.

D. Slope Stability Model Performance

1. Model performance based on correct prediction of air photo identified landslides

Air photo identified landslides compiled as part of Phase I of the Freshwater Creek TMDL sediment source assessment were used to examine the performance of the three models. The air photo identified landslides were mapped on the stereo pairs of the 1987, 1997 and 2003 historic aerial photography. The time frame for the air photo analysis is approximately 26 years, assuming that landslides identified on the oldest historic air photo set (1987) could be identified up to 10 years prior, that is, as early as 1977.

One hundred and eighty-five (185) shallow road-related and non road-related landslides delivering sediment to streams were mapped in the slope stability study area, including 63 landslides observed in Freshwater Creek and 122 landslides observed in the Ryan Slough planning watershed. Shallow landslides identified as part of the analysis consisted of mainly of debris landslides and debris flows. The average landslides identified as part of the Phase I Freshwater TMDL air photo analysis was approximately 600 m^2 in size (PWA 2006). Earthflows and shallow landslides with no sediment delivery to streams were not mapped or evaluated in the slope stability modeling analysis.

The population of air photo identified landslides used to examine the performance of the models is small and may not allow an accurate evaluation of slope stability model performance. Road-related landslides were included in the comparison sample in order to increase the size of the small sample dataset. The inclusion of road-related landslides could affect the evaluation of model performance, because the 4-m LiDAR may not be able to pick of subtle drainage patterns on the road surface. Many road-related debris slides and debris flows are initiated by concentrated road surface runoff at low points on the road surface. These low points would be indiscernible at the scale of the 4-m LiDAR, and depending on the adjacent area, the low points may be considered “stable” or not prone to landsliding. As a result, road-related landslides caused by subtle road drainage patterns may not be predicted by the model outputs.

The scope of the Phase I Freshwater Creek TMDL sediment source assessment did not include the analysis of historic landslides from air photo years prior to the 1987 air photo time period, and did not include the landslides exhibiting no sediment delivery. A complete historic air photo analysis spanning available historic aerial photography, and identifying all translational landslides with and without sediment delivery would clearly provide a better population to evaluate model performance.

Figure 15, Figure 16 and Figure 17 show each of the three model outputs and the distribution of 185 air photo identified landslides. To facilitate comparative analyses, air photo identified landslide polygons mapped in ArcView were overlapped with each of the three model outputs to determine their concurrence. The area of each modeled instability class present within each of the 185 landslides was then tabulated to determine the level of concurrence. Landslides were categorized by instability/probability class according to the presence of the highest instability/probability class in the landslide polygon. That is, if a mapped landslide contained even a single SMORPH "unstable" model output unit within its boundaries, then the model prediction for that landslide polygon was classified as "unstable". This conservative methodology has been employed in other slope stability modeling studies (i.e. Stillwater Sciences 2007; Shaw and Vaugeois 1999). Using this method, the three models were evaluated as to whether they correctly predicted the occurrence of air photo identified landslides in the moderate to high instability classes (SHALSTAB and SMORPH) or the higher probability classes (PISA-m). The models were considered to be good predictors if a high percentage (>60%) of landslides were correctly classified in the higher instability/probability classes.

Table 9, Table 10 and Table 11 describe the percent watershed area in each instability/probability class and the percent frequency of air photo identified landslides classified by instability/probability class for all three models. For the purposes of this study air photo identified landslides were considered correctly predicted by the SHALSTAB model if they were classified as unstable to moderately unstable (-10 (-log(q/T) to -2.5 (-log(q/T)). SHALSTAB results show that 75% of the air photo identified landslides in the slope stability modeling study area were classified between the range of unstable and moderately unstable and therefore are considered to be correctly predicted (Table 9). Approximately 18% of the air photo identified landslide polygons were classified as moderate low instability, and 7% of the air photo identified landslide polygons were classified as low instability or stable (6.5% and 0.5%, respectively) (Table 9). Overall, SHALSTAB was proven to be a good predictor of slope stability for areas classified as unstable to moderately unstable (-10 (-log(q/T) to -2.5 (-log(q/T)) in the Freshwater Creek TMDL slope stability study area.

According to the PISA-m results, approximately 32% of the air photo identified landslide polygons were classified in the highest probability category "p>0.2" (less than 5 year landslide recurrence interval) (Table 10). In addition, 31% of the air photo landslides were classified in the second highest probability class (0.1 to 0.2 or 5 to 10 year recurrence interval). Combined, the 2 highest probability classes ("p>0.2" and "0.1 to 0.2") represent only 5% of the slope stability study area, but account for correctly classifying 63% of the landslides that occurred in the 26 year study period. This suggests that PISA-m is a good predictor for landslides in the two highest probability class areas. Only 7% of the air photo identified landslides were observed in the

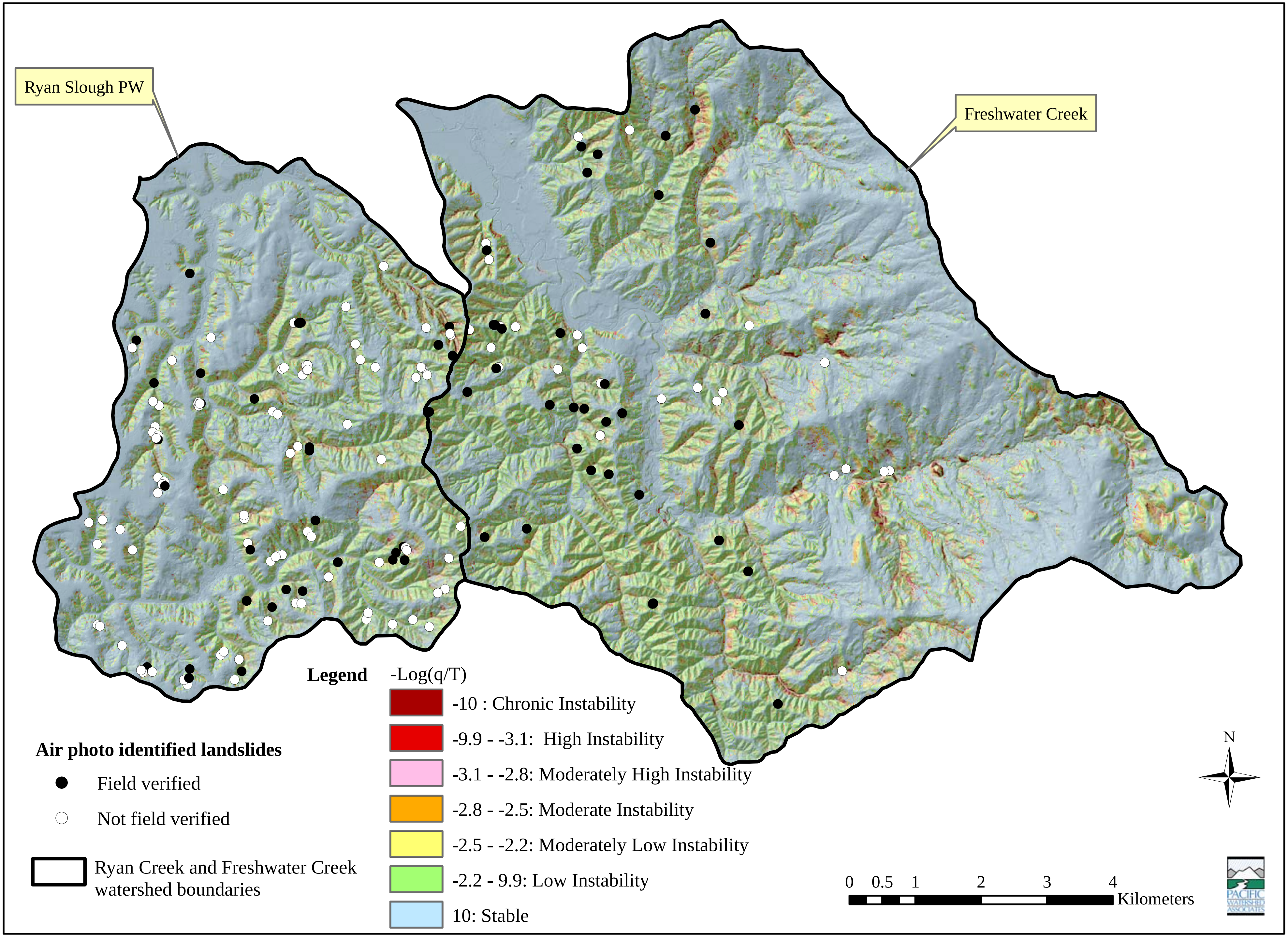


Figure 15. Air photo identified landslides and Shalstab results for Freshwater Creek and the Ryan Slough planning watershed

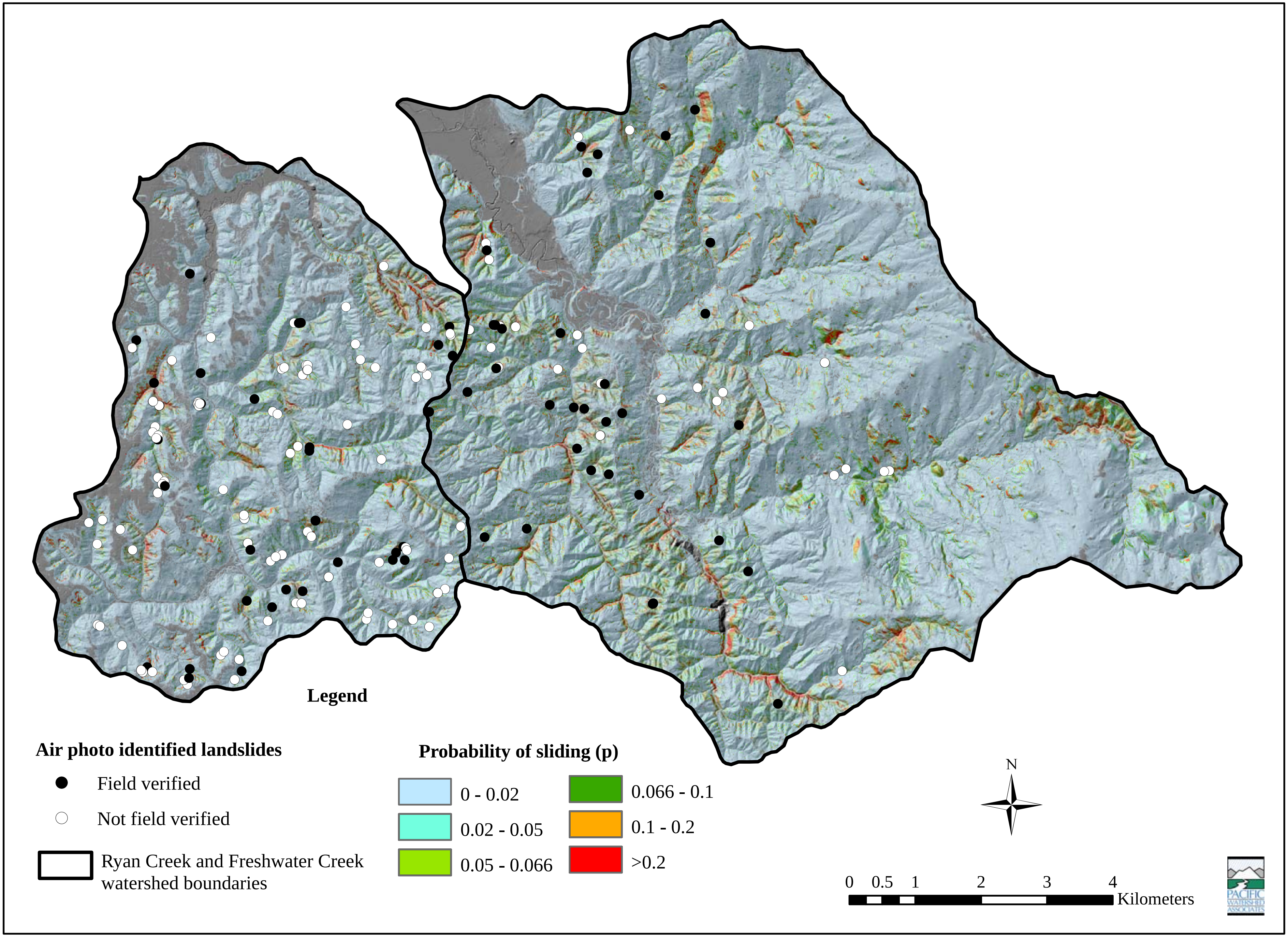


Figure 16. Air photo identified landslides and PISA results for Freshwater Creek and Ryan Slough planning watershed

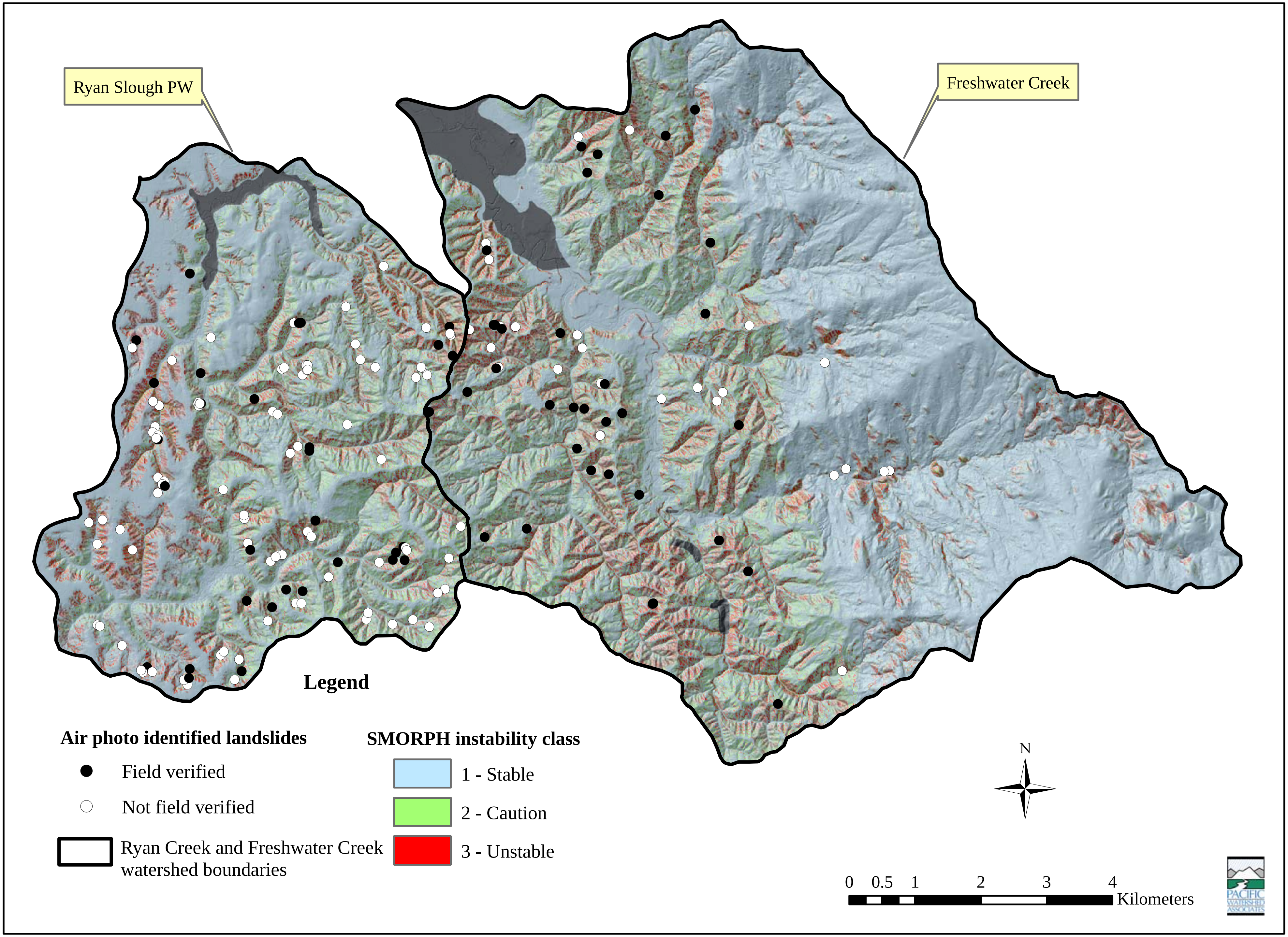


Figure 17. Air photo identified landslides and SMORPH results for Freshwater Creek and Ryan Slough planning watershed

Table 9. Percent watershed area and percent frequency of air photo identified landslides by SHALSTAB instability class and subwatershed.

