Multi-Interval Data Mining of At-A-Station Hydraulic Geometry to Quantify Temporal Stability

Leong Lee

Department of Computer Science

Mark Lees

Gregory S. Ridenour

Department of Geosciences

Austin Peay State University

Clarksville, Tennessee, USA

Introduction

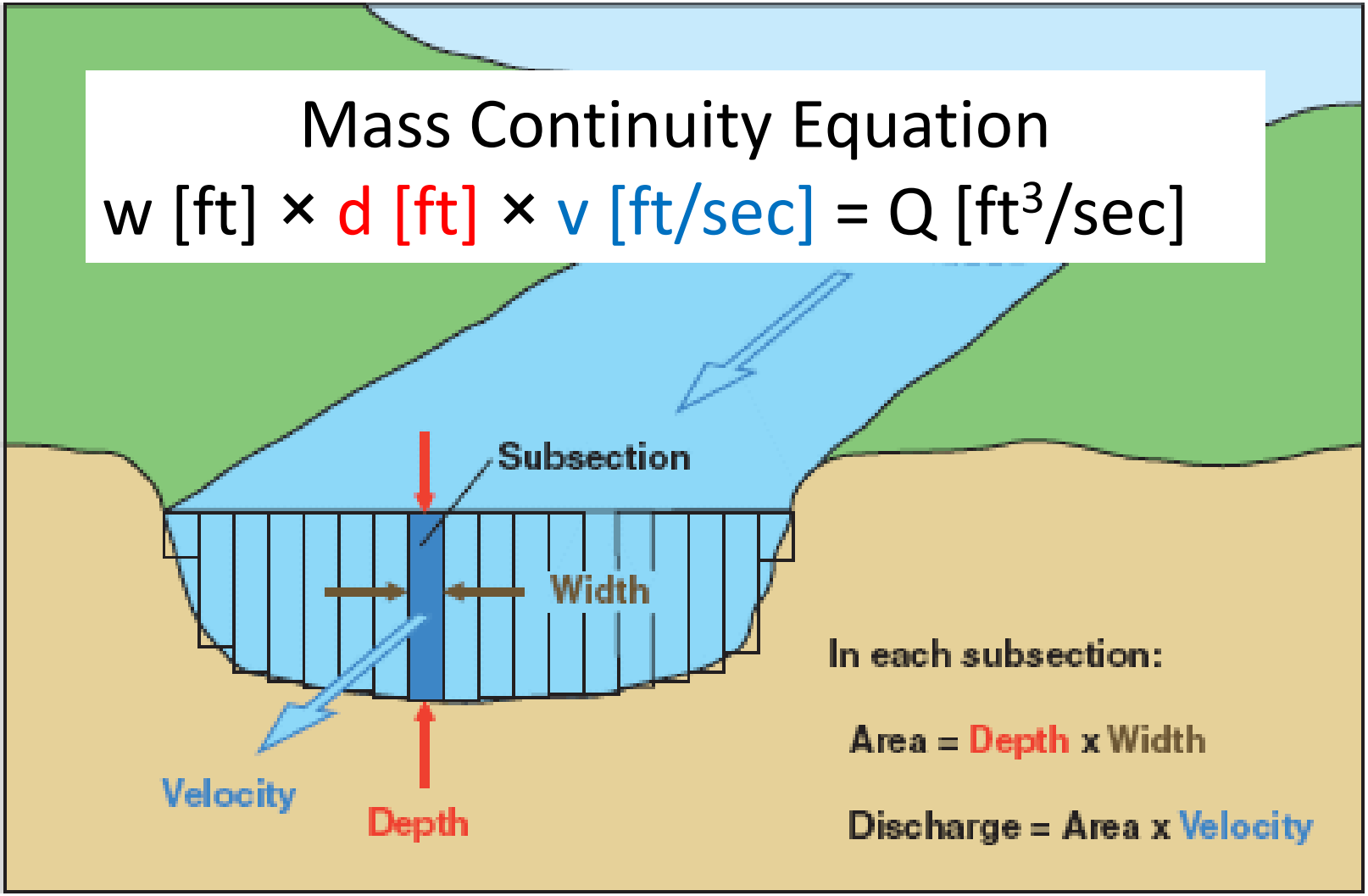
Hydraulic geometry equations express relationships between the *flow* (velocity and discharge) and *geometry* (width and depth) of a stream.

Variables may change in *time* at a single cross section (**at-a-station**) or in *space* along a stream channel for a given frequency of discharge (**downstream**).

Introduction

Data was mined in multiple, 2-year intervals to compute and classify hydraulic geometry, as well as to quantify stability.

The temporal stability of at-a-station hydraulic geometry at cross-sections from rivers in mountainous and non-mountainous watersheds in Tennessee was compared.



Current-meter discharge measurements are made by determining the discharge in each subsection of a channel cross section and summing the subsection discharges to obtain a total discharge.

Hydraulic Geometry

Width (w) $w = aQ^b$

mean depth (d) $d = cQ^f$

mean velocity (v) $x \quad v = kQ^m$

discharge (Q) $Q = ackQ^{b+f+m}$

Thus, it is evident that $a \times c \times k = 1$

$b + f + m = 1$

Hydraulic exponents are **compositional** (*unit-sum constrained*) **data**.

