

1) Background:

- Shapes of valleys are described using observed cross-valley elevation profiles
- Described shape is of interest to infer the dominant surficial process in the formation of the valley
- Glaciated valleys are U-shaped, in contrast to V-shaped valleys, formed by other processes
 - (earliest reference McGee (1894))
- Approaches to describing curvature
 - Polynomial and nonlinear regression models used in Geomorphology literature; most commonly used are versions of the power law regression model:
 - Power law model for each side separately (shape through value of exponent)
 - Generalized Power Law (GPL) model: $y = a|x - x_0|^b + y_0 + \epsilon$ (Pattyn & Heule, 1998)
 - Frequent violation of assumptions underlying models (e.g. heteroskedasticity)
- Method of data collection (of profiles) could pre-determine the conclusions reached

2) Previous Work (Greenwood & Humphrey, 2002)

A. Improved Sampling (new for Geomorphology literature)

- Within each valley, a relatively homogeneous section is defined
 - (blue "center" lines in column 2 of Figure 1)
- A location on this center line is **randomly** selected
 - Allows inference to entire sampling region
- Profiles at different angles through central location are selected
 - (green lines in column 2 of Figure 1)
- Minimum distance across the valley is used to determine the optimal profile for a given location
 - Choosing non-optimal profiles could bias results to U-shapes by extending width of base of valley
- Minimum of the ridge heights is used to determine the width of the valley
 - Comparison of results across studies is difficult due to inconsistent or unspecified sampling methods
- Greenwood and Humphrey (2002) contains further details on the selection of profiles and the effects on conclusions using nonlinear regression methods, illustrated in the Dhona Khola
 - Chosen wall height can affect conclusions (reduced wall height, more U shaped)
 - Orientation of cross-valley profile can affect conclusions

B. Nonlinear Regression and Model Selection

- Different nonlinear regression models were considered
 - modifications to the GPL and the catenary curve (new for Geomorphology literature)
- Model selection (new for Geomorphology literature) was discussed
- Assumptions underlying these models (new for Geomorphology literature)

Both allow more powerful conclusions than previously used methods

Drawbacks of Nonlinear regression:

Wall height can affect conclusions about valley shape
Models available may be too restrictive in shape

3) Data:

Himalayan Valleys used to illustrate methods:

Dhona Khola Drainage (Known to be glaciated, should be U-shaped)

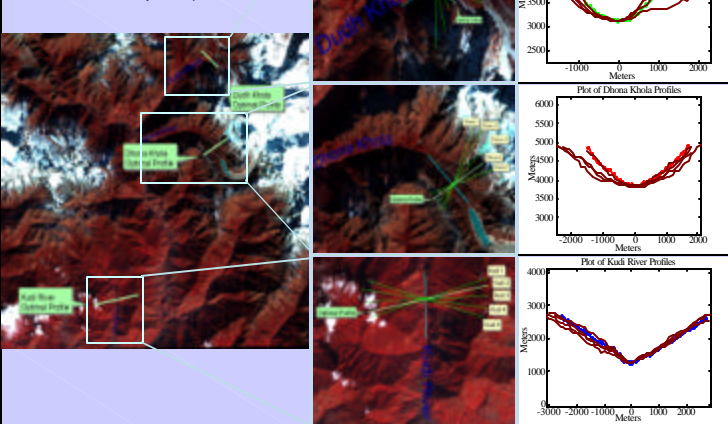
Dudh Khola Drainage (Known to be glaciated, should be U-shaped)

Kudi River Valley (Known to be an incised river valley, should be V-shaped)

Sampling from a Digital Elevation Map (DEM):