Sub watershed	**SHALSTAB model results (-log(*q*/*T*))**														
	-10 (Unstable)		**<-3.1 (High instability)**		**-3.1 - -2.8 (High-moderate instability)**		**-2.8 - -2.5 (Moderate instability)**		**-2.5 - -2.2 (Moderate-low instability)**		**>-2.2 (Low instability)**		**10 (Stable)**	
	% area	% slides	% area	% slides	% area	% slides	% area	% slides	% area	% slides	% area	% slides	% area	% slides
Freshwater Creek	0.4	22.2	1.6	19.0	1.9	22.2	4.1	14.3	8.1	19.0	24.8	3.2	59.2	0.0
Ryan Slough PW	0.3	11.6	1.2%	26.4	1.6	16.5	3.5	19.0	7.5	17.4	23.7	8.3	62.2	0.8
Total	**0.4**	**15.2**	**1.5**	**23.9**	**1.8**	**18.5**	**3.9**	**17.4**	**7.9**	**17.9**	**24.4**	**6.5**	**60.1**	**0.5**

Table 10. Percent watershed area and percent frequency of air photo identified landslides by PISA-m probability class and subwatershed.

Sub watershed	**PISA-m model results (*p*)**											
	0 – 0.02 (Low probability)		**0.02 – 0.05**		**0.05 – 0.065**		**0.066 – 0.1**		**0.1 – 0.2**		**>0.2 (High probability)**	
	% area	% slides	% area	% slides	% area	% slides	% area	% slides	% area	% slides	% area	% slides
Freshwater Creek	81	3	8	6	2	0	3	10	4	37	2	44
Ryan Slough PW	85	9	7	12	2	12	2	14	3	28	2	25
Total	**82**	**7**	**8**	**10**	**2**	**8**	**3**	**12**	**3**	**31**	**2**	**32**

Table 11. Percent watershed area and percent frequency of air photo identified landslides by SMORPH instability class and subwatershed.

Subwatershed	SMORPH model results					
	Stable		Caution		Unstable	
	% area	% slides	% area	% slides	% area	% slides
Freshwater Creek	77.5	0.0	14.2	1.6	8.3	98.4
Ryan Slough PW	71.3	0.8	18.0	4.9	10.7	94.3
Total	75.5	0.5	15.4	3.8	9.1	95.7

lowest probability class <0.02 (greater than 50 year landslide recurrence interval). The lowest probability class represents more than 80% of the total slope stability study area (Table 10).

The SMORPH model appeared to predict landslide occurrence better than the SHALSTAB or PISA-m slope stability models in the moderate instability ("Caution") to unstable categories. According to the SMORPH model, landslides are correctly predicted if they are classified as moderate instability ("Caution") or "Unstable". Approximately 96% of the air photo identified landslides observed in the slope stability study area were classified as unstable and nearly 4% of the landslides were classified as "Caution" or moderate instability (Table 11). Less than 1% of the air photo identified landslides were classified as occurring on stables slopes. Thus, while correctly classifying 96% of the landslides as occurring in "unstable" areas, only 9% of the watershed area was classified as "unstable". Although this model appears to successfully discriminate existing landslide areas, the 9% of modeled unstable terrain within the watershed is visually spread over large areas, in a shotgun-like pattern, thereby reducing its practical predictive value.

2. Model performance based on comparison of air photo identified landslides and randomly generated polygons

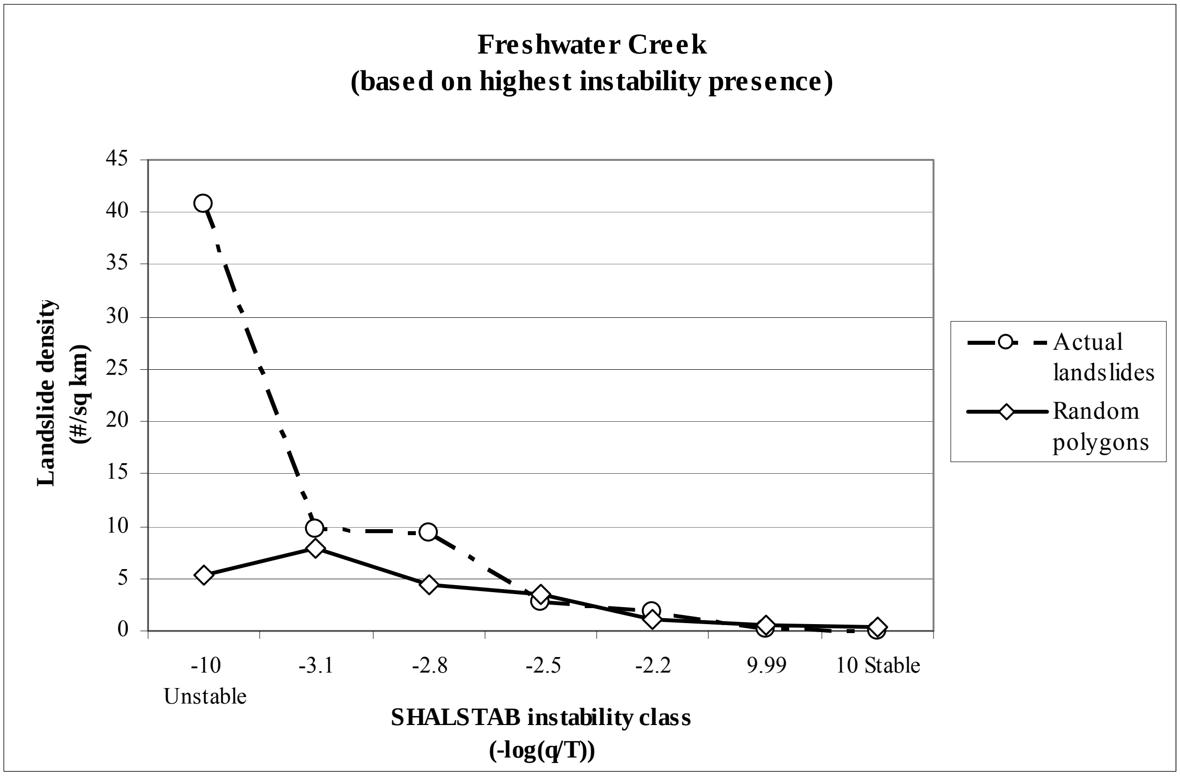
In addition to the compilation of air photo identified landslides, a sample of randomly generated polygons was developed for Freshwater Creek (N=383) and Ryan Slough planning watershed (N=381). Randomly located points were created using ArcMap and polygons were generated by developing a 15 m buffer around the random points. The 15 m buffer was chosen to best approximate the average size of the air photo identified landslides mapped in the study area. The random polygons were then overlain onto the model layers and attributed according to the same method applied to the air photo identified landslides. Comparing a randomly located sample of landslide-sized polygons in each subwatershed allows for further evaluation and comparison of the three model results.

Landslide density graphs plotting air photo landslide and random polygon densities (# slides or random polygons/instability class area (km^2)) are a useful for determining whether a model is considered a "good" predictor of landslide occurrence. Models are considered to be "good" predictors of landslide occurrence if stability/probability class densities showed evidence of divergence towards increased landslide density away from the typical constant density of random generated landslide polygons (see also Section 3C1). If air photo identified landslide densities showed little difference or no divergence from the randomly generated polygon densities graphs, then they were determined to be "poor" predictors of landslide occurrence. Figure 18, Figure 19 and Figure 20 illustrate the density of air photo identified landslides and randomly generated simulated landslides per square kilometer of the subwatershed by instability/probability class for the three models.

The SHALSTAB air photo landslide density graph shows a large divergence (440% increase in both Freshwater Creek and the Ryan Slough planning watershed) in landslide density compared to the model prediction for the randomly generated polygons from the high moderate instability class (-2.8 –log (*q/T*)) to the unstable class (-10 –log (*q/T*)) (Figure 18). This suggests that the model is a better predictor for landslides occurring in the moderate (-2.5 –log (*q/T*)) to unstable (-10 –log (*q/T*)) instability classes. Landslide densities occurring in the lower SHALSTAB instability ranges (-2.2 –log (*q/T*) to 10 –log (*q/T*)) do not differ substantially from randomly generated polygon densities. This indicates that the SHALSTAB model is not a good predictor for landslides occurring in the lower instability classes. Densities for randomly generated polygons showed an overall minor increase across SHALSTAB lower instability classes (>-2.5 –log(*q/T*)) in Freshwater Creek and the Ryan Slough planning watershed.

Landslide density graphs for PISA-m also show a major divergence in air photo identified landslide density from randomly generated polygon density, but varied considerably between the two subwatersheds. Air photo identified landslide density increased by nearly 375% from the probability range "0.1 to 0.2" to the highest instability class "p>0.2"in the Freshwater Creek subwatershed (Figure 19). The PISA-m model proved to be a poor predictor of landslide occurrence in the probability classes less than 0.1 in the Freshwater Creek watershed as witnessed by the lack of divergence from the randomly generated polygon density graph. In contrast, the Ryan Slough planning watershed exhibited a large divergence from the randomly generated polygon densities between the "0.05 to 0.065" and "0.1 to 0.2" probability ranges. At the "0.1 to 0.2" probability range, the air photo landslide density increased dramatically from 23 slides/km^2 to 57 slides/km^2. The Ryan Slough planning watershed density graph suggests that PISA-m is a good indicator of slope instability in the probability classes greater than 0.05, but not as good a predictor in the lower probability classes between 0 and 0.05 (Figure 19).

Using the modified classification scheme, PISA-m was a better landslide predictor in the Ryan Slough planning watershed as compared to Freshwater Creek. The different results from the 2 watersheds may be due to the smaller population (N=63) of representative air photo identified landslides observed in Freshwater Creek as compared to the larger population (N=122) of air photo identified landslides observed in the Ryan Slough planning watershed. The small population of air photo identified landslides in the Freshwater Creek watershed may represent an inadequate (and therefore non-representative) sample to provide comparable results with the Ryan Slough planning watershed.



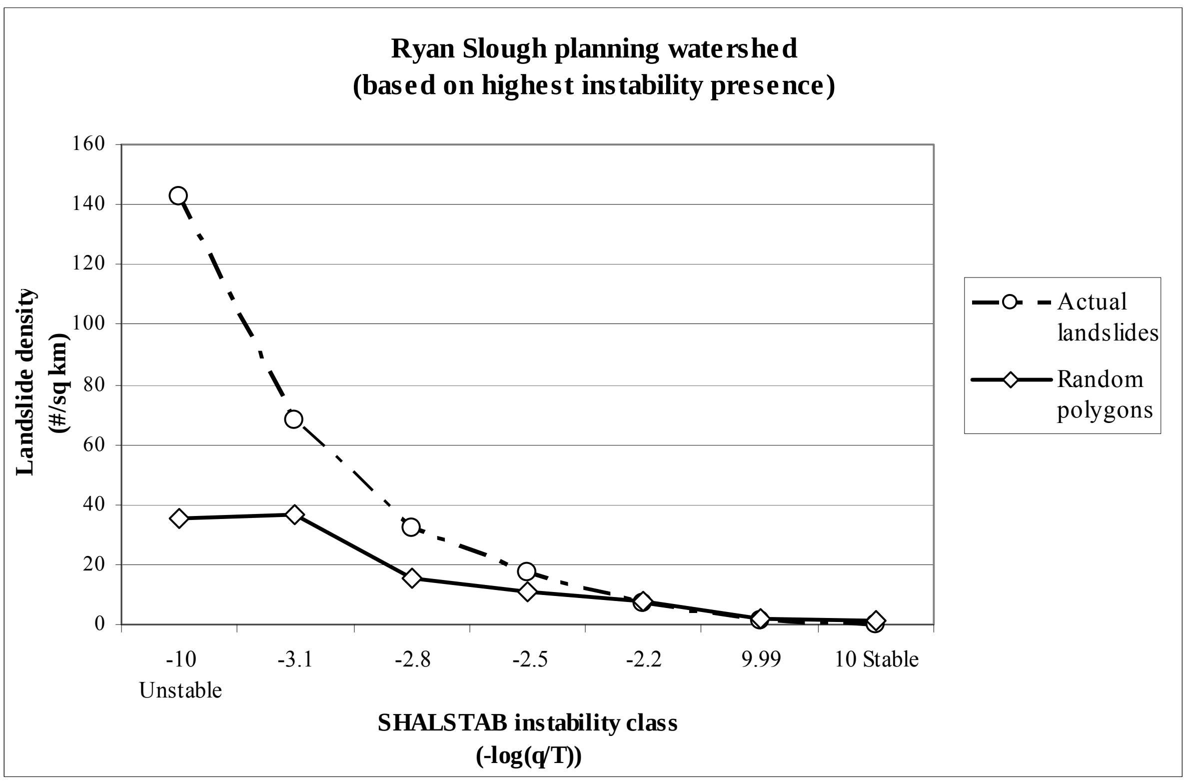
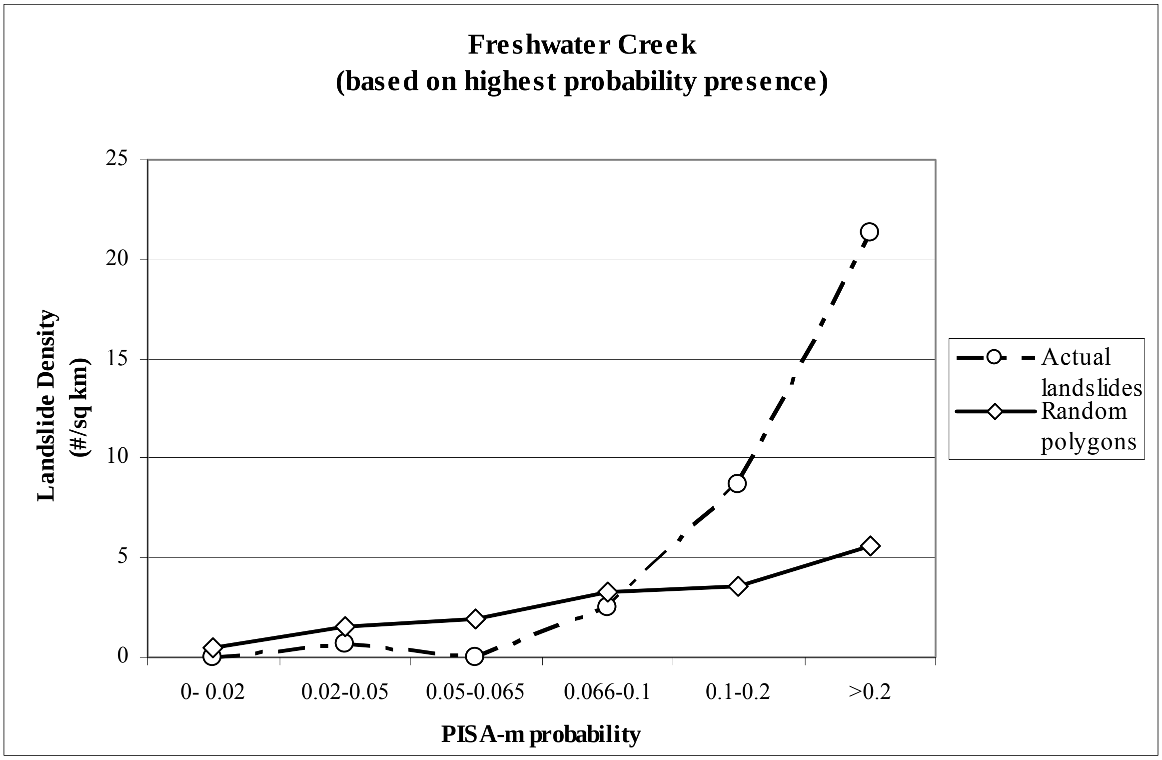


Figure 18. Density of air photo identified landslides and random slide-size polygons by SHALSTAB instability class, classification based on presence of highest instability class, Freshwater Creek and Ryan Slough planning watershed.