Linearized Power Functions

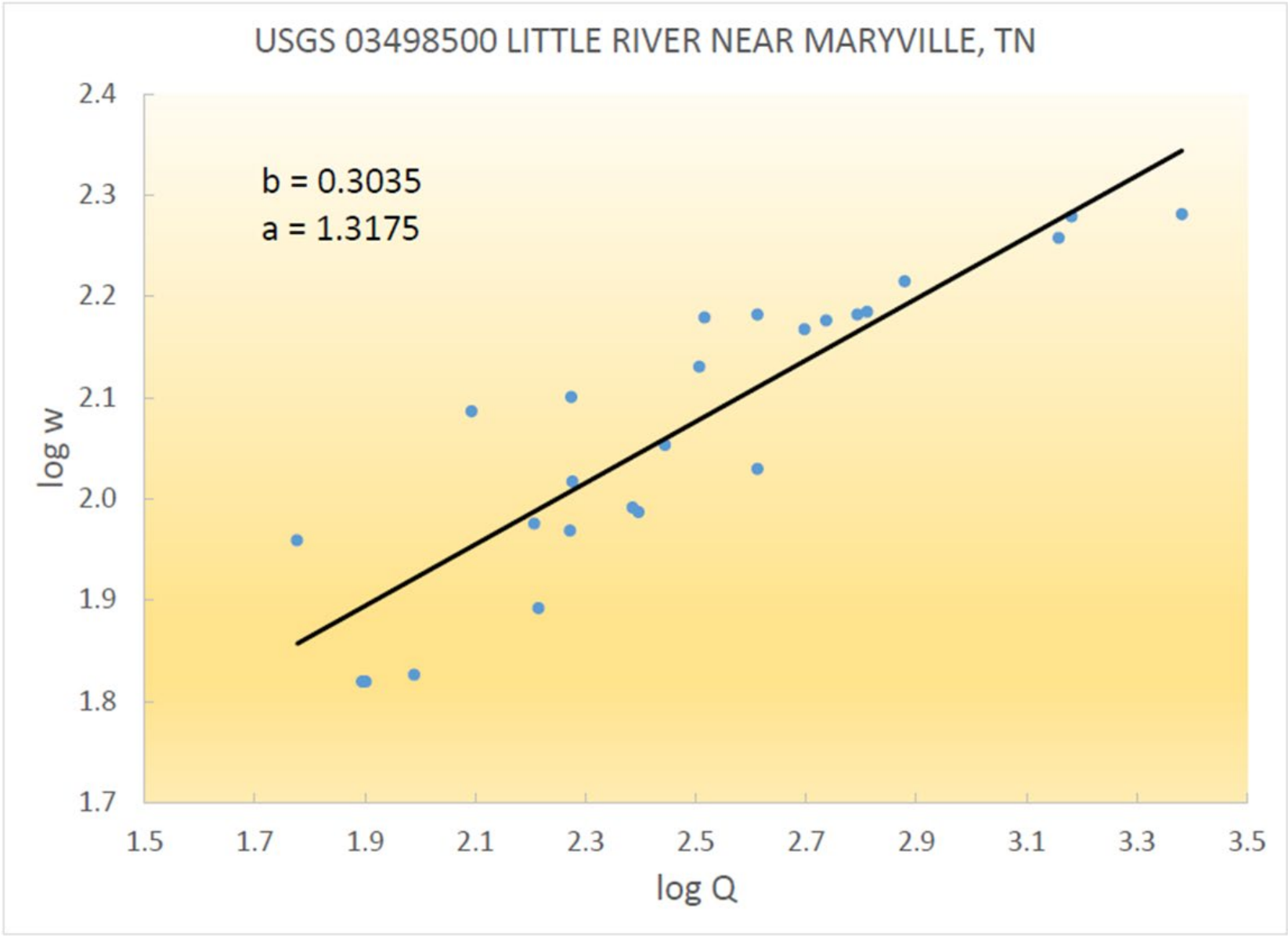
$$\log w = b \log(Q) + \log(a)$$

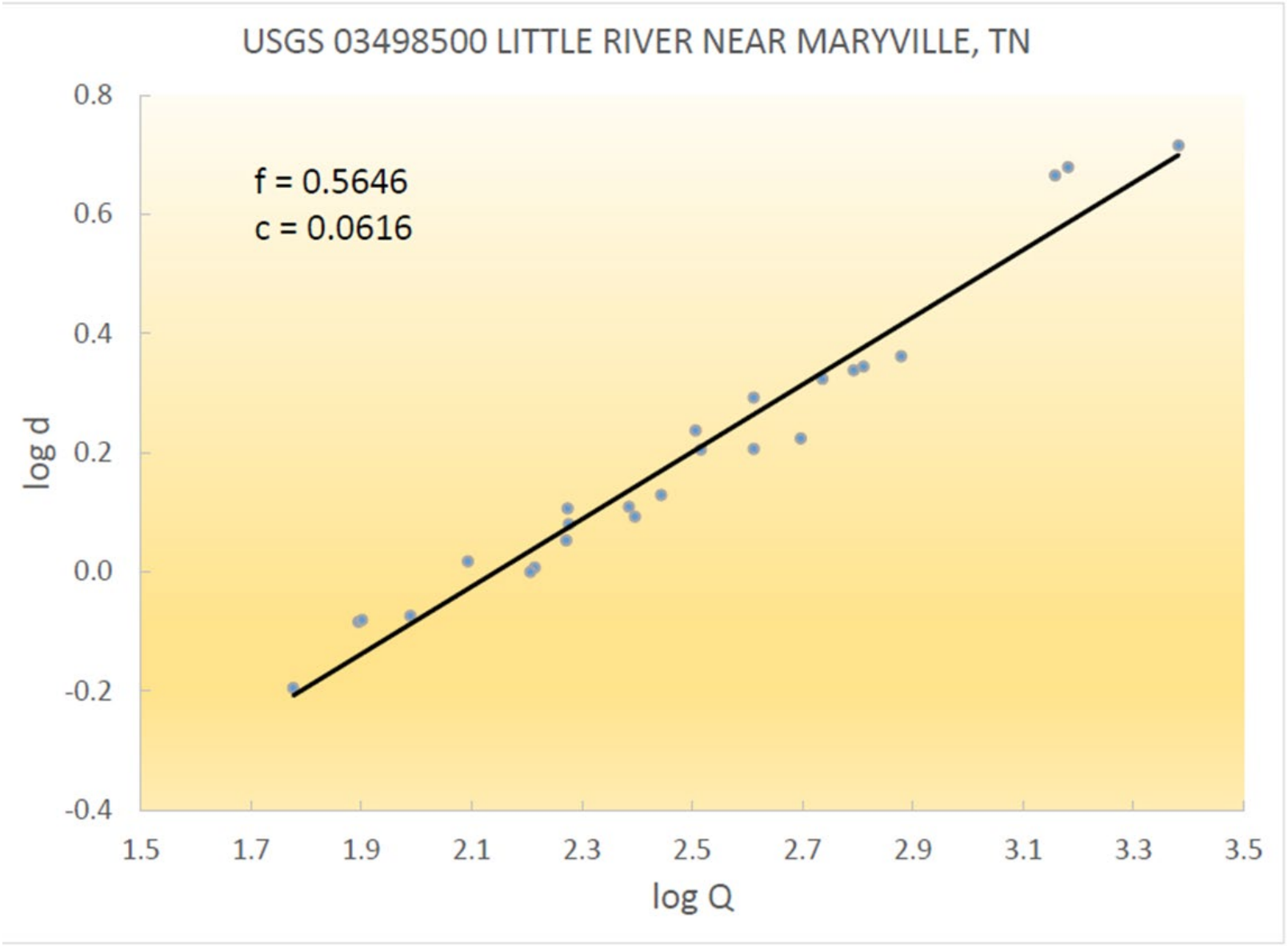
$$\log d = f \log(Q) + \log(c)$$

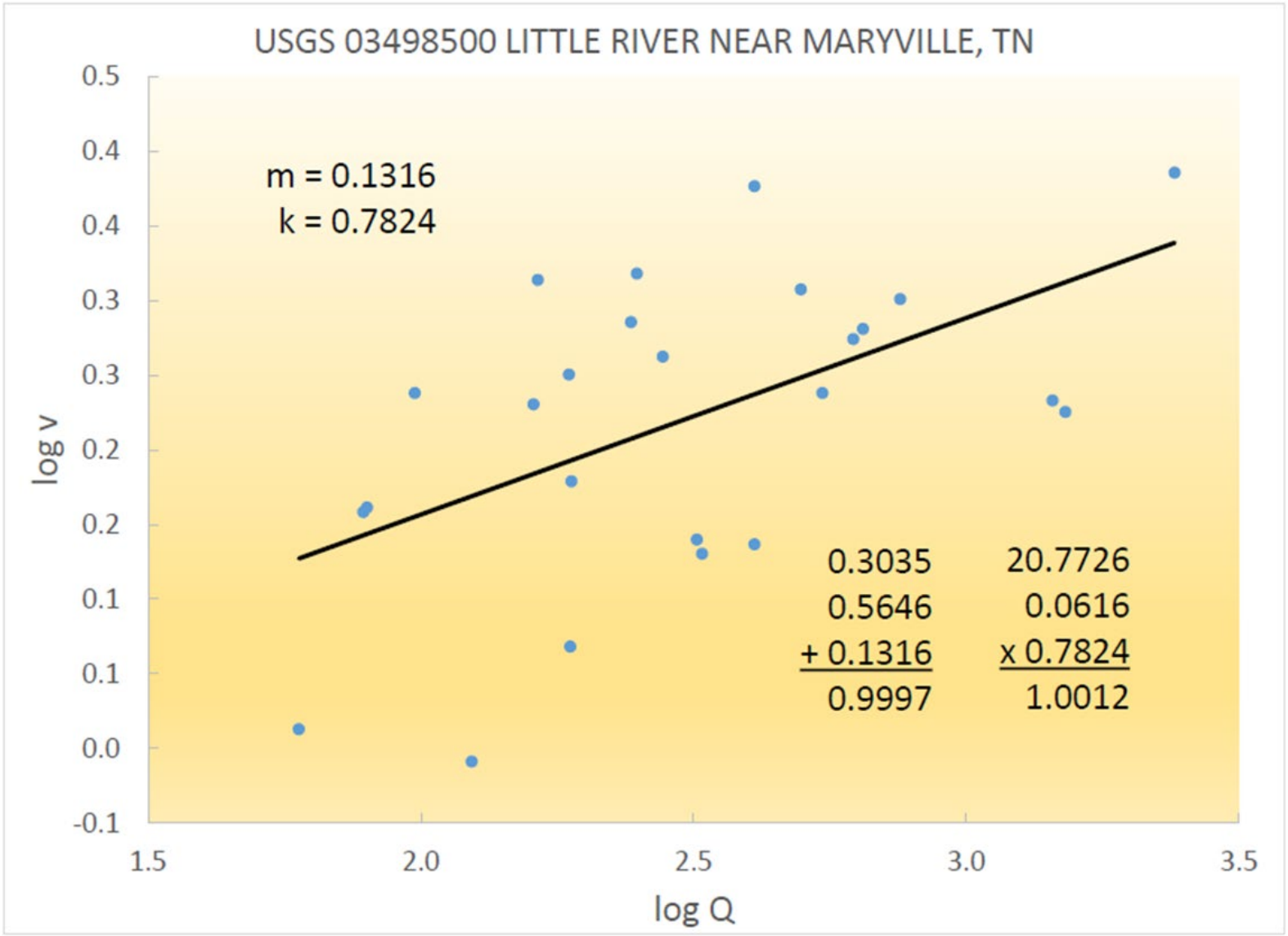
$$\log v = m \log(Q) + \log(k)$$

Linear regression of log-transformed data is used to determine

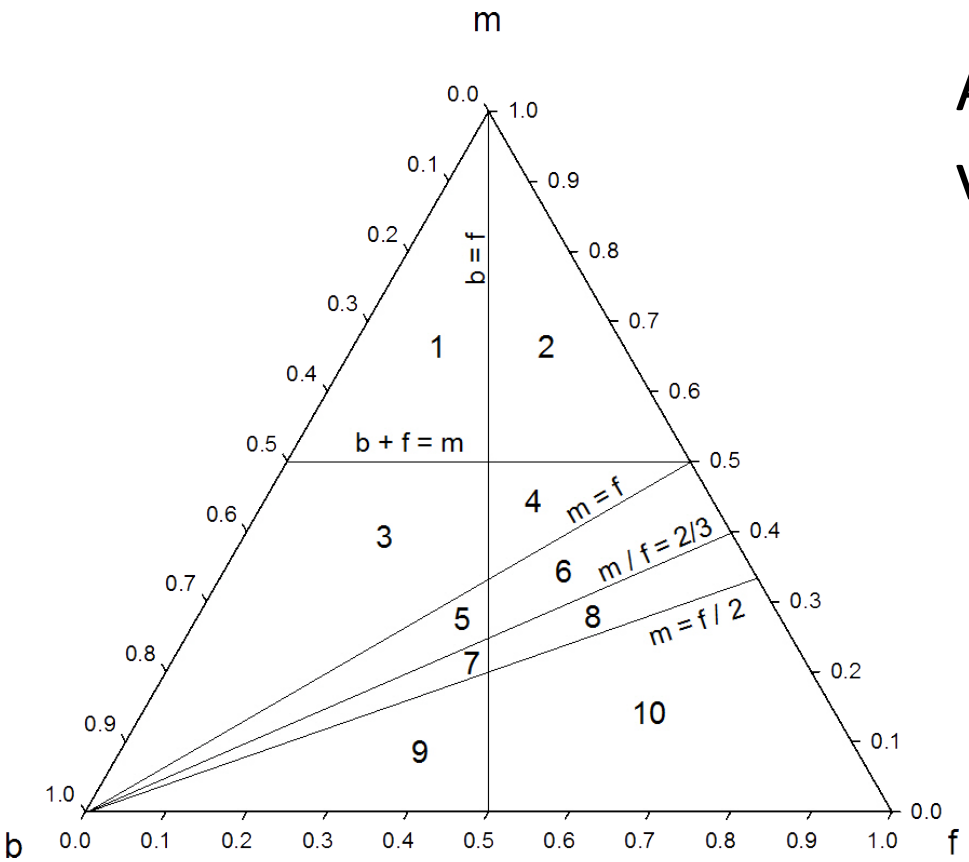
- a, c, and k (antilog of the y-intercept)
- b, f, and m (slope)


USGS 03498500 LITTLE RIVER NEAR MARYVILLE, TN
b = 0.3035
a = 1.3175
log w
log Q
2.4
2.3
2.2
2.1
2.0
1.9
1.8
1.7
1.5
1.7
1.9
2.1
2.3
2.5
2.7
2.9
3.1
3.3
3.5


USGS 03498500 LITTLE RIVER NEAR MARYVILLE, TN
f = 0.5646
c = 0.0616
log d
log Q
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
1.5
1.7
1.9
2.1
2.3
2.5
2.7
2.9
3.1
3.3
3.5


USGS 03498500 LITTLE RIVER NEAR MARYVILLE, TN
m = 0.1316
k = 0.7824
0.3035
0.5646
+ 0.1316
0.9997
20.7726
0.0616
x 0.7824
1.0012
log v
log Q
0.5
0.4
0.4
0.3
0.3
0.2
0.2
0.1
0.1
0.0
-0.1
1.5
2.0
2.5
3.0
3.5

Ternary Diagram (*b-f-m* diagram) and Channel Types



A barycentric plot, 3 variables sum to 1.

Ternary Diagram (*b-f-m* diagram) and Channel Types

Channel Type	Criteria
I	b + f < m AND b > f
II	b + f < m AND b < f
III	b + f > m AND b > f AND m > f
IV	b + f > m AND b < f AND m > f
V	f > m AND b > f AND m/f > 2/3
VI	f > m AND b < f AND m/f > 2/3
VII	m > f/2 AND b > f AND m/f < 2/3
VIII	m > f/2 AND b < f AND m/f < 2/3
IX	m < f/2 AND b > f
X	m < f/2 AND b < f

Rhodes Classification of Hydraulic Geometry.

Five hydrologically significant lines differentiate the ternary diagram into 10 fields (channel types).

Ternary Diagram (*b-f-m* diagram) and Channel Types

The **first line** is ***f* = *b***. If a point plots to the **left** of the line (*b* > *f*), the width-depth ratio **(*w*/*d*) increases** with **increasing discharge**; to the **right** of the line (*b* < *f*) the ratio would **decrease**.

The second line is *m* = *f*. If a point plots above the line (*m* > *f*), competence (the largest particle size a stream can transport) should increase with increasing discharge; below the line competence should decrease.