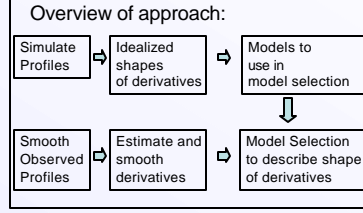
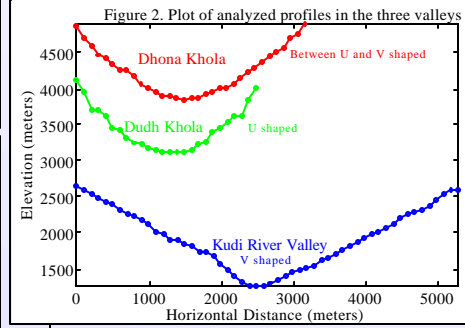
- Profile extraction tool in ArcView
- DEM has 100m x 100m pixel size, mean elevation on pixel used
- Elevation sampled every 100m on transect
- Analyzed Profiles: **Dhona: 33 observations**; **Dudh: 26 observations**; **Kudi: 54 observations**

Figure 1. Satellite image of valleys of interest, locations of profiles and sampled profiles (x-axis shifted to minimum of each profile).



Modeling Shapes of Valley Profiles.

Mark Greenwood (markg@uwyo.edu); PhD Student, Dept. of Statistics
Advisor: S. Huzurbazar, Dept. of Statistics; Committee Member: N. Humphrey, Dept. of Geology and Geophysics
University of Wyoming



4) Functional Data Analysis (FDA)

- Curvature/ shape can be described through derivatives of the profile
- Differencing to estimate the derivatives is not recommended
 - noise in observations can make derivative estimates difficult to interpret
 - "Although a common practice, derivative estimation by differencing is, in fact, seldom a good idea!" (p 41, Ramsay & Silverman, 1997)
- FDA in this context assumes a function generated the discrete observations
- We assume that the true function is smoother than observed values
- 6th Order B-spline basis is used to provide continuous derivatives
- 1st and 2nd derivatives should be relatively smooth
 - roughness penalized in the 4th derivative
 - interest lies in the shape of 1st and 2nd derivatives
- Estimate (a smooth) function for each profile as well as λ by minimizing the penalized residual sum of squares for a value of λ : $Penalized \text{ Sum of Squares} = \sum (y_j - x(t_j))^2 + \lambda \int \{D^4 x(s)\}^2 ds$
- The amount of smoothing (value of λ) can be chosen by minimizing the cross validation criteria (CV): $CV(I) = \sum_{i=1}^n (y_i - \hat{x}^{(i)}(t_i))^2$ where $\hat{x}^{(i)}(t)$ is the spline estimate at location t_i with that observation removed

Figure 3 displays the estimated functions and derivatives for different λ :

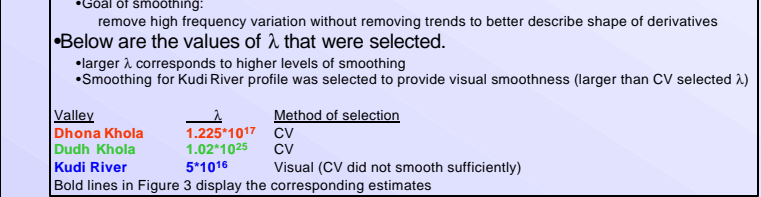
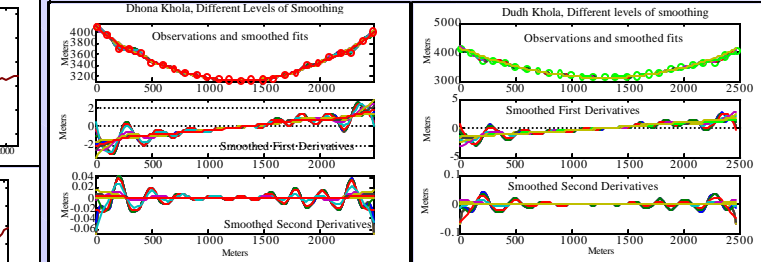


Figure 3. Plot of Observed and Smoothed Values, with Derivatives.