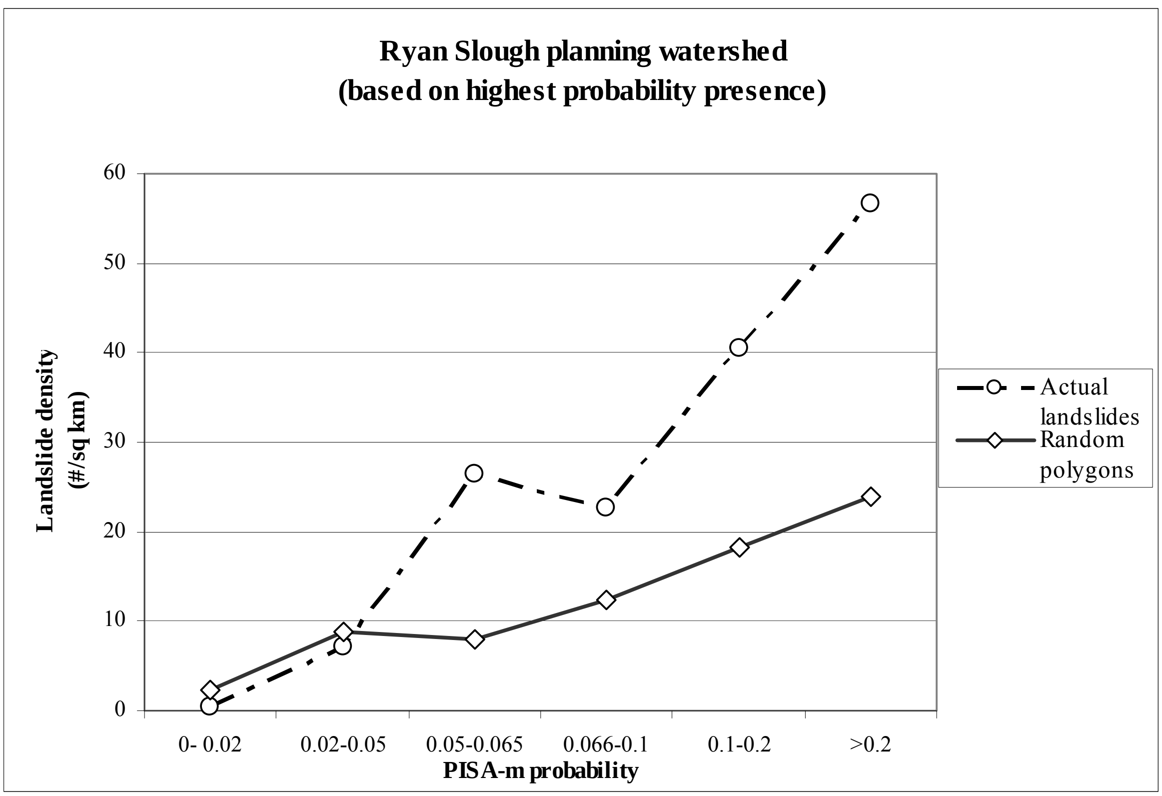
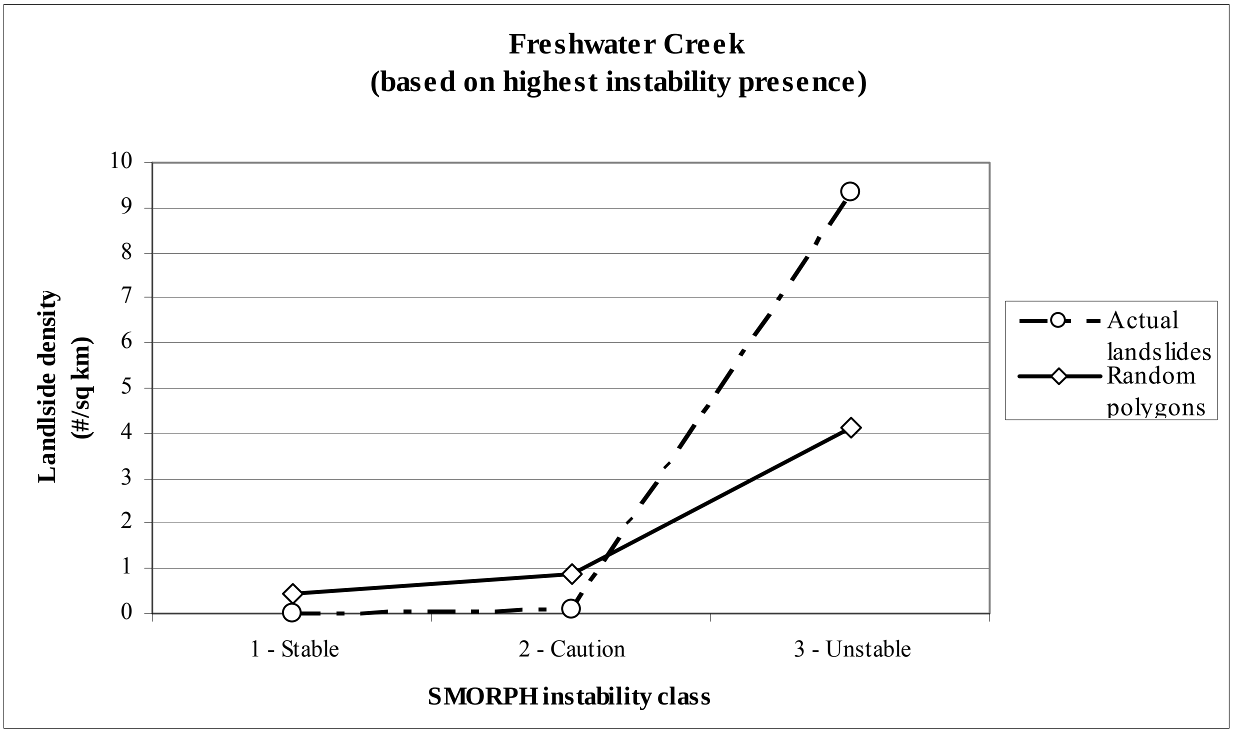


Figure 19. Density of air photo identified landslides and random slide-size polygons by PISA-m probability, classification based on presence of highest probability class, Freshwater Creek and Ryan Slough planning watershed.



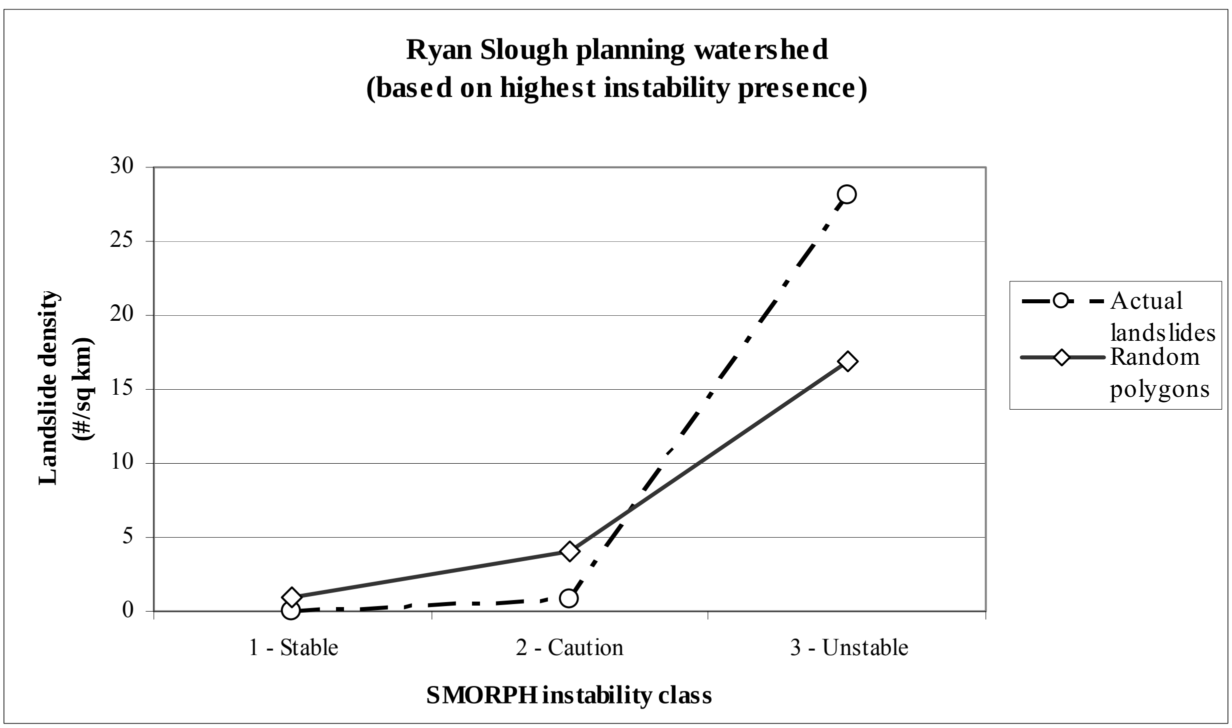


Figure 20. Density of air photo identified landslides and randomly generated polygons by SMORPH instability class, Freshwater Creek and Ryan Slough planning watershed.

Of the three slope stability models, the SMORPH model correctly predicted the highest percent (96%) of air photo identified landslides in the highest instability class ("Unstable") (Table 11). The high percent of correctly predicted air photo identified landslides may be an over prediction due to the "shot gun" pattern of high instability model-output raster polygons coupled with using the presence of the highest instability classification method.

Figure 20 illustrates the landslide densities for the air photo identified landslides and randomly generated polygons by the three SMORPH instability classes. Although the SMORPH model correctly predicted landslides in the "Unstable" class, the randomly generated polygons exhibited a similar substantial increase in density in the "Unstable" class. This suggests that SMORPH may not be a good predictor for slope instability in the highest instability class and could result in the over prediction of landslides in this instability class. In general, the majority of both air photo identified landslides and randomly generated polygons contained a distribution of all three instability classes.

As stated previously, air photo identified landslides were attributed by instability/probability class based on presence of the highest instability/probability class in the landslide polygon, regardless of the polygon area. This approach is typically employed in other slope stability models (i.e. SMORPH and SHALSTAB) and assumes that any size area of the highest instability class within a polygon defines the entire landslide instability/probability classification. Table 12 lists the general statistics generated for the percent area of the highest instability class defining the instability/probability classification of the air photo identified landslides. For all three models, the range of percent area of the highest instability/probability class that defined the overall instability/probability classification for each landslide polygon ranged from 0.004% to 100% (Table 12). The minimum percent area of instability/probability class defining the model classification ranged from 0.004% (SHALSTAB) to 2% (SMORPH).

Table 12. General statistics percent area of highest instability/probability class defining the classification of air photo identified landslides.

Statistic	SHALSTAB (%)		PISA-m (%)		SMORPH (%)	
	Freshwater Creek	Ryan Slough	Freshwater Creek	Ryan Slough	Freshwater Creek	Ryan Slough
Minimum % area	0.05	0.004	0.1	0	2	1
Maximum % area	98	100	100	100	76	100
Average % area	18	23	18	23	32	40
Median % area	10	10	12	12	30	35

The minimum percent area defining the instability/probability classification is quite low for all three models and could result in the over prediction of air photo landslides in the higher instability/probability classes if a fraction of a 4-m polygon (LiDAR DEM raster size) is used for predictive classification. For example, one landslide mapped in the Ryan Slough planning watershed exhibited 0.05 m^2 in the second highest SHALSTAB instability class (>3.1 $-\log(q/T)$).

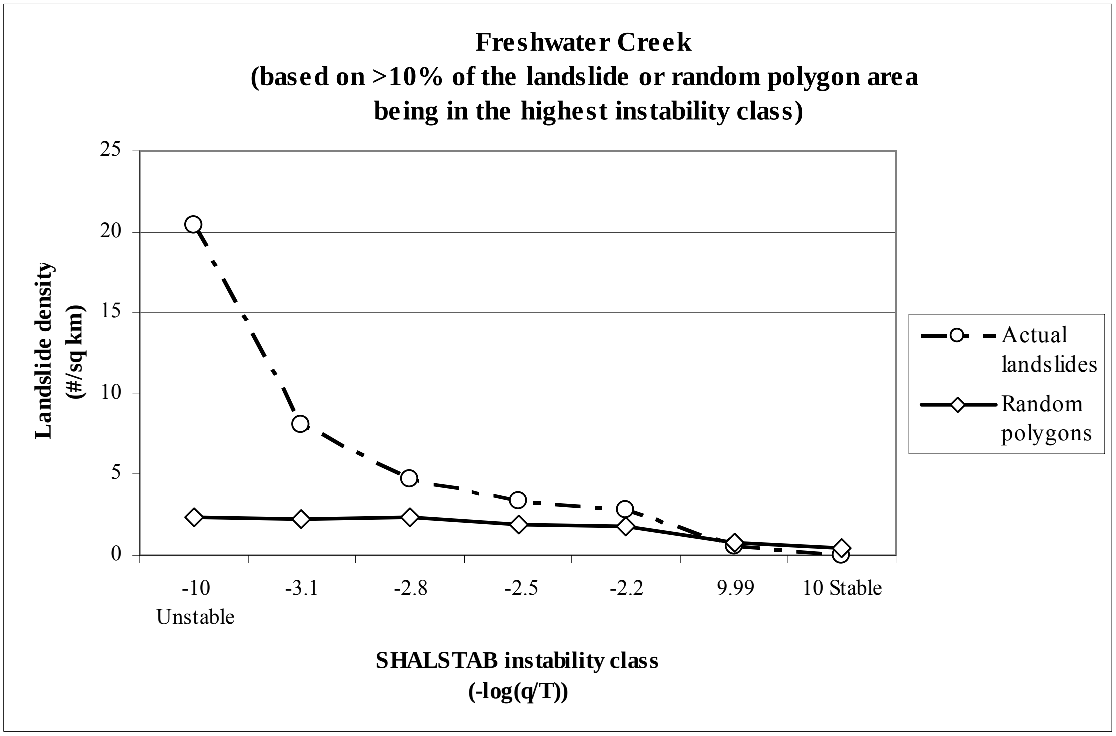
Using the presence of the highest stability class classification method, this landslide was classified in the ">3.1 –log(*q/T*)" instability category. The 0.05 m^2 area only represents 0.004% of the total landslide area (1,340 m^2). This very small area could represent a minor topographic irregularity that may not affect slope stability. Over 80% of the landslide area (1,070 m^2) was classified in the 2 lowest instability classes (>2.2 –log(*q/T*) and 10 –log(*q/T*)).