The third line is *m* = *f* /2. If a point plots above the line (*m* > *f*/2), the Froude number (which differentiates supercritical and subcritical flow) increases with increasing discharge; below the line the Froude number decreases.

The fourth line is *m* = *b*+*f*. If a point plots above the line (*m* > *b*+*f*), velocity increases more rapidly than cross-sectional area with increasing discharge; below the line it decreases.

The fifth line is *m*/*f* = 2/3, which is related to the Manning equation. If a point plots above the line (*m* > (2/3) *f*), the ratio of the square root of slope (*s*) to the roughness coefficient ($s^{½}/n$) increases with increasing discharge; below the line it decreases.

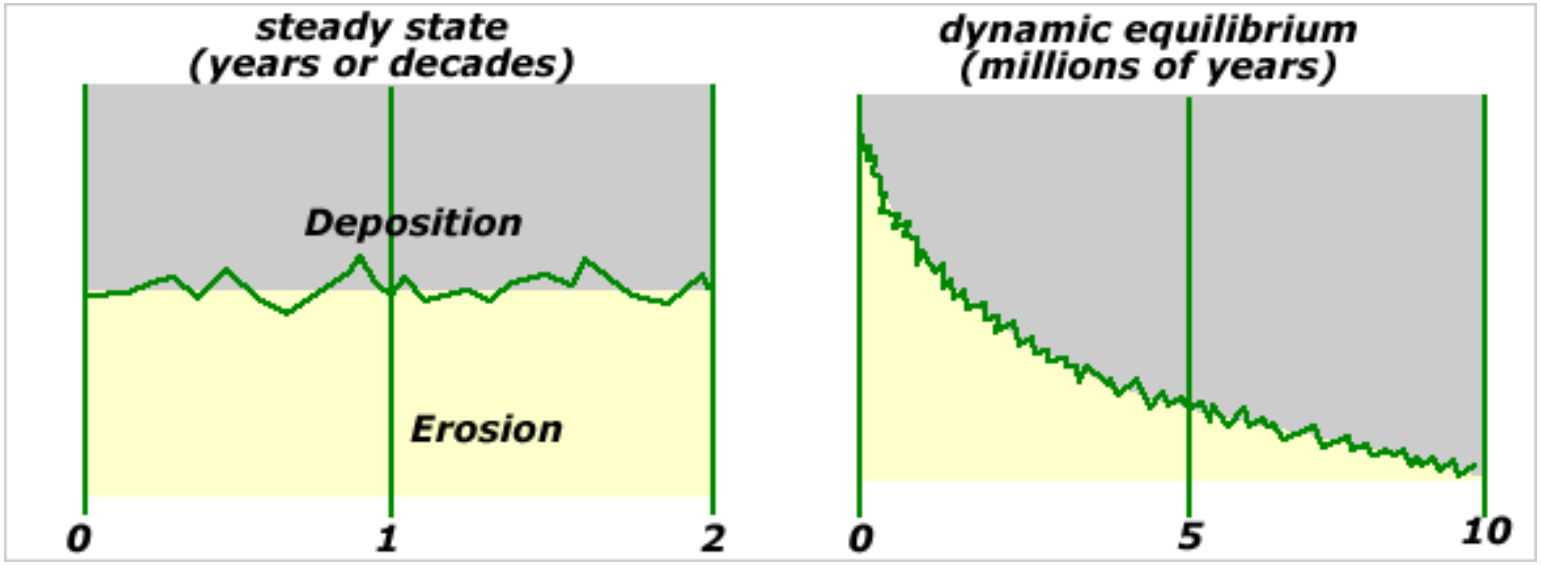
Landform Equilibrium and Stability

Landforms adjust, or self-regulate, in response to available energy and mass input.

Equilibrium state can be perceived as changes in the state of a system as a function of time (Fig. 1).

As shown in Fig. 2, equilibria can also be classified as (a) stable, (b) metastable, and (c) unstable.