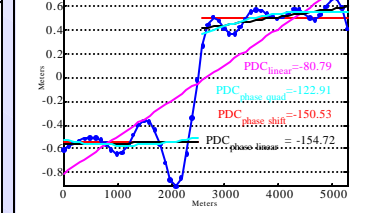
What do estimates of derivatives tell us about curvature of a profile?

- U shaped (simulated) profile:** Catenary curve used to generate U shaped profile
 - First Derivative: linear for slope of valley
 - Second Derivative: constant, small positive (slope of first derivative)
- Between U and V shaped (simulated) profile:**
 - First Derivative: phase change linear or quadratic reasonable approximation
 - Second Derivative: small bump at profile minimum, constant positive elsewhere
- V shaped (simulated) profile:**
 - First Derivative: jump discontinuity (small deviation from theoretical derivative: continuous function to estimate discontinuous)
 - Second Derivative: spike at function minimum (again, small deviation from theoretical derivative)
- Dudh Khola:** derivatives follow U shape closely
- Dhona Khola:** tends to U shape, but not as clearly as the Dudh Khola profile
- Kudi River Valley:** tends to V shape, but not perfectly; Jump discontinuity in first derivative, related spike in second derivative near function minimum

Modeling (discrete) 1st derivatives:

- Estimate of 1st derivatives at (discrete) locations of observation
- Smoothed as described above
- Regression models considered to describe first derivatives:
 - Theoretical V shape: two phase mean model
 - Theoretical U shape: linear regression model
 - Between U and V shape: approximated by two phase quadratic regression model
 - Also considering two phase linear regression model to allow asymmetric, but linear 1st derivatives
- Model Selection amongst the previous regression models for 1st derivatives:
 - Predictive Discrepancy Criterion (PDC) used to select the best model
 - Cross-Validation type model selection criteria (Davies and Cavanaugh, 2002)
 - Under the assumption of normal, independent errors for the models, the PDC = $\sum_{i=1}^n \ln S_{-i}^2 + \sum_{i=1}^n (y_i - \hat{y}_{i,-i})^2 / S_{-i}^2$, where $\hat{y}_{i,-i}$ is the fitted value for y_i and S_{-i}^2 is the MLE for the variance, both based on the data set with y_i removed.
 - PDC is used here as it has less restrictive assumptions than AIC

Figure 5. Plot of First Derivative Kudi River, with PDCs and Model Fits



Kudi River Profile (pictured on left):

- Phase change linear regression model is favored
- indicates distinct jump with a trend in right side of the derivative (Phase shift model also reasonable)
- Suggests profile is relatively V shaped
- Dudh Khola Profile (not pictured):** Phase change linear regression model is favored; Profile is U shaped with minor asymmetry
- Dhona Khola Profile (not pictured):** Phase change linear regression model is favored; All models relatively poor fitting; Describes profile as close to U shaped, but not as U shaped as Dudh Khola

Advantages of FDA methods over Nonlinear regression:

- Flexible (nonparametric) smoothing of observations does not bias results to certain shapes
- Restriction to portion of profile may not be as critical as with nonlinear regression models
- Other possible shapes and behaviors may be described using smoothing through derivatives
- Smoothing is now a part of the analysis and not part of pre-processing data as in Li, Liu, and Cui (2001)

5) Conclusions and Further Work

- Nonlinear regression and model selection criteria provide a method for describing curvature
- FDA methods provide useful estimates of derivatives
 - information about the curvature displayed and wall height not as influential (in the visual description of derivatives)
- Dudh Khola U shaped, Dhona Khola between U and V shaped (more U than V), Kudi River relatively V shaped**
- Model selection for the shape of 1st derivatives has potential for description/ comparison of curvature
- Determination of change from U to V in a valley possible by FDA methods and model selection**
- Some Future Work:**
 - Multiple profiles from a valley for use in Functional Linear Modeling (change in shape over the length of the valley)
 - Incorporation of covariate information (direction of valley, bedrock type, etc.)