To evaluate model performance using a more rigorous instability classification method, air photo identified landslides and randomly generated polygons were re-classified based on the median percent area of highest SHALSTAB instability class (10% median area), PISA-m probability class (12% median area), and SMORPH instability class (33% median area) (Table 12). This additional analysis was conducted to determine whether instability/probability classification based on a larger percent area of instability/probability class within the air photo and random landslides would result in better model performance, as compared to classification based on strictly presence of the highest instability/probability class. Landslide density graphs were re-plotted for the SHALSTAB, PISA-m and SMORPH models and are illustrated in Figure 21, Figure 22, and Figure 23. An instability/probability threshold of the median percent area of the highest instability/probability class for each sample population was derived after several iterations of instability/probability classification using different thresholds of percent area of the highest instability/probability class (mean percent area; 50%; 25% of the highest instability/probability class) in each sample population. Using landslide density graphs, it appeared that a threshold of the median percent area of the highest instability/probability class defining instability classification resulted in overall better model performance for the project watersheds in comparison to the other larger percent area thresholds.

The SHALSTAB model for the Freshwater Creek subwatershed appeared to perform better using the criteria of a minimum 10% median area of the highest instability class (Figure 21). Using this alternative method there was increased divergence (i.e., better differentiation and thus performance) in the -2.8 - -2.5 –log (q/T) instability range. In comparison, applying the minimum 10% median classification area to air photo landslides and randomly generated polygons in the Ryan Slough planning watershed did not result in better model performance. The Ryan Slough planning watershed results showed an overall poorer performance with an increased density of air photo landslides in the lower probability classes (-2.5 - -2.2 –log (q/T)).

The PISA-m model for the Freshwater Creek subwatershed appeared to perform better using the criteria of a minimum 12% median area of the highest instability class. There is considerable increased divergence of the air photo landslide density in the "0.05 to 0.065" probability class range (Figure 22). Conversely, applying the minimum 12% median classification area to air photo landslides and randomly generated polygons in the Ryan Slough planning watershed did not result in better model performance. Similar to the SHALSTAB model, the PISA-m model results for the Ryan Slough planning watershed results showed an overall poorer performance with an increased density of air photo landslides in the lower probability class "0.02-0.05" (20-50 year landslide recurrence interval)

The SMORPH model appeared to perform better using the criteria of a minimum 33% median area of the highest instability class in both subwatersheds. The randomly generated polygons did not follow the increasing landslide density trend as seen in Figure 20 using the presence of the



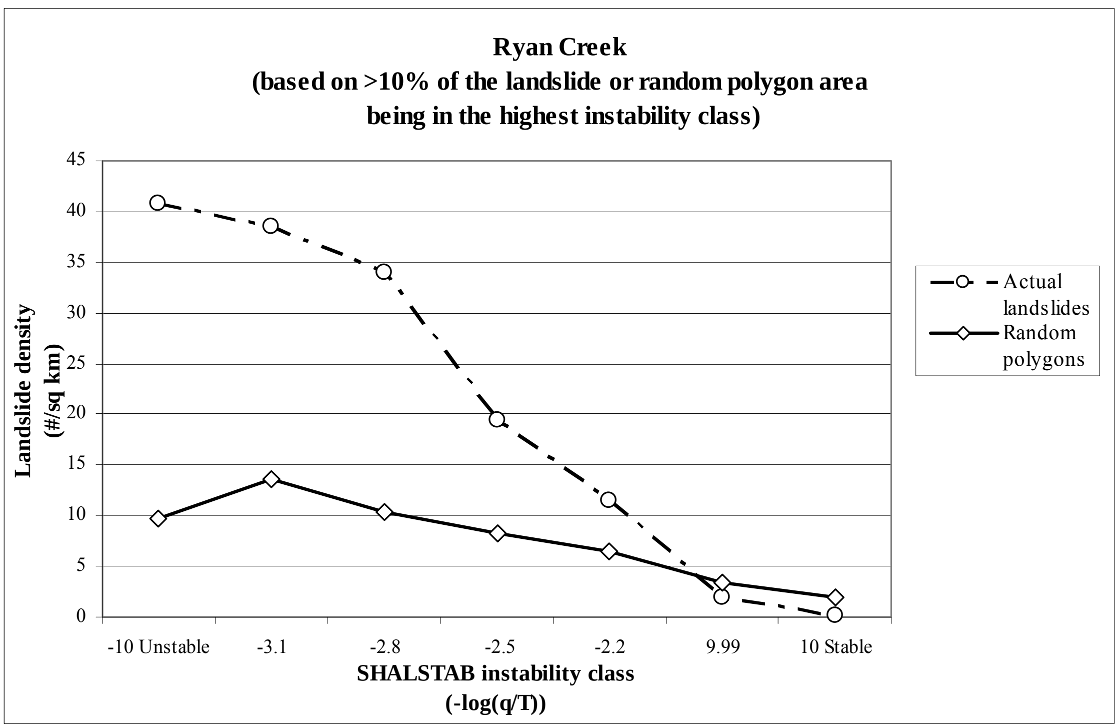
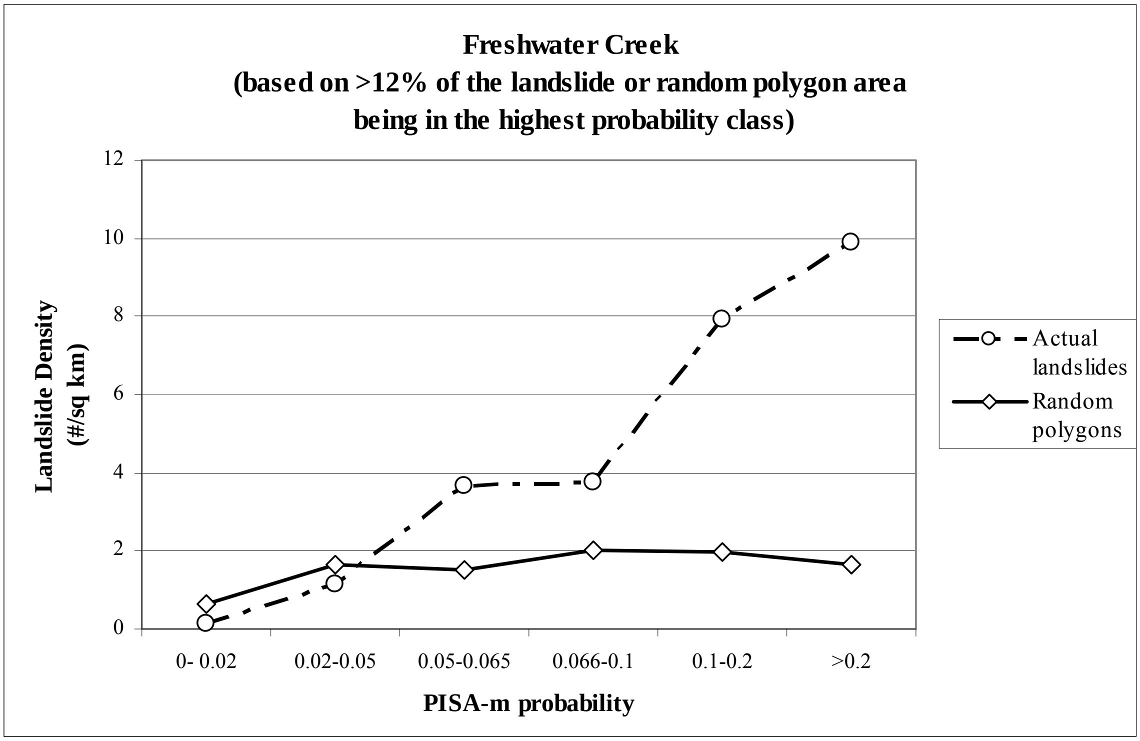


Figure 21. Density of air photo identified landslides and randomly generated polygons by SHALSTAB instability class, classification based on minimum 10% area of highest probability class, Freshwater Creek and Ryan Slough planning watershed.



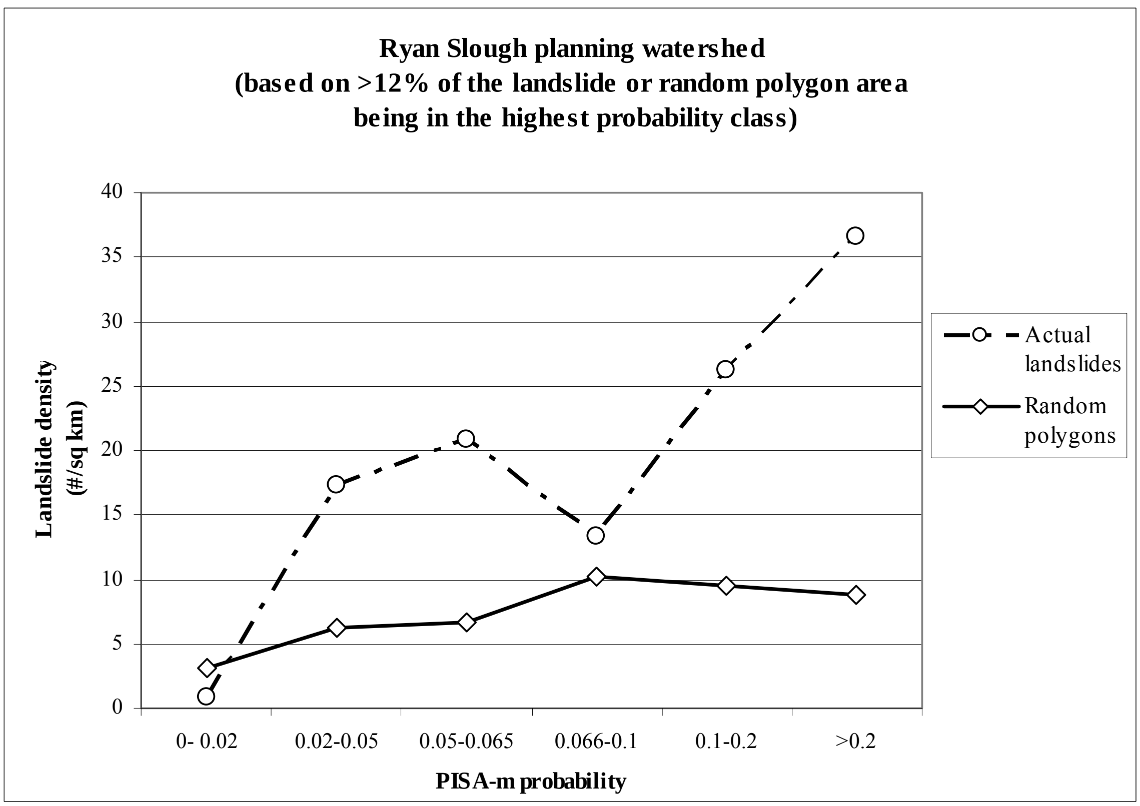
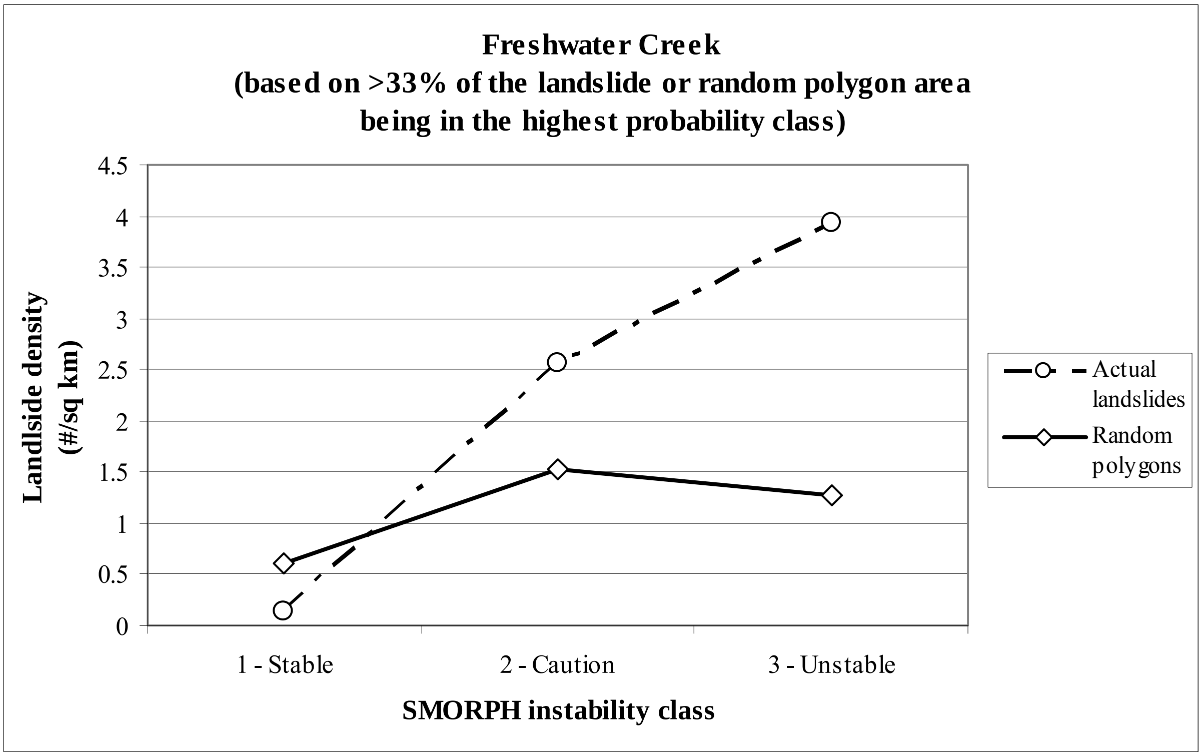


Figure 22. Density of air photo identified landslides and randomly generated polygons by PISA-m probability class, classification based on minimum 12% area of highest probability class, Freshwater Creek and Ryan Slough planning watershed.