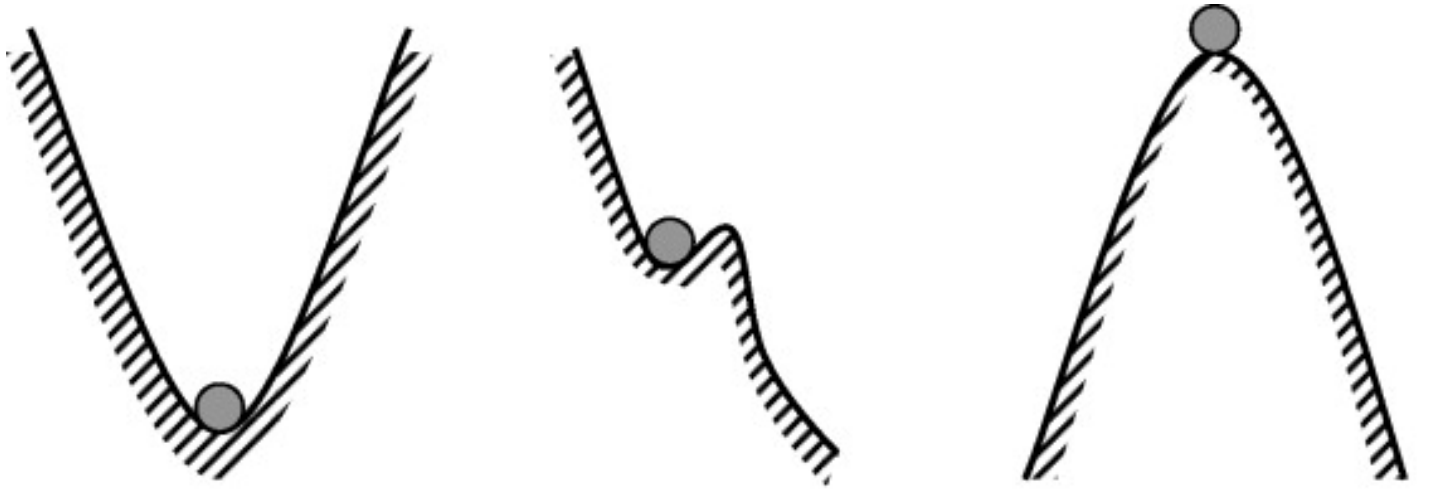
Landform Equilibrium and Stability



Static state equilibrium: fluctuation about a constant mean

Dynamic equilibrium: fluctuation about a moving average, or trend line

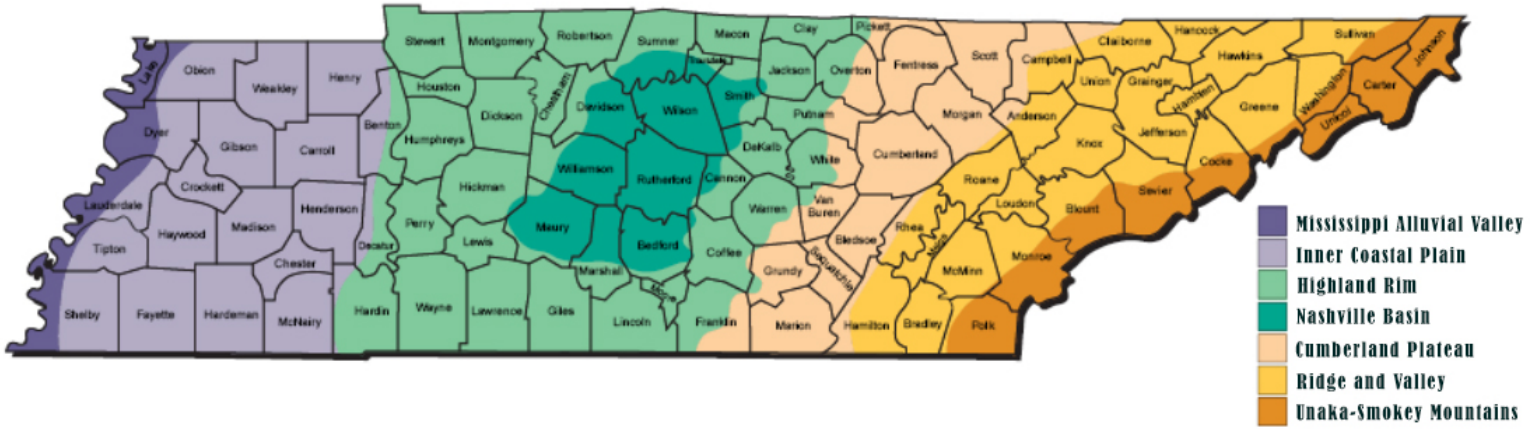
Landform Equilibrium and Stability



(a) stable equilibrium. (b) metastable equilibrium. (c) unstable equilibrium.

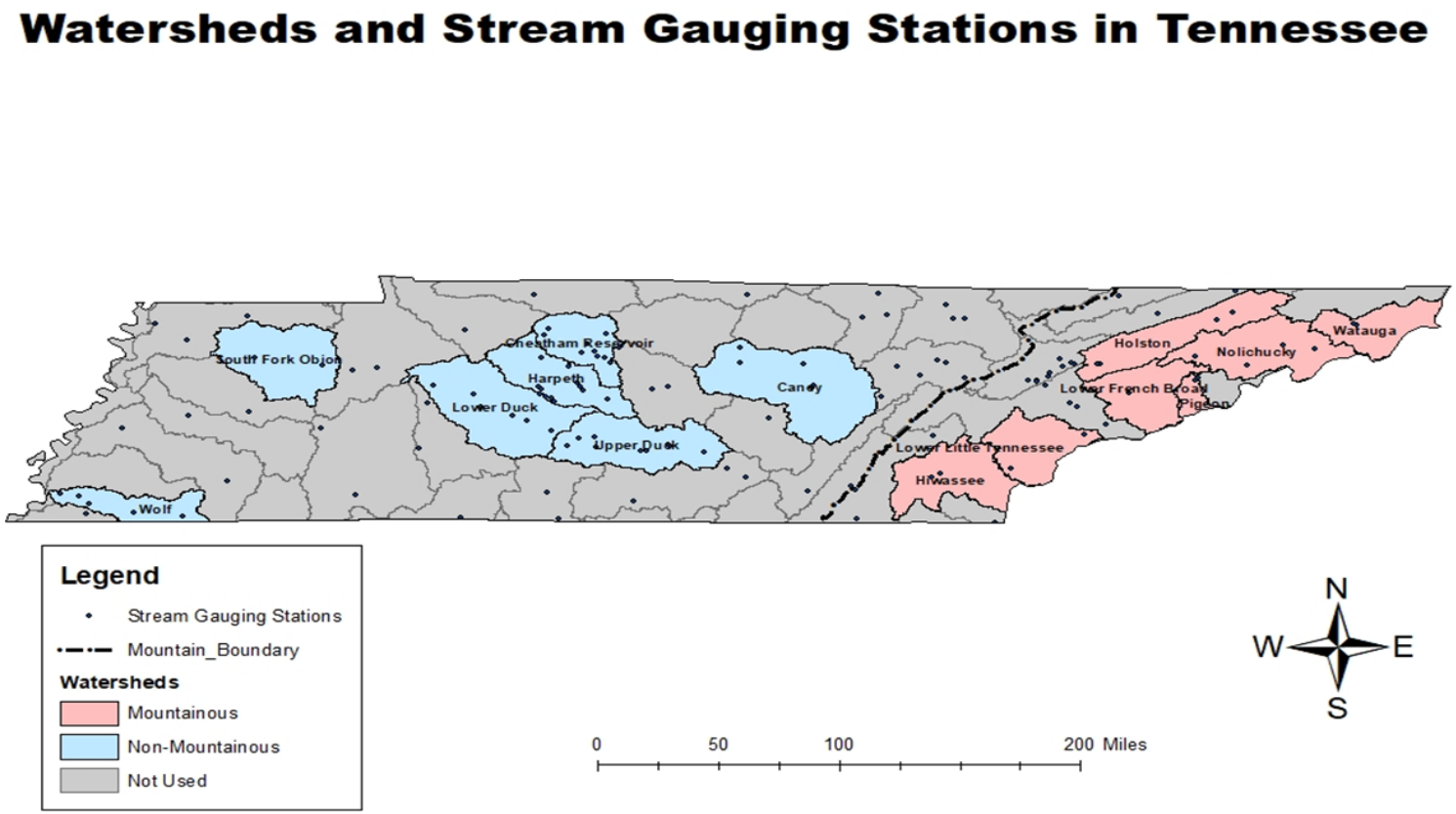
Position of the ball represents a system state, i.e. a particular set of (b, f, m) values.