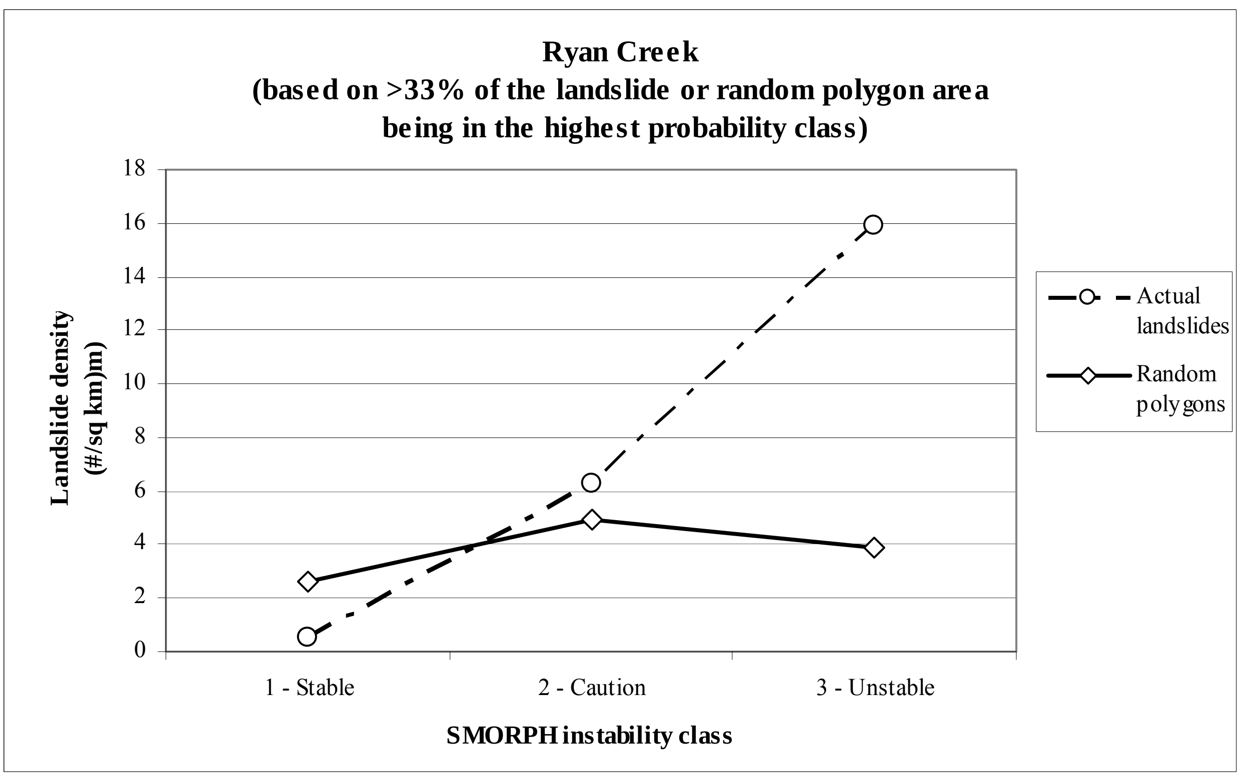


Figure 23. Density of air photo identified landslides and randomly generated polygons by SMORPH instability class, classification based on minimum 12% area of highest instability class, Freshwater Creek and Ryan Slough planning watershed.

highest instability class classification method, and as a result there was a major divergence between the air photo landslide and randomly generated polygon densities in the "caution" or moderate instability class (Figure 23). As a result, the SMORPH model was a better landslide predictor in the "caution" and unstable classes using the minimum median percent area of the highest stability/probability method as compared to presence of the highest stability class method.

Classifying landslides using a larger area of the highest instability/probability method may or may not result in better model performance. Overall, using the minimum percent area of the highest instability probability class can result in better model performance through increased divergence and air photo landslide and random polygon densities in the lower probability classes. This validation test has shown that classifying landslides and randomly generated polygons by minimum percent area of the highest instability/probability method can have varying results between the three models, and within the one model between subwatersheds (e.g. PISA-m).

3. Model performance based comparison between the three models

The performance of the SHALSTAB, PISA-m, and SMORPH models were compared to each other using three methods. The first method involved comparing each model's landslide predictive performance. Landslides were classified as correctly or incorrectly predicted using model-specific criteria. The second method uses the cumulative percentage graph technique. The third method involved overlaying the three model outputs to determine the amount of overlap of unstable areas between all three or between 2 models. The following section discusses the results of the model comparison analysis.

a. Model predictive performance method 1 - model-specific landslide prediction

The three models were compared to each other by evaluating their predictive performance through analysis of the frequency of "correctly" and "incorrectly" predicted air photo identified landslides. Air photo identified landslides were classified as correctly and incorrectly predicted using criteria specific to each model. For the SMORPH model, landslides were considered to be correctly predicted if they occurred on moderate instability ("caution") and unstable slopes. Air photo landslides occurring on "stable" slopes were considered incorrectly predicted. According to the SMORPH model, approximately 76% of the study area was classified as "stable" 15% was classified as "caution", and only 9% was classified as "unstable" (Table 11). The SMORPH model correctly predicted the most air photo landslides on slopes classified as "caution" and "unstable" compared to any other model. Over the entire slope stability modeling study area, SMORPH corrected predicted 99% of the air photo landslides (Table 13). The SMORPH model only incorrectly predicted 1 landslide on slopes classified as stable.

For the purposes of this study, landslides were considered correctly predicted by the SHALSTAB model if they occurred on slopes with an instability classification of $\leq$ -2.5 – log(q/T). This incorporates all slopes that are classified with moderate, moderate high and high instability, and unstable slopes. According to the SHALSTAB model outputs, slopes classified with SHALSTAB instability values $\leq$-2.5 –log(q/T) represent nearly 8% of the study area. Approximately 75% of the air photo identified landslides were correctly predicted, and 25% were incorrectly predicted according to the SHALSTAB model (Table 13). Of the 25% of air photo landslides that were incorrectly predicted, 18% were classified as moderate low instability, 6.5% were classified as low instability, and 0.5% were classified as stable.

Choosing a probability threshold value for instability for PISA-m was problematic. For the purposes of this study, we chose an instability threshold probability value of 0.066 to represent the break between stable ($p<0.066$) and unstable ($p\geq0.066$) terrain. This probability value was chosen because it represented the period of greatest mass wasting vulnerability following vegetation removal until transpiration and root strength have substantially recovered. The PISA-m model variables for root strength and tree surcharge are based on stand ages developed from the CALVEG (CDF 2001) dataset coupled with stand age attributes compiled by PWA using the Fire and Resource Assessment Program's (FRAP) Timber Harvest Plans (THP). Based on the information provided in Table 3, root strength and tree surcharge values are lowest in non-forested areas, and reduced in forested areas with stand ages <15 years. The highest root strength and tree surcharge values were assigned to forested areas with stand ages >15 years. Because Freshwater Creek and the Ryan Slough planning watershed are primarily forested, we chose to classify slopes as unstable if their probability of landsliding failure equaled or exceeded 0.066 ($p\geq0.066$).

Slopes classified as unstable ("$p\geq0.066$") represent 8% of the total study area. According to the PISA-m model, approximately 75% of the air photo identified landslides were correctly predicted, and 25% were incorrectly predicted (Table 13). Of the 25% of air photo landslides that were incorrectly predicted, 8% were classified in the "0.05-0.065" probability class (16 to 20 year landslide recurrence interval), 10% were classified in the "0.02-0.05" probability class (20 to 50 year landslide recurrence interval), and 7% were classified in the "<0.02" probability class (>50 year landslide recurrence interval).

Table 13. Air photo identified landslides and model predictive performance

Sub watershed	Number of landslides (#)	SHALSTAB				PISA-m				SMORPH			
		Landslides correctly predicted		Landslides incorrectly predicted		Landslides correctly predicted		Landslides incorrectly predicted		Landslides correctly predicted		Landslides incorrectly predicted	
		(#)	(%)	(#)	(%)	(#)	(%)	(#)	(%)	(#)	(%)	(#)	(%)
Freshwater Creek	63	49	78	14	22	57	90	6	10	63	100	0	0
Ryan Slough	122	89	73	33	27	82	67	40	33	121	99	1	1
Total	**185**	**138**	**75**	**47**	**25**	**139**	**75**	**46**	**25**	**184**	**99**	**1**	**1**

b. Model predictive performance method 2 -- cumulative percentage graphs

In addition to the method discussed above, model predictive performance was also measured by developing cumulative percentage graphs. This is an approach that reveals the ability of the models to predict landslides by selectively capturing areas with a high frequency of landsliding, and excluding areas that exhibit low frequency of landsliding. The method, as used in this study, is based on a study conducted by Reid et al (2007), who measured the predictive success of four different topographic models (SHALSTAB, slope, proximity to stream, and landslide prone landforms) using cumulative percentage graphs developed from 16 data sets of past debris slides.

Cumulative percentage graphs compare the proportion of the study area associated with each slope instability or probability class with the percentage of landslide area (based on air photo documentation) captured in each class. The method does not require the classification of landslides by the highest instability or probability class, and makes no assumption of the location of landslide initiation within landslide polygons. As a result, the analysis of cumulative percentage graphs is a more conservative approach to the assessment of model predictive performance.

Cumulative percentage graphs are developed by plotting the highest instability or probability class first, which is represented in the lower left hand corner of the plot (e.g., Figure 24). Each instability or probability class from highest to lowest is then cumulatively added to the graph to total 100% of the instability or probability class area. The 1:1 line shown on the graph represents “random success” (Figure 24). On any point of the 1:1 line, the landslide area within the instability or probability class is proportional to the total instability or probability class area in the subwatershed. According to this method, models perform well if their results plot above the 1:1 line or toward the upper left-hand corner of the graph. Results plotting in the upper left-hand portion of the graph show good agreement between the percentage of verified landslide area and the percentage of the total subwatershed area within the same instability or probability class.

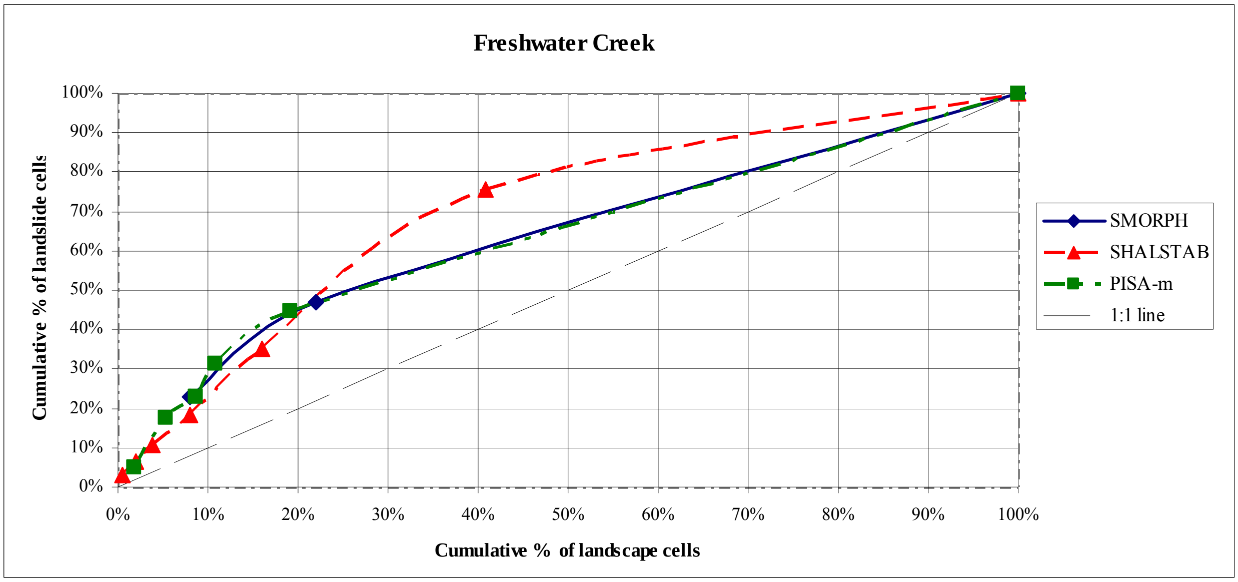
Figure 24 shows cumulative percentage graphs for the SHALSTAB, PISA-m, and SMORPH models. Results for all three models plotted above the 1:1 line (“random success”) and therefore appear to perform moderately well as compared to each other. For example, all three models predicted approximately 45% of the air photo identified landslide area before encompassing 20% of the Freshwater Creek subwatershed. In comparison, the models predicted approximately 38% to 55% of the air photo identified landslide area before encompassing 20% of the Ryan Slough planning watershed area (Figure 24). Overall results for the Freshwater Creek subwatershed and Ryan Slough planning watershed show that the SHALSTAB model performed slightly better than the SMORPH and PISA-m models. The PISA-m model showed the poorest performance for the Ryan Slough planning watershed.

c. Model predictive performance method 3 - comparison by area of overlap

The three model outputs were also analyzed to determine if the model outputs are spatially consistent with each other. This was achieved using ArcMAP to overlay the model outputs and determine the area of overlap between unstable and stable slope stability areas for the three models (Table 14 and Table 15). Unstable slopes were defined for the three models using the same criteria used for landslide classification where:

- SHALSTAB: unstable slopes were classified with SHALSTAB instability values of ≤ -2.5 –log(q/T).
- PISA-m: unstable slopes were classified with probability values greater than 0.066 (≤15 year landslide recurrence interval).

- SMORPH: unstable slopes were classified by the "Caution" or "Unstable" SMORPH instability classes.



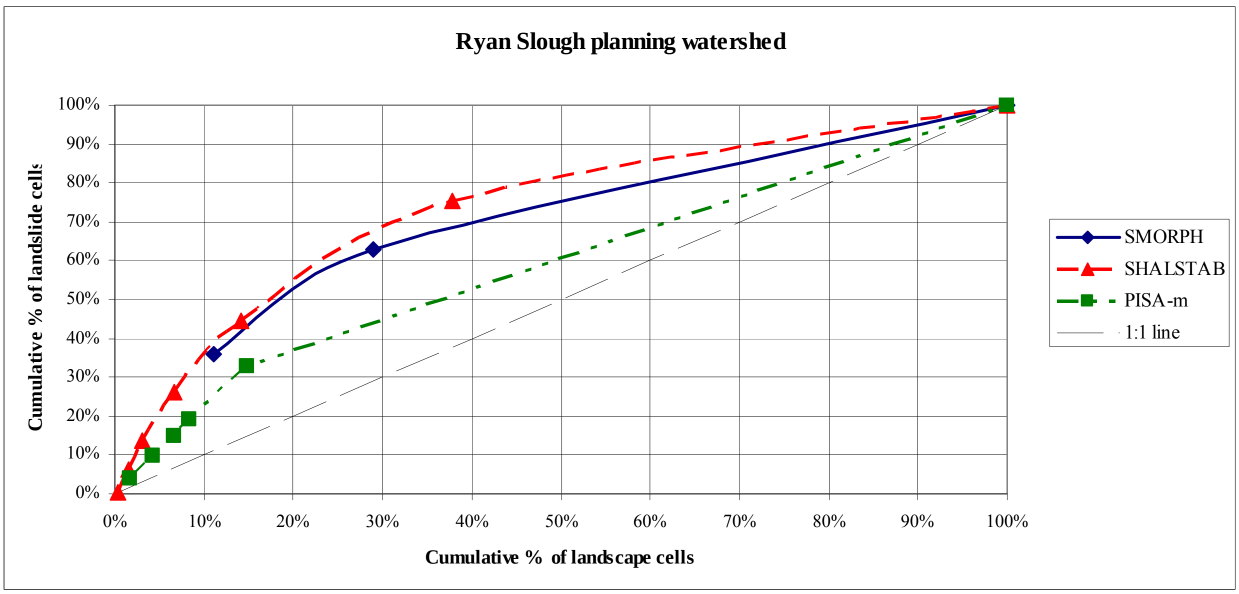


Figure 24. Cumulative percentage plots comparing predictive success of the three models, Freshwater Creek and Ryan Slough planning watershed.