Study Area

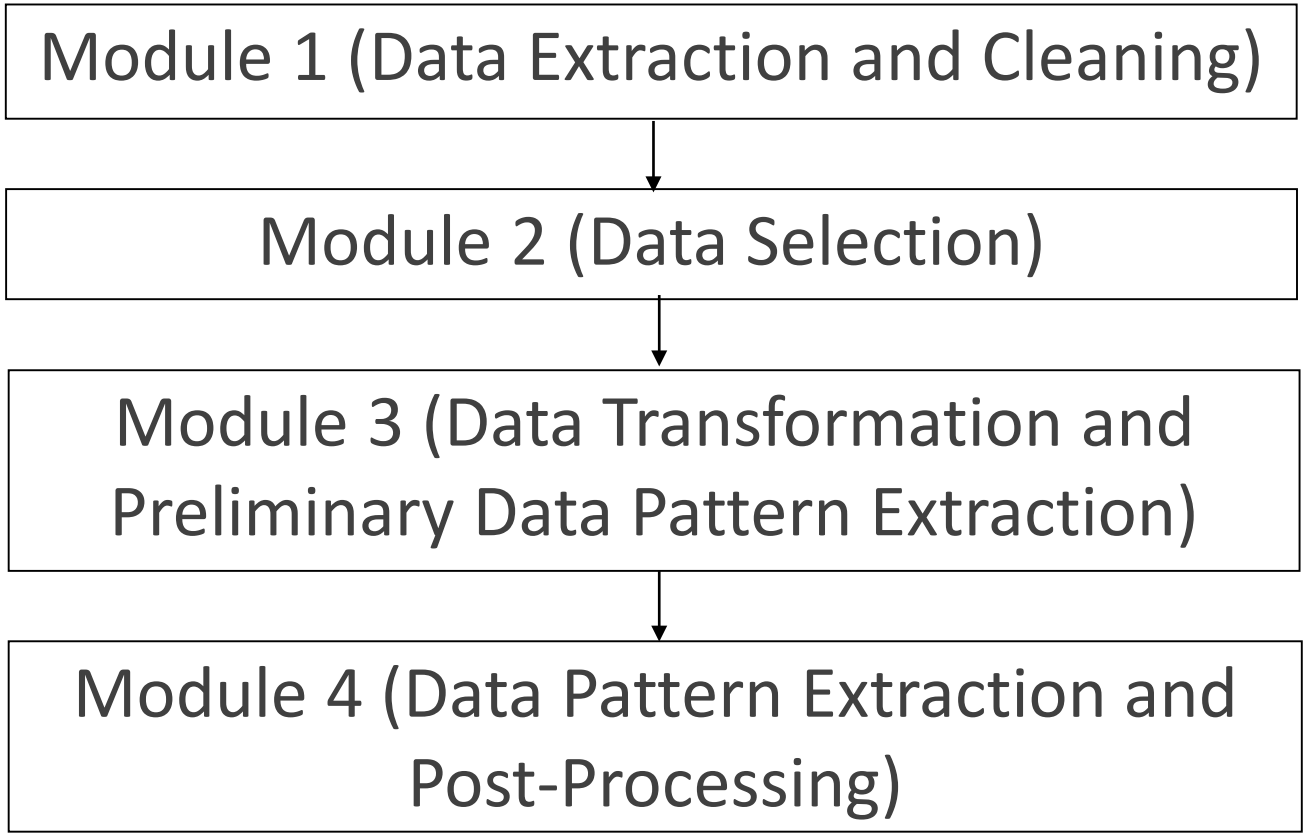


The geomorphic provinces in Tennessee are shown.

Study Area : 7 watersheds (mountainous vs non-mountainous provinces)



Date Mining Approach



Date Mining Approach

The USGS started collecting stream information in 1889 by using a stream gage.

A stream gage measures water level, but also collects information about the water quality and amounts of sediment.

Measurements are usually recorded every 15 minutes and uploaded to the USGS.

Date Mining Approach

Module 1 acquired preliminary data in the form of tab-separated plain text data files from the USGS site.

Performed a filtered search constrained by a selected state, and a date range of January 1, 1921, to December 31, 2018, at two-year intervals.

The data set was then “cleaned” by removing records with missing data and “filtered” by excluding stations with fewer than twelve records.

Module 2 & 3 computed hydraulic geometry (b, f, m) from stream flow and channel geometry measurements collected by the USGS from at least **5** successive **2-year** intervals.

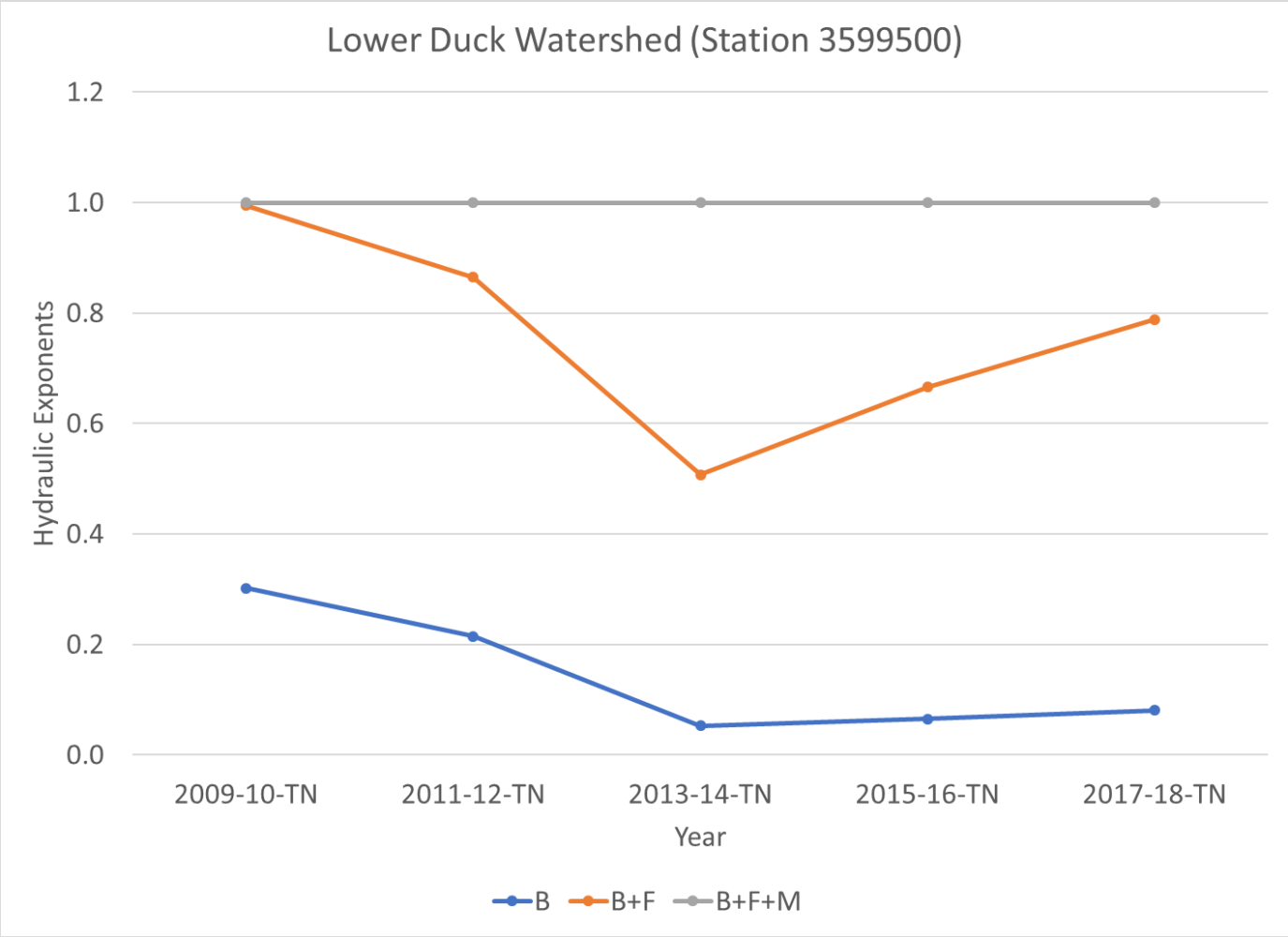
Results

In the watersheds selected from non-mountainous region, 50 stations met our criteria.