Table 14 illustrates the overlap of unstable areas classified according to each of the three models by each subwatershed and for the entire study area. The SHALSTAB model identified approximately 8.84 km^2 of unstable slopes (<-2.5 –log(q/T)) (Table 14). The SHALSTAB model exhibited approximately 23% overlap between the other 2 models (PISA-m and SMORPH). The SMORPH model had the highest percent (90%) overlap with unstable areas classified by

SHALSTAB. This includes the percent area overlap between the three models (23%) and the overlap between 2 models: SMORPH and SHALSTAB (67%) (Table 14). Only 8% (0.67 km^2) of unstable areas classified by SHALSTAB exhibited no overlap between the 2 other models (Table 14).

Approximately 8.5 km^2 of unstable slopes were defined by the PISA-m model based on the instability threshold of "p≥0.066" (Table 14). Twenty-four percent (24%) of the unstable area defined by PISA-m exhibited overlap between the other 2 models (SHALSTAB and SMORPH). Similar to the SHALSTAB overlap results, the PISA-m model exhibited the highest percent (69%) overlap with unstable areas classified by the SMORPH model (percent area overlap between the three models (24%) and the overlap between PISA-m and SMORPH (45%) (Table 15). Twenty-eight percent (28%) of the unstable area defined by the PISA-m model exhibited no overlap with the other 2 models (Table 14).

Compared to SHALSTAB and PISA-m, the SMORPH model classified the highest percentage of unstable area within the slope stability study area. Unstable areas defined by the SMORPH model were classified as "caution" and "unstable" according to the model outputs. Approximately 28.9 km^2 of unstable area was defined by the SMORPH model, representing more than 3 times the unstable area defined by each of the SHALSTAB and PISA-m models (Table 14). However, the SMORPH unstable areas showed minor (7%) overlap with both SHALSTAB and PISA-m models and approximately 59% (17 km^2) of the unstable areas defined by the SMORPH model did not exhibit any overlap with the other 2 models.

Table 15 describes the results of the model output overlay comparison by total subwatershed area. Only 2% of the entire study area exhibited overlap of all the three models on slopes classified as unstable. The SMORPH and SHALSTAB models exhibited minor overlap of approximately 8% for the entire study area (overlap between the three models (2%) and the overlap between 2 models: SMORPH and SHALSTAB (6%)). The PISA-m model also exhibited very little overlap (6%) with the other 2 models (overlap between the three models (2%) and the overlap between 2 models: SMORPH and PISA-m (4%), for a total of 8% overlap). Approximately 69% of the study area exhibited overlap for areas classified as not unstable by the three models.

The model overlap analysis of unstable areas suggests that the three models do not correspond well with each other. The SHALSTAB and PISA-m models classified approximately the same areas of instability (8.84 km^2 and 8.49 km^2, respectively), but exhibited only 25% of overlap between the two models (Table 14). This implies that the 2 models define different slope locations as unstable. The dissimilarity in classification of unstable slopes is likely due to the differences in the application of the infinite slope equation in the SHALSTAB and PISA-m models. The SHALSTAB model is based on an infinite slope equation that uses a simulated variable steady state pore water pressure, but does not allow for spatial variability in other soil properties (i.e. soil thickness, friction angle, and cohesion). The model assumes that water

Table 14 Model overlap of unstable areas by model type and subwatershed.

Model		Freshwater Area (km^2)					
		Total unstable area	**All 3 models**	**PISA/ SMORPH**	**PISA/ SHALSTAB**	**SMORPH/ SHALSTAB**	**No overlap**
SHALSTAB	km^2	6.32	1.58	--	0.19	4.24	0.31
	%	100%	25%	--	3%	67%	5%
PISA-m	km^2	6.35	1.58	2.60	0.19	--	1.98
	%	100%	25%	41%	3%	--	31%
SMORPH	km^2	17.91	1.58	2.60	--	4.24	9.49
	%	100%	9%	15%	--	24%	53%
Model		**Ryan Area (km^2)**					
		Total unstable area	**All 3 models**	**PISA/ SMORPH**	**PISA/ SHALSTAB**	**SMORPH/ SHALSTAB**	**No overlap**
SHALSTAB	km^2	2.53	0.47	--	0.02	1.68	0.36
	%	100%	19%	--	1%	66%	14%
PISA-m	km^2	2.14	0.47	1.23	0.02	--	0.42
	%	100%	22%	58%	1%	--	20%
SMORPH	km^2	10.97	0.47	1.23	--	1.68	7.50
	%	100%	4%	11%	--	15%	68%
Model		**Entire study area Area (km^2)**					
		Total unstable area	**All 3 models**	**PISA/ SMORPH**	**PISA/ SHALSTAB**	**SMORPH/ SHALSTAB**	**No overlap**
SHALSTAB	km^2	8.85	2.05	--	0.21	5.92	0.67
	%	100%	23%	--	2%	67%	8%
PISA-m	km^2	8.49	2.05	3.83	0.21	--	2.40
	%	100%	24%	45%	2%	--	28%
SMORPH	km^2	28.88	2.05	3.83	--	5.92	16.99
	%	100%	7%	13%	--	20%	59%

Table15. Model overlap of unstable and stable areas by subwatershed

Sub watershed	**Area of sub watershed/ %**	**Unstable areas with overlapping models**				**Unstable areas with no model overlap**			**Not unstable areas defined by all 3 models**	**Total**
		All 3 models	**PISA/ SMORPH**	**PISA/ SHALSTAB**	**SMORPH/ SHALSTAB**	**PISA**	**SHALSTAB**	**SMORPH**		
Freshwater Creek	Area (km^2)	1.58	2.60	0.19	4.24	1.98	0.31	9.49	52.76	73.15
	%	2%	4%	0.3%	6%	3%	0.4%	13%	72%	100%
Ryan Slough PW	Area (km^2)	0.47	1.23	0.02	1.68	0.42	0.36	7.50	20.30	31.97
	%	1%	4%	0.1%	5%	1%	1%	23%	63%	100%
Total	**Area (km^2)**	**2.05**	**3.83**	**0.21**	**5.91**	**2.41**	**0.68**	**16.99**	**73.05**	**105.12**
	%	2%	4%	0.2%	6%	2%	1%	16%	69%	100%

accumulates in concave hollows where soils are thicker and exhibit higher pore water pressures. As a result, the SHALSTAB model considers slope angle and slope curvature as the primary driving forces affecting slope stability.

In contrast, the PISA-m model calculates instability using an infinite slope stability equation that incorporates soil shear strength, pore water pressures, and tree root shear strength. The PISA-m model assumes that pore water pressure is random across the landscape with no areas persistently wetter or drier than others. As a result, the location of unstable areas classified by the SHALSTAB and PISA-m models are very different from each other. Figure 25 illustrates the difference between the three model outputs and the classification of unstable areas in an upper portion of Little Freshwater Creek, tributary to Freshwater Creek. The PISA-m model output shows more unstable areas along planar, steep inner gorges and streamside slopes, whereas the SHALSTAB model shows unstable areas primarily in areas of steep convergence.

The SMORPH model exhibited a high percentage of overlap with both the SHALSTAB and PISA-m models (90% and 69%, respectively), but also classified 3 times the area of unstable slopes in the study area. Nearly 60% of the area of unstable slopes classified by the SMORPH model exhibited no overlap with the other 2 models. The SMORPH model may over predict unstable slopes in the study area, and as a result incorporate a higher percentage of unstable slopes classified by the SHALSTAB and PISA-m models. Figure 25 illustrates the distribution of unstable slopes defined by SMORPH output values of “unstable” and “caution”. Obviously, the SMORPH model classifies the majority of convex and steep topography as “unstable”. Similar to the SHALSTAB model, the SMORPH model considers slope and topographic curvature as the main driving forces affecting slope stability. The model parameters and the instability calculation matrix utilized by SMORPH are very simplistic (Table 4 and Table 5). At the resolution of the 4-m LiDAR DEM, the SMORPH model does not appear to discriminate areas of instability, and as a result, any areas of moderate to high steepness and convexity are classified as “unstable”.

Overall, the SMORPH model appeared to over-predict unstable slopes in the study area, and should not be used in the development of future landslide hazard maps. The SHALSTAB and PISA-m models also did not show strong correspondence with each other, but based on the physics that each model embodies they may provide important insight when used together. Each model (SHALSTAB and PISA-m) might predict landslides missed by the other, even though their general lack of overlap is likely to cause an over-identification of unstable hillslopes. One could develop a three-tiered slope stability map showing: 1) unstable areas predicted by PISA-m, 2) unstable areas predicted by SHALSTAB, and 3) unstable areas predicted by both models**.** However, the integration of these three analyses will not be straightforward and will not result in a uniquely defined landscape hazard map.

E. Supplemental PISA-m analysis to test the effects of clearcut timber harvest activities

At the request of the NCRWQCB, the PISA-m model was run a second time in order to determine the effects of clearcut timber harvesting in the 2 study subwatersheds. The second PISA-m model run assumed a vegetation age of <15 years for all forested areas within the study subwatersheds. A vegetation age of “<15 years” is assumed to represent clearcut harvest conditions where root cohesive strength and tree surcharge are reduced to the lowest

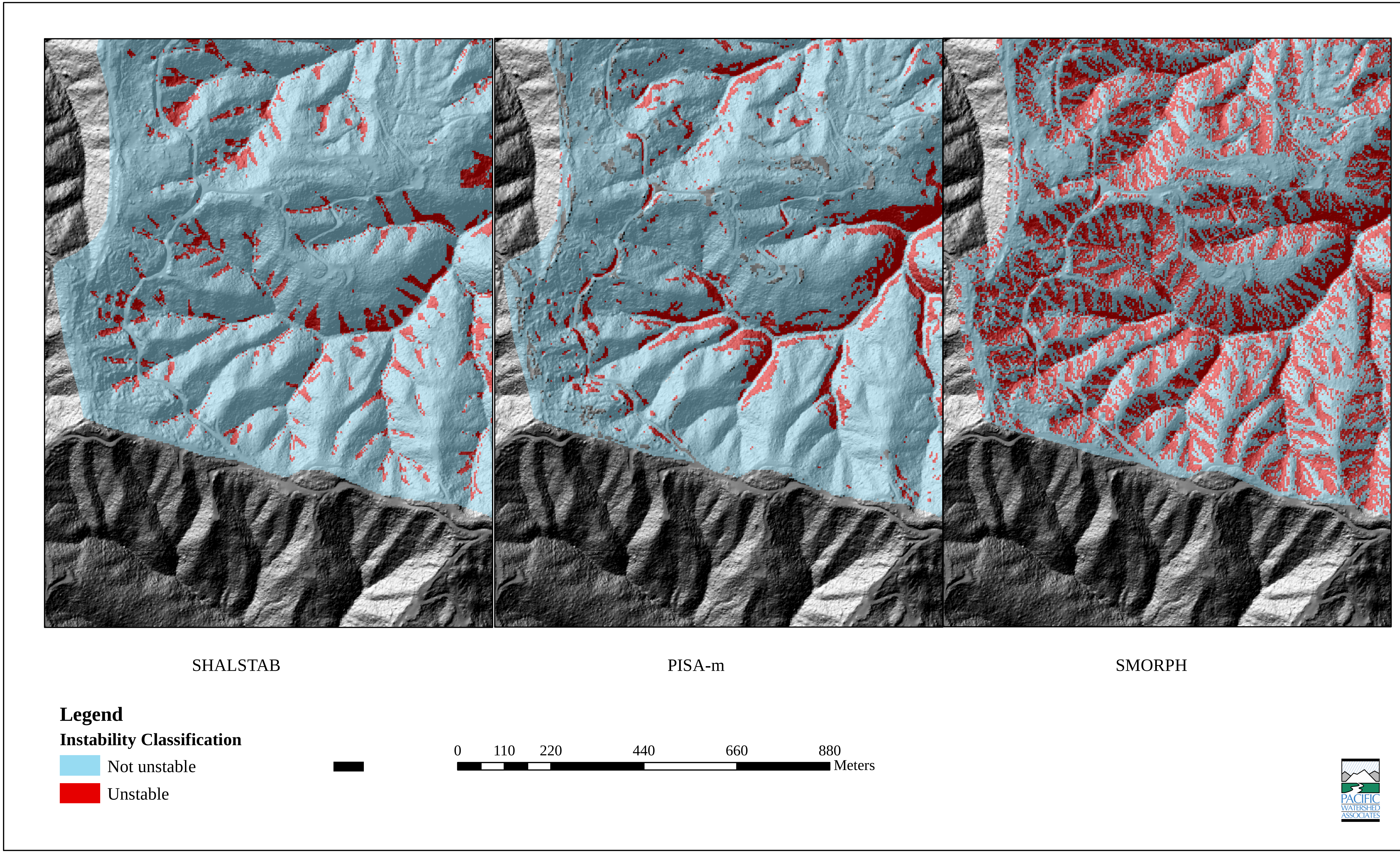


Figure 25. Model output by instability classification, Little Freshwater Creek

"minimum", "likely" and "maximum" values for forested lands as outlined in Table 3. The SHALSTAB and SMORPH models were not run a second time to test the effects of clearcut timber harvesting because they do not include root cohesive strength, tree surcharge, or vegetation age as model parameters.

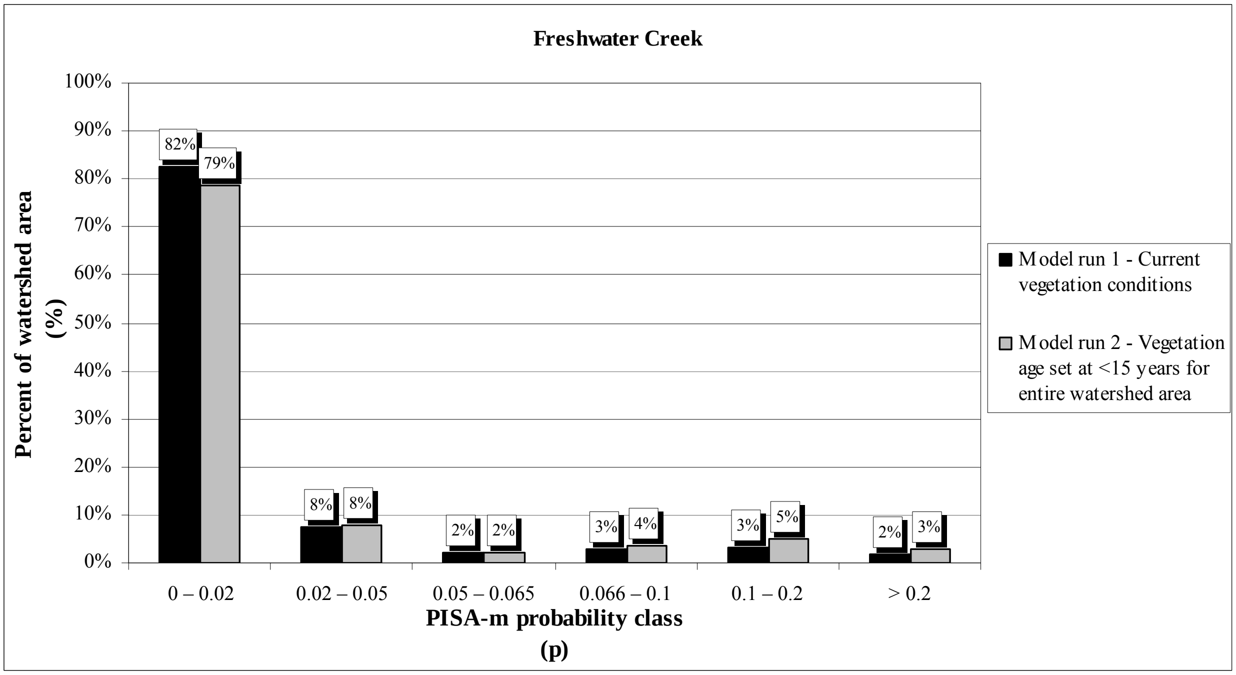
Table 16 shows the results for the second-run PISA-m model designed to simulate clearcut conditions (<15 year old vegetation age) for all forested lands within the Freshwater Creek subwatershed and the Ryan Slough planning watershed. Approximately 11% (13.08 km^2) of the 2 subwatersheds are classified as unstable (p>0.066), as defined in "Section IV-D-3-a" of this report (Table 16). This is a 54% increase in unstable area as compared to the results of the first model run, which identified 8% of the watershed area as "unstable" assuming current vegetation age conditions. According to the PISA-m second model run, increases of 57% and 102% in probability class area occurred in the "0.1 – 0.2" and ">0.2" PISA-m probability classes respectively. The only decrease in PISA-m probability class area was observed in the lowest probability class (p=0-0.02), in which changing the vegetation age to represent clearcut conditions resulted in an overall decrease in 5%.