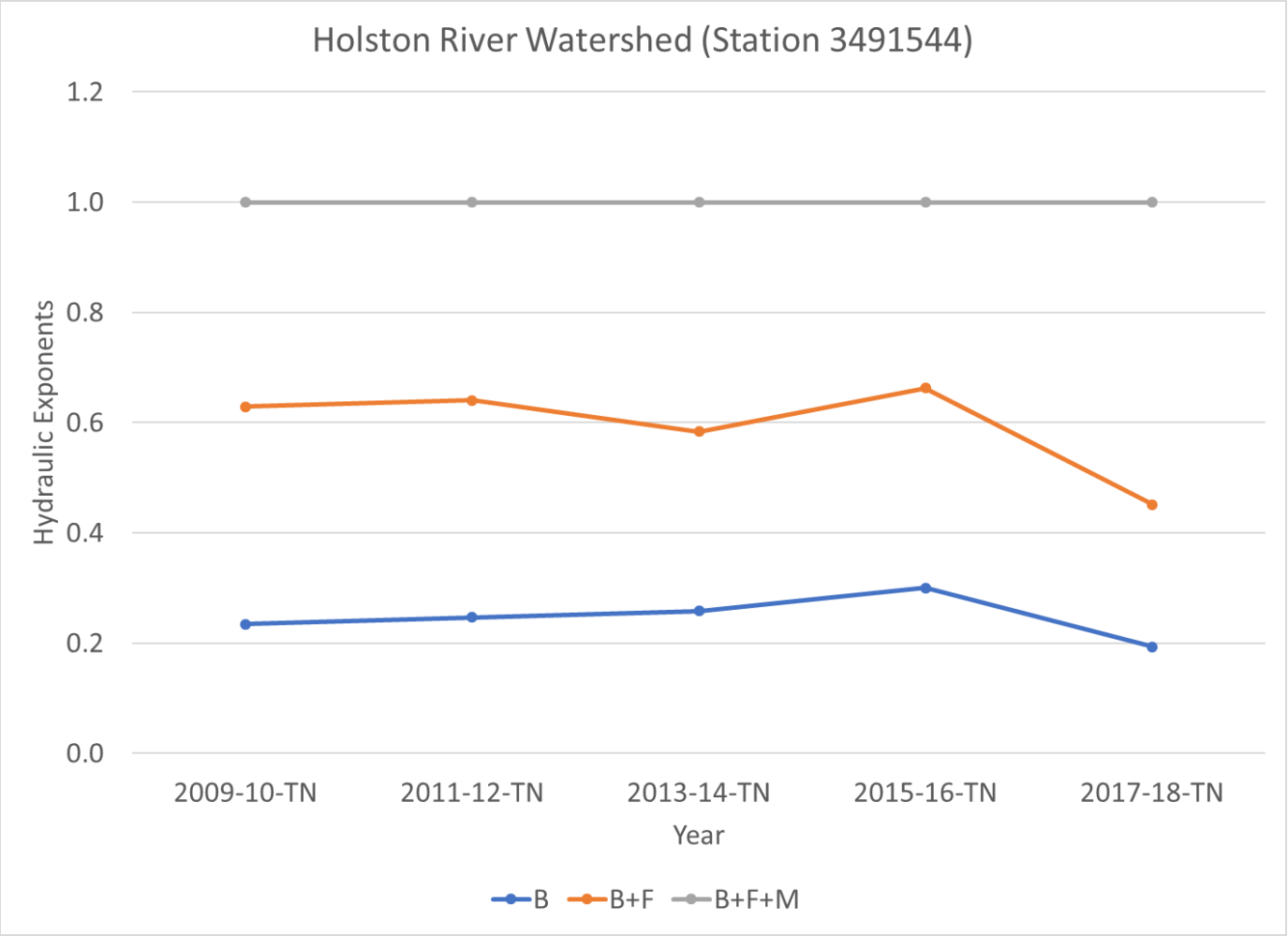
In the watersheds selected from the mountainous region, 16 stations met our criteria.

One station from each region was arbitrarily chosen to illustrate the variability in hydraulic geometry over time:

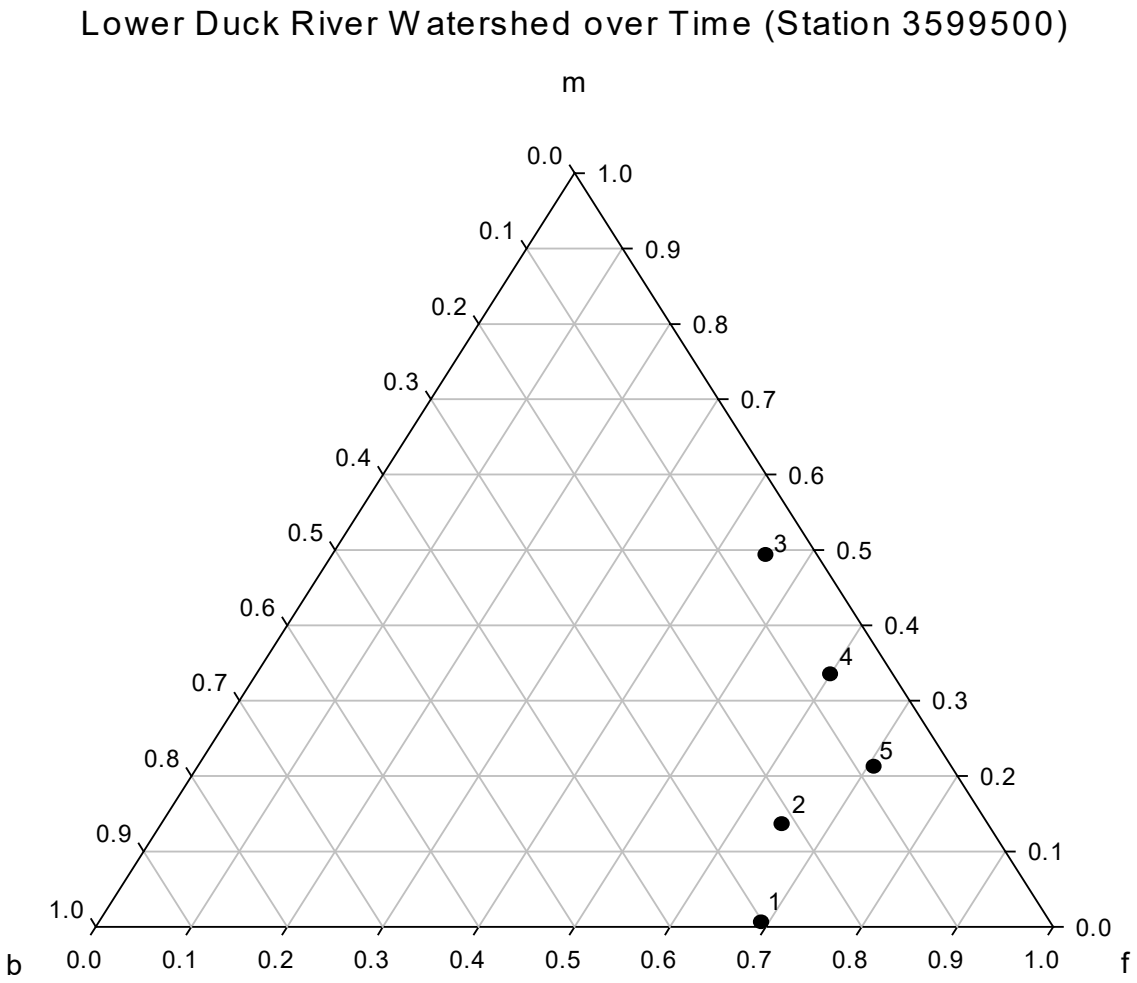
- station 3599500 from the Lower Duck Watershed (non-mountainous)
- station 3491544 from the Holston River Watershed (mountainous)