Table 16. PISA-m model results for second model run assuming <15 year old vegetation age for all forested areas, Freshwater Creek and Ryan Slough planning watershed

Subwatershed		**PISA-m probability class**						
		0 – 0.02 (Low probability)	**0.02 – 0.05**	**0.05 – 0.065**	**0.066 – 0.1**	**0.1 – 0.2**	**> 0.2 (High probability)**	**Total**
Freshwater Creek	area (km^2)	62.62	6.11	1.77	2.82	4.02	2.41	79.75
	%	79%	8%	2%	4%	5%	3%	100%
Ryan Slough planning watershed	area (km^2)	31.62	2.24	0.65	1.03	1.46	1.34	38.34
	%	82%	6%	2%	3%	4%	3%	100%
Total	area (km^2)	94.24	8.35	2.42	3.85	5.48	3.75	118.09
	%	80%	7%	2%	3%	5%	3%	100%

Figure 26 compares the results of model run 1 (current vegetation conditions) versus model run 2 (clearcut conditions) for each subwatershed. Comparing the 2 study subwatersheds, the Ryan Slough planning watershed exhibited the greatest increase (79%) in "unstable" areas (p≥0.066) from 2.14 km^2 to 3.83 km^2. Freshwater Creek exhibited a 46% increase in "unstable" areas from 6.35km^2 to 9.25 km^2. Both subwatersheds exhibited a similar decrease of approximately 4% to 5% in "stable" areas (p<0.066) (Figure 26).

The results of the PISA-m second model run clearly show that clearcut timber harvest in the 2 subwatersheds can result in an increase in "unstable "areas classified by the PISA-m model. Although the total percent watershed area classified as "unstable" (p≥0.066) by both PISA-m model runs is relatively small (8.48 km^2 using model run 1 and 13.08 km^2 using model run 2) in



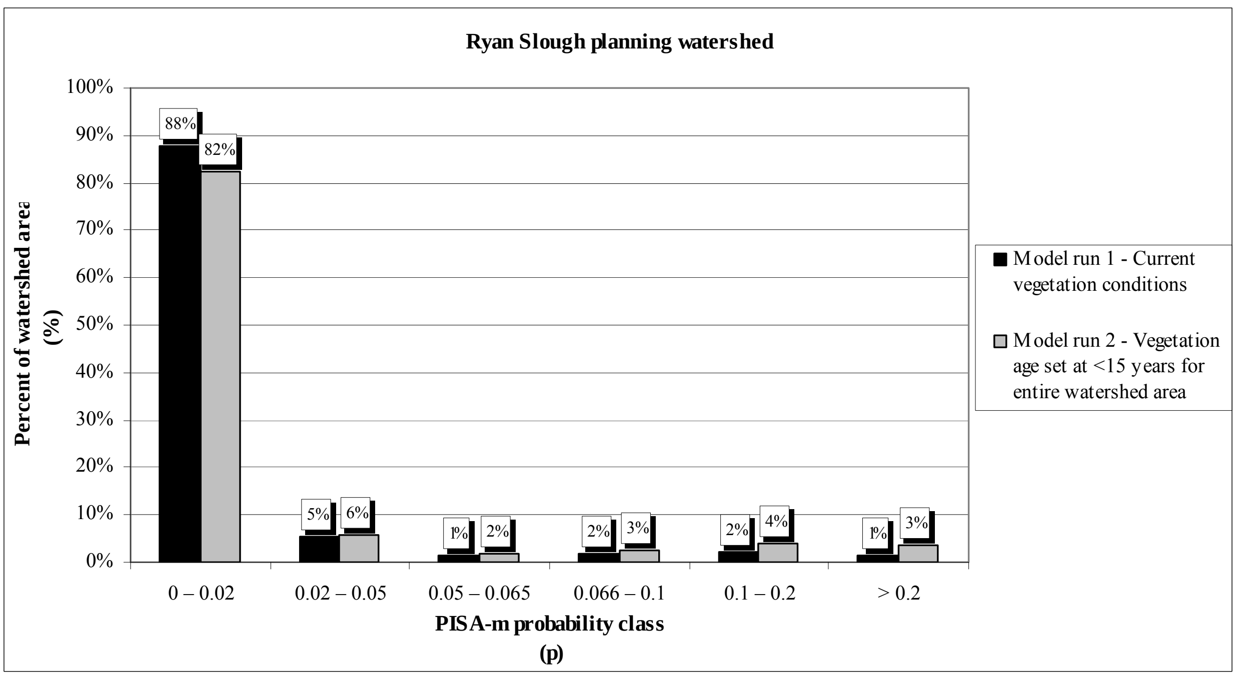


Figure 26 Comparison of percent watershed area by PISA-m probability class for 2 PISA-m model runs, Freshwater Creek and Ryan Slough planning watershed.

comparison to the entire watershed area (118.09 km^2), the percent increase in the area of the highest PISA-m probability classes (p = 0.01-0.02 and p>0.02) increased dramatically by 67% and 50%, respectively. The secondary analysis using a vegetation age of "<15 years old" to represent clearcut conditions may provide a useful tool in assessing slope stability in areas containing proposed Timber Harvest Plans.

V. Summary and Conclusions

Sediment inputs from mass wasting are an important issue in Coast Range watersheds of Northern California due to their impacts on water quality, fisheries habitat, and environmental degradation, as well as potential damaging impacts to property and the potential for injury or loss of life. As a result, it is important to have a good physical understanding of the past and current conditions of the watershed. The available predictive tools may then be employed to aid in determining potentially unstable areas that will require additional attention and inspection by resource agencies, land managers, and the public.

Three slope stability models (SHALSTAB, PISA-m, and SMORPH) were employed in the Freshwater Creek and Ryan Slough planning watershed to assist in categorizing the subwatersheds into landslide "hazard" or mass wasting susceptibility categories. The model outputs are intended to be used as a planning tool by resource agencies to identify potentially unstable areas, and provide guidance with regards to future land management.

According to the model outputs, the majority (between 75% and 85%, depending on the model) of Freshwater Creek and the Ryan Slough planning watershed are composed of stable terrain and terrain of low instability (Table 6, Table 7, and Table 8). Approximately 8% of the study area terrain was classified as moderate to high instability/probability by the SHALSTAB and PISA-m models. The SMORPH model classified approximately 25% of the study area terrain as "caution" or "unstable" (15% and 10%, respectively).

Air photo identified landslides with sediment delivery compiled as part of the Phase I Freshwater Creek TMDL sediment source assessment were used to evaluate the performance of the three models. In addition, a sample of randomly generated polygons was generated using ArcMap to further evaluate the three model output results. The sample of air photo identified landslides is an incomplete representation of the historic landslide record for the study area. As defined in the project scope, the air photo identified landslides used in the slope stability analysis only included slides from the 1987 to 2003 air photo time periods (26 air photo year record) and only slides with sediment delivery to streams. Landslides from previous air photo periods that bracket significant historic storms (i.e. 1964 storm) and landslides with no sediment delivery were not included in the study. The analysis of older air photo time periods and the inclusion of landslides with no sediment delivery may have improved model results.

Air photo identified landslides and randomly generated polygons were classified by instability/probability class using the presence of the highest instability/probability class within the landslide or random polygon. Overall, the three models appear to predict the air photo identified landslides better than the randomly generated landslides in the moderate to unstable stability classes (SHALSTAB classes: -10 –log (*q/T*) to -2.5 –log (*q/T*) and SMORPH classes: "Caution" and "Unstable") and in the p>0.066 PISA-m instability probability class (Figure 18, Figure19, and Figure 20).

A modified instability/probability classification method using a minimum threshold value of the median percent area of highest hazard area within the air photo identified landslide and random polygons was applied to further test the comparative performance of the SHALSTAB, PISA-m and SMORPH models. This modified and more rigorous classification method was applied to determine whether it is more accurate to classify landslides by the presence of the highest instability/probability class or by a minimum median percent area of the highest instability/probability class. The SMORPH model for both subwatersheds and PISA-m for Freshwater Creek appeared to perform better using the modified classification method as compared with the singular presence method (Figure 22). The SHALSTAB and PISA-m model runs for Freshwater Creek performed better than the model runs for the Ryan Slough planning watershed using the modified classification method. The different results from the two watersheds may be due to a smaller and less statistically significant population (N=63) of representative air photo identified landslides in Freshwater Creek as compared to the larger population (N=122) of air photo identified landslides observed in the Ryan Slough planning watershed.

The three model outputs were also compared to each other to determine how well they predicted air photo identified landslides and the degree of model overlap of unstable and stable areas. Overall, the SMORPH model correctly predicted the highest percentage (99%) of air photo identified landslides. Although the SMORPH model appears to be a "good" predictor of landslide occurrence, the model identified more than 3 times the area of unstable slopes in comparison to the SHALSTAB and PISA-m models. As a result, the SMORPH model appeared to overestimate the area of unstable slopes in the study area and therefore may not be a good slope stability model for the Freshwater Creek TMDL study area. Both SHALSTAB and PISA-m models correctly identified 75% of the air photo identified landslides (Table13).

Predictive performance of the three models was also measured from cumulative percentage graphs that plot the percentage of the study subwatershed area occupied by each slope instability or probability class against the percentage of air photo identified landslide area captured in each slope instability or probability class. .Overall, the cumulative percentage graphs suggest that all three models perform moderately well in comparison to one another. The SHALSTAB model performed only slightly better than the SMORPH and PISA-m models in Freshwater Creek and the Ryan Slough planning watershed. The PISA-m model performed slightly poorer than the SHALSTAB and SMORPH models in the Ryan Slough planning watershed.

As a final evaluation of model performance, the outputs of the three models were overlain using ArcMap to determine the amount of overlap of unstable and stable areas classified by each model. The SHALSTAB and PISA-m models exhibited only 25% of overlap of unstable areas classified by each model type (Table 14). The low percentage of overlap between the SHALSTAB and PISA-m models may be due to the difference in the application of infinite slope equations by each model. The SHALSTAB model uses a hydrologic-slope stability equation that uses a simulated variable steady state pore water pressure, but does not allow for spatial variability in other soil properties (i.e. soil thickness, friction angle, and cohesion). The model assumes that water accumulates in concave hollows where soils are thicker and exhibit higher pore water pressures.

In comparison, the PISA-m model as implemented in this project utilizes an infinite slope stability equation that does not consider topographic convergence as a driving factor in slope stability. The infinite slope equation used in the PISA-m model incorporates soil shear strength, pore water pressures, and tree root shear strength. The PISA-m model assumes that pore water pressure is random across the landscape. As a result, the two models identify very different areas of slope instability (Figure 25). Although the SHALSTAB and PISA-m models did not correlate well together, they both exhibited a high percentage of overlap with the SMORPH model unstable areas (93% and 69%, respectively).

Areas classified as unstable by the SMORPH model exhibited only 27% and 20% unstable area overlap with SHALSTAB and PISA-m, respectively. Nearly 60% of the slopes classified as unstable by the SMORPH model exhibited no overlap with the other two models. This suggests that SMORPH may be over predicting unstable slopes in the TMDL slope stability study area. The SMORPH model is based on a simplistic algorithm using slope and topographic convergence as the main driving forces affecting slope stability. At the resolution of the 4-m LiDAR DEM, the SMORPH model did not discriminate areas of slope instability, and as a result any area of moderate-steep slope gradient and topographic convergence was classified as "caution" or "unstable". It is apparent that there is a significant difference in the each model's classification of unstable areas, and as a result the models do not correspond well to each other and do not appear to provide an accurate representation of landslide hazard risk in Freshwater Creek and the Ryan Slough planning watershed.

The PISA-m model was run a second time at the request of the NCRWQCB to evaluate the effects of clear cut timber harvesting in the 2 study subwatersheds. The model parameter for vegetation age was set at "<15 years" for all forested areas within the Freshwater Creek and Ryan Slough planning watershed. A vegetation age of "15 years" simulates clear cut timber harvest conditions and assumes minimum root cohesive strength and tree surcharge values for forested areas of the 2 subwatersheds. By setting the age of forested areas to "<15 years, areas classified as "unstable" ($p>0.066$) increased by 54%. These results may provide an additional tool for regulators in assessing slope stability in areas with proposed Timber Harvest Plans.

A landslide hazard analysis is intended to provide information to resource agencies, landowners, and the public about potential hazard prone areas in the watershed. Landslide hazard maps developed from slope stability model outputs can be a beneficial tool when assessing areas intended for future land management activities and should be used in combination with field-based investigations by qualified professionals. Slope stability models are only as accurate as the assumptions used in their construction and the data that are used as input variables. The models themselves contain many assumptions and simplified conditions that adversely affect their performance alone, and in comparison with each other. Limiting factors that can affect the accuracy of slope stability models include: 1) model assumptions that are not appropriate for this terrain and climate, 2) inaccurate or incomplete watershed data inputs, 3) poor quality spatial data (e.g. inaccurately mapped air photo identified landslides used in model evaluations and comparisons), and 4) model outputs that are necessarily tested against already failed landscapes. As a result of these limiting factors, landslide hazard maps derived solely from slope stability model output can be inaccurate (either under predicting, over predicting or inaccurately predicting unstable landforms) and unrepresentative of actual watershed characteristics and conditions. For this reason, models are best employed as one of a number of possible tools that are most appropriately used in combination to produce a generalized mass wasting hazard map.

The three slope stability models (SHALSTAB, SMORPH, and PISA-m) employed in the Freshwater Creek TMDL slope stability study did not provide conclusive or reliable results by themselves, or in combination, that could be directly used in their present state to develop an accurate landslide hazard map for Freshwater Creek and the Ryan Slough planning watershed. Further refinement of the slope stability models and their input parameters (i.e. pore water pressure and the angle of internal friction) would be necessary before any attempt to develop a landslide hazard map from model output is made. The most effective hazard zonation for the study area might employ the results of the probabilistic and deterministic modeling in combination with more traditional geomorphic analyses and field-based observations.

VI. Limitations of analysis

The use of slope stability models in coastal watersheds in Northern California is problematic. In spite of the theoretical strengths of each model, the poor correspondence between the performance of the three models (SHALSTAB, PISA-m and SMORPH) employed in this study does not provide confidence in the choice or utility of the models in defining landslide hazard areas. The SMORPH and SHALSTAB slope stability models were developed in very different terrains and lithologies as compared to Freshwater Creek and the Ryan Slough planning watershed, and rely on local slope and topographic convergence as significant variables that define slope instability. Their performance is not surprising when, according to the Phase I Freshwater Creek TMDL sediment source assessment report, shallow landsliding (landslides delivering to streams) in the study area is dominated by inner gorge and planar steep streamside debris landslides (91%), and very few debris flows (9%). In addition, only 28% of the air photo identified landslides occurred on convergent hillslopes (i.e. swales and headwall areas). The majority (72%) of the landslides occurred on inner gorge and steep streamside (>65% gradient); streamside (<65% gradient); and other planar slope locations.

According to Dietrich et al. (1998), the SHALSTAB model does not perform well in coastal watersheds in Northern California due to the poor representation of the ratio of slope and the degree of topographic convergence. The model developers suggest delineating inner gorge areas and other steep planar zones from aerial photography or by field mapping, and automatically assigning these areas with a high hazard rating. If models that define slope instability based on slope and topographic convergence do not perform well in these watersheds, then it is apparent that the simple geomorphic indicators of slope steepness and inner gorge and streamside locations may be better predictors of landslide hazard areas.

The lack of correspondence of the three models may also be due to a variety of other factors. The Freshwater Creek TMDL slope stability modeling validation tests used a relatively small population (N=185) of air photo identified landslides from a short time span (26 years). In addition, the 185 air photo identified landslides only include landslides that delivered sediment to streams. Landslides with no sediment delivery were not mapped or evaluated as part of the Phase I Freshwater Creek TMDL sediment source assessment. In the 1999 Freshwater Creek Sediment Source Investigation (PWA 1999) and the PALCO Freshwater Creek watershed analysis (PALCO 2001), 443 landslides were identified between 1942 and 1997 air photo time periods. Of the 443 landslides, 21% were identified as landslides with no sediment delivery to streams.

The Freshwater Creek TMDL slope stability model validation tests may have been more conclusive had all 443 air photo identified landslides with and without sediment delivery from the PALCO Freshwater Creek watershed analysis been included in the modeling analysis for this study. In addition, the Phase I Freshwater Creek TMDL air photo analysis for the Ryan Slough planning watershed could have been expanded to incorporate both delivering and non-delivering landslides from the 1942 -2003 air photo time period. Although not within the scope of this study, a larger population of air photo identified landslides might have resulted in more accurate and more useful slope stability model validation results.

Slope stability modeling validation tests rely upon the prediction of past landsides mapped from historic aerial photography. This method involves overlaying the landslide polygon over the slope stability model output and classifying the landslide polygons by the presence of the highest instability/probability class in the landslide polygon. The landslide polygons represent landslide voids and their perimeters are usually delineated by steep head scarps and lateral landslide scarps. The topographic expression of past landslide voids can be significantly different from the adjacent “natural” hillslope topography. At the resolution of the 4-m LiDAR DEM, the slope stability models could classify the steep landslide scarps or steep landslide faces as high hazard areas. Landslide voids may be classified as high hazard areas because they may indicate the potential for landslide reactivation. This may result in the prediction of where landslides may reactivate, but does not necessarily represent where future landslides would likely occur. If past landslide voids are used solely to evaluate the performance of the slope stability models, then this could result in an over-prediction of landslides in high landslide instability/probability classes. For example, in the case where landslide voids are classified as unstable, but the surrounding “natural” slopes are classified with a lower instability/probability class, a more accurate validation test would involve comparing the model outputs with landslides generated in the next major storm event. Using this method, pre-landslide slope hazard classification can be verified by actual landslide occurrence. The existence of the high resolution LiDAR DEM for Freshwater Creek now provides an accurate topographic picture of the landscape against which future landslides can be more accurately analyzed.

Finally, the slope stability validation tests rely upon accurate mapping of air photo identified landslide polygons on the 1-m LiDAR DEM. Air photo identified landslides compiled as part of this study were identified during stereoscopic analysis of historic air photo pairs. Observed landslides were mapped on mylars overlain on the air photos as points at the mid point of the head scarp or polygons outlining the landslide void. Mapped landslides were then transferred and digitized “heads-up” as polygons on the 1-m LiDAR DEM using ArcMAP GIS. Even though landslide locations and void boundaries were delineated as best as possible on the LiDAR DEM coverage, there is likely some human error associated with the accurate spatial location of each landslide, as well as the "heads-up" digitizing process itself. In the future, air photo identified landslides used in validation tests should be field verified for accurate location using high resolution GPS surveying. Field investigations will also allow for the collection of more accurate landslide void dimensions and site conditions.

VII. Future analyses and recommendations

Further refinement of the slope stability model variables and parameters is recommended before using model outputs to directly develop landslide hazard maps for Freshwater Creek and the Ryan Slough planning watershed. Recommendations for further analyses include:

1) Develop landslide hazard maps in part using slope stability models that incorporate assumptions and variables that are suited for Freshwater Creek and the Ryan Slough planning watershed. For example, the SHALSTAB model and SMORPH models may not be appropriate for the study watersheds due to the model assumption of topographic convergence as one of the primary factors driving slope stability. Studies conducted in Freshwater Creek and nearby watersheds (PWA 1998(a); PWA 1998(b); PWA 1999) suggest that slope steepness and slope position are the main driving factor affecting slope stability. Thus, effective hazard zonation for the study area might employ the results of modeling in combination with more traditional geomorphic analyses and field-based observations.
2) The PISA-m model may be a better slope stability model for the Freshwater Creek TMDL slope stability study area because it does not consider topographic convergence as a main driving factor for slope instability. In addition, the PISA-m model offers greater flexibility because it can be set up to model extreme conditions, such as extreme saturation or a large earthquake. The historic air photo landslide inventories and the SHALSTAB or SMORPH models can not model extreme conditions. The PISA-m model coupled with a geomorphic or landscape sensitivity assessment may provide a better approximation of landslide risk hazard for this study area. The landscape sensitivity assessment discussed in the PALCO SYP/HCP (PALCO 1999) may also be an appropriate and useful model for identifying areas of potential landslide hazard or sensitivity and involves delineating the landscape according to sensitivity ratings of very low to very high. The sensitivity ratings are derived from GIS analysis of slope steepness, slope position, bedrock geology, active landslide terrain, and unstable and erodible soils (e.g. Atwell) (PALCO 1999). The PISA-m model and a landscape sensitivity assessment are tools that can be used to educate an overall depiction of mass wasting hazard for the study area that is eventually verified or confirmed at the project level through the application of professional field observation and analysis.
3) The evaluation of future slope stability models employed in the Freshwater Creek TMDL slope stability study area should incorporate all available air photo identified landslides observed and mapped from existing studies, including landslides with and without sediment delivery to streams. Complete air photo histories spanning the available historic aerial photography (1942 – present) can provide a more complete landslide data set for comparison to the slope stability model outputs. Any areas within the study area that do not have existing past landslide data should be analyzed for landslides with and without

sediment delivery over the entire historic air photo period (1942-present). Landslide locations should be GPS-located wherever possible.

4) The evaluation of model performance should include future landslides that will occur in the next major storm event. This will determine the model's performance comparing accurate pre-landslide topography with actual landslide occurrence. This can easily be done by analyzing the available imagery from the National Agricultural Imagery Program (NAIP) provided annually by the U.S. Department of Agriculture.
5) If slope stability models are employed to produce landslide hazard maps, then they should only be used as one tool among several to identify potential locations of instability. Areas identified as "unstable" should be inspected by qualified professionals before management decisions are made.

VIII. References

Abe, K., and Ziemer, R.R., 1991, Effect of tree roots on a shear zone: modeling reinforced shear stress. Canadian Journal Forest Research, Volume 21, Number 7, pp. 1012-1019.

Benda, L., Bigelow, P. and Worsley, T, 2002, Recruitment of wood to streams in old- growth and second-growth redwood forests, northern California, U.S.A. Canadian Journal of Forest Research 32, p. 1460-1477.

California Department of Forestry and Fire Protection, 2001, California Vegetation (CALVEG & WHR) GIS, Available online: fttp://frap.cdf.ca.gov/data.

Carver, G.A., 1987, Late Cenozoic tectonics of the Eel River basin region, coastal northern California, *in* Schymiczek, H., and Suchland, R., eds., Tectonics, sedimentation and evolution of the Eel River and coastal basins of northern California: San Joaquin Geological Society Miscellaneous Publication 37, p. 61-72.

Carver, G.A. and Burke, R.M., 1992, Late Cenozoic Deformation on the Cascadia Subduction Zone in the Region of the Mendocino Triple Junction, *in* Burke, R.M., and Carver. G.A., eds., A look at the southern end of the Cascadia Subduction Zone and the Mendocino Triple Junction: Pacific Cell, Friends of the Pleistocene field trip guidebook, p. 31-63.

Clarke, S.H., Jr., 1992, Geology of the Eel River Basin and adjacent region: Implications for late Cenozoic tectonics of the southern Cascadia Subduction Zone and Mendocino triple junction: American Association of Petroleum Geologists Bulletin, v. 76, p. 199-224.

Dietrich, W. E., Bellugi, D., Real de Asua, R., 2001, Validation of the shallow landslide model, SHALSTAB, for forest management, *in* Land use and watersheds: human influence on hydrology and geomorphology in urban and forest areas, American Geophysical Union, p. 195-227.

Dietrich, W. E., and Montgomery, D.R., 1998, SHALSTAB: a digital terrain model for mapping shallow landslide potential, Prepared for publication as a technical report by NCASI.

Harden, D.R., Colman, S.T,, and Nolan, K.M., 1995, Geomorphic Processes and Aquatic Habitat in the Redwood Creek Basin, Northwestern California in the Redwood Creek basin; in Geomorphic processes and aquatic habitat in the Redwood Creek basin, northwestern California; U.S. Geological Survey Professional Paper 1454, Nolan, M., Kelsey, H., and Marron, D., eds., p. G1-11.

Haneberg, W.C., 2004, A rational probabilistic method for spatially distributed landslide hazard assessment, Environmental and Engineering Geoscience X, p. 27-43.

Haneberg, W.C., 2005, PISA: probabilistic infinite slope analysis, User manual, Version1.0, Haneberg Geoscience.

Hunt, R.E., 1994, Geotechnical Engineering Investigation Manual, McGraw-Hill, Inc., 674 p.

Keefer, D.K., 1984. Landslides caused by earthquakes. Bulletin of the. Geological Society of America 95, 406–421.

Kelsey, H.M., Coghlan, M.C., and Pitlick, J., 1995, Geomorphic analysis of streamside landsliding in the Redwood Creek basin; in Geomorphic processes and aquatic habitat in the Redwood Creek basin, northwestern California; U.S. Geological Survey Professional Paper 1454, Nolan, M., Kelsey, H., and Marron, D., eds., p. J1-J12.

Krogstad, F., 1995, A physiology and ecology based model of lateral root reinforcement of unstable hillslopes, M.Sc. thesis, University of Washington, Seattle, Wash., 47 p.

McLaughlin, R.J., Ellen, S.D., Blake, Jr., M.C. , Jayko, A.S., Irwin, W.P., Aalto, K.R., Carver, G.A., and Clarke, Jr., S.H., 2000, Geology of the Cape Mendocino, Eureka, Garberville, and Southwestern part of the Hayfork 30 x 60 Minute Quadrangles and Adjacent Offshore Area, Northern California, Miscellaneous Field Studies MF-2336, Version 1.0.

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

Montgomery, D. R., Sullivan, K., and Greenberg, H. M., 1998, Forest Regional test of a model for shallow landsliding, Water Resources Research 30, p. 943-955.

Nilsen, T.H., and Clarke. S.H., Jr., 1987, Geologic evolution of the late Cenozoic basins of northern California, *in* Schymiczek, H. and Suchsland, R., eds., Tectonics, sedimentation and evolution of the Eel River and associated coastal basins of northern California, San Joaquin Geological Society Miscellaneous Publication, n. 37, p. 15-29.

PWA (Pacific Watershed Associates), 1998(a), Sediment source investigation and sediment reduction plan for the Bear Creek watershed, Humboldt County, California, Unpublished report prepared for The Pacific Lumber Company, Pacific Watershed Associates, Arcata, CA.

PWA (Pacific Watershed Associates), 1998(b), Sediment source investigation and sediment reduction plan for the North Fork Elk River watershed, Humboldt County, California, Unpublished report prepared for The Pacific Lumber Company, Pacific Watershed Associates, Arcata, California.

PWA (Pacific Watershed Associates), 1999, Sediment Source Investigation and Sediment Reduction Plan for the Freshwater Creek Watershed, Humboldt County, California. Unpublished report prepared for the Pacific Lumber Company, Scotia, California.

PWA (Pacific Watershed Associates), 2006, Freshwater Creek TMDL Sediment Source Assessment Phase I. PHASE I, Unpublished report prepared for the Sanborn and North Coast Regional Water Quality Control Board, 75 p.

PALCO (Pacific Lumber Company), 1999, Habitat Conservation Plan, Prepared by the Pacific Lumber Company, Scotia Pacific Holding Company, and Salmon Creek Corporation.

PALCO (Pacific Lumber Company), 2001, Final Report: Freshwater Creek watershed analysis, Prepared by Pacific Lumber Company, Scotia, California.

PALCO (Pacific Lumber Company), 2003, Reconnaissance level forensic landslide investigation, Geology SOP-08 (Standard operating procedure 08), Version 1.0.

Reid, L.M., 1998, Review of the Sustained Yield Plan / Habitat Conservation Plan for the properties of The Pacific Lumber Company, Scotia Pacific Holding Company, and Salmon Creek Corporation, Unpublished report, USDA Forest Service Pacific Southwest Research Station Redwood Sciences Laboratory, 33 p.

Reid, L.M. and Dunne, T. 1996. Rapid Evaluation of Sediment Budgets. Reiskirchen, Germany, Catena Verlag GMBH, 164 p.

Reid, M.E., Ellen, S.D., Brien, D.L., de la Fuente, J., Falls, J.N., Hicks, B.G., and Johnson, E.C., 2007, Predicting Debris-Slide Locations in Northwestern California, In: Standiford, R.B., Giusti, G.A., Valachovic, Y., Zielinski, W.J., Furniss, M.J., technical editors, 2007. Proceedings of the redwood region forest science symposium: What does the future hold? Gen. Tech. Rep. PSW-GTR-194, Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture; p. 347-356.

Roering, J. J., Schmidt, K. M., Stock, J. D., Dietrich, W. E., and Montgomery, D. R., 2003, Shallow landsliding, root reinforcement, and the spatial distribution of trees in the Oregon Coast Range, Canadian Geotechnical Journal 40: p. 237-253.

Shaw S.C. and Vaugeois L. M, 1999, Comparison of GIS-based models of shallow landsliding for application to watershed management. TFW report PR10-99-001. Washington Department of Natural Resources, Olympia, WA, 110 p.

Sidle, R.C., 1992, A theoretical model of the effects of timber harvesting on slope stability, Water Resources Research 28: p. 1897-1910.

Stillwater Sciences, 2007, Landslide Hazards in the Elk River Basin, Humboldt County, California, Unpublished report prepared for the North Coast Regional Water Quality Control Board, 62 p.

Wieczorek, G.F., 1996, Landslide triggering mechanisms, in Turner, A.K., and Schuster, R.L., eds., Landslides: investigation and mitigation: Special Report 247; Transportation Research Board, National Research Council, p. 76-90.