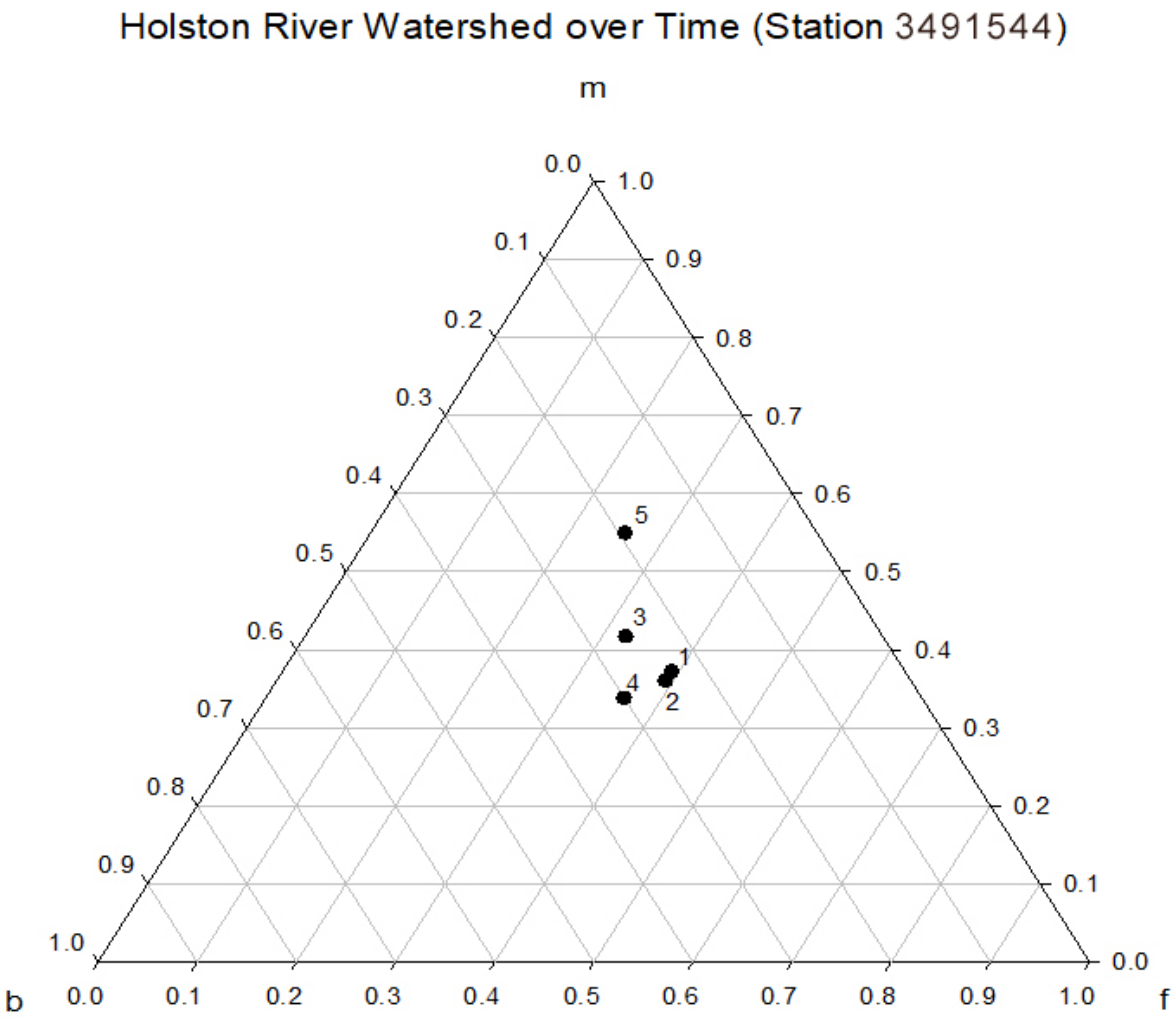
Variability of hydraulic geometry at a single station (**non-mountainous**) over 5 successive 2-year periods, beginning 2009-2010, ending 2017-2018



Variability of hydraulic geometry at a single station (**mountainous**) over 5 successive 2-year periods, beginning 2009-2010, ending 2017-2018



Variability of hydraulic geometry at a single station (**non-mountainous**) over 5 successive 2-year periods, on a ternary diagram.



Variability of hydraulic geometry at a single station (**mountainous**) over 5 successive 2-year periods, on a ternary diagram.

Results

Over 5 successive 2-year intervals from 2009 through 2018, the Rhodes channel type of the station in the Lower Duck River Watershed (non-mountainous) transformed in the sequence X → X → IV → VIII → X, while channel type of the station in the Holston River Watershed transformed in the sequence VI → VI → IV → VI → II.

One station from each region (Data of 50 stations from non-mountainous region, 16 stations from mountainous region)

Statistical Hypothesis Tests

Pooled variances were computed from variances in successive 2-year intervals for hydraulic exponents *b*, *f*, and *m* from selected mountainous and non-mountainous watersheds.

F-statistic was calculated as the ratio of the pooled variances from the two regions.

All three calculated F-statistics were less than the critical value of F, indicating that there was no significant difference between the pooled variances of hydraulic exponents in mountainous and non-mountainous watersheds at the 5% level of significance.

Statistical Hypothesis Tests

Next Slide: Pooled variances were computed from variances in successive 2-year intervals for hydraulic exponents *b*, *f*, and *m* from selected mountainous and non-mountainous watersheds.

F-statistic was calculated as the ratio of the pooled variances from the two regions.

All three calculated F-statistics were less than the critical value of F, indicating that there was **no significant difference** between the pooled variances of hydraulic exponents in mountainous and non-mountainous watersheds at the 5% level of significance.

Statistical Hypothesis Tests

	---- Pooled variance ----				
Hydraulic Exponent	**Non-Mountain**	**Mountain**	**Fcalc**	**Fcrit α=0.05**	**p-value**
b	0.01078	0.00783	1.38	2.18	0.25
f	0.01156	0.00869	1.33		0.28
m	0.02152	0.01289	1.67		0.14
number of stations	50	16			
degree of freedon	49	15			

Discussion

The high variability of at-a-station hydraulic geometry over a relatively short time span (a decade) that includes large jumps, e.g., from Interval 2 (2011-2012) to two Interval 3 (2013-2014) on the ternary diagram, suggests that investigators utilizing hydraulic geometry in their applications should work with stream flow and channel geometry data that is as **recent** as possible.

Conclusion

Statistical hypothesis tests indicated that there was **no significant difference** between pooled variances computed from variances in successive 2-year intervals for hydraulic exponents *b*, *f*, and *m* from selected mountainous and non-mountainous watersheds in Tennessee.

Thus, differences in regional **topographic setting** did not produce differences in **magnitude of the temporal stability** of at-a-station hydraulic geometry.

Conclusion

Data mining cannot replace application-specific field measurements conducted by investigators who require high density, survey-quality data for localities without USGS data archives.

Nevertheless, data mining can make an important contribution to understanding statistical relationships between environmental variables and hydraulic geometry, and can provide readily-available input for models based on hydraulic geometry.

Conclusion

Output from applications that utilize hydraulic geometry as input might be sensitive to short-term (annual) perturbations in hydraulic geometry.

Ridenour determined seasonal variability in the magnitude and location of maximum oxygen deficit and remaining oxygen demand by using hydraulic geometry, discharge, and water temperature at two gaging stations situated upstream and downstream.

Conclusion

Thompson et al. found that it is possible to create a strategically designed system that maximizes velocity and energy reductions, providing insight that is currently lacking regarding the design of regenerative stormwater conveyances (RSCs).

Rosenfeld et al. found that optimal flows based on hydraulic geometry were higher for larger fish and increased proportionally as streams became smaller, and they decreased downstream.

These applications would benefit from multi-interval hydraulic geometry data for sensitivity and time series analysis.

Conclusion

While recognizing its spatial limitations, data mining can also provide **models** that incorporate hydraulic geometry parameters with data for calibration.

Additionally, data mining can facilitate**validation** of theoretically derived hydraulic geometry relationships predicted from environmental variables by comparison with field data.

Finally, data mining provides an efficient method to **quantify** the **stability** of empirically determined relationships.

Thank You!

Questions?