GIS Approaches for Channel Typing in the Columbia River Basin: Carrying Fine Resolution Data to a Large Geographic Extent

Hiroo Imaki, Tim Beechie Northwest Fisheries Science Center, NOAA Fisheries, Seattle, Washington, USA
(hiroo.imaki@noaa.gov) (tim.beechie@noaa.gov)

Introduction

We are developing a series of GIS and statistical techniques to predict 8 channel types based on geomorphic principles (Figure 1) and improving prediction accuracy in the Columba River basin (Figure 2). Our stream layer can be used as a geomorphic reference condition on a 50 ~ 100k scale. Four major steps are included:

(1) Reconstruct a historic stream layer based on current stream GIS layers,
(2) Calculate reach attributes and construct channel prediction models,
(3) Examine prediction errors, and
(4) Develop GIS stream data with channel types and other attributes

Figure 1. Biophysical controls on channel pattern.

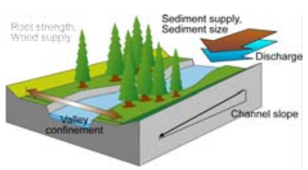
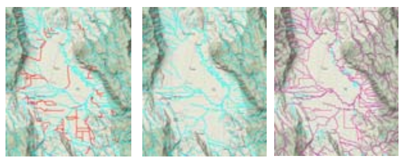


Figure 2. The Columbia River basin and the scale of channel typing. The total length of reaches in the GIS layer is 444,121,266 m and the number of reaches are 2,273,010. A blue area in the Walla Walla watershed (red area) includes 295 reaches, all of which are 200 m long.

Approach

Historic stream reconstruction

We built our historic stream network based on current stream GIS layers. US and Canadian stream layers were merged and artificial channel types were removed from the network. To reconnect isolated natural channels, we restored the minimum number of artificial channels back to the network (Figure 3).



1. Natural (cyan) and artificial (red) channels.
2. Artificial (red) channels were removed.
3. Isolated channels were reconnected

Figure 3. Historic channel reconstruction processes. Artificial channels (red lines) were removed, and isolated streams were reconnected with a minimum amount of artificial channels.

Reach Attributes

We started with a minimum reach length of 200 m and attached six attributes to each reach: 1) slope, 2) accumulated precipitation (surrogate of discharge), 3) relative shear stress to the above reach, 4) confinement, 5) % of fine sediment area in its drainage, and 6) % of alpine sediment supply area in its drainage (Figure 4). All 6 attributes were calculated based on a 10 ~ 30 m DEM (US and Canada). Where adjacent 200 m segments had similar slope and bankfull width, we aggregated multiple segments into longer geomorphically meaningful reaches. This also increased accuracy in slope calculations in low gradient floodplains.

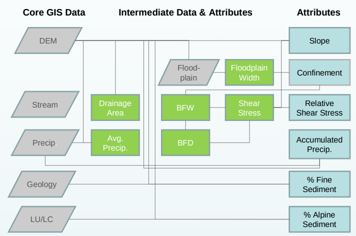


Figure 3. GIS processes to calculate reach attributes.

Channel Typing and Model Selection

Our channel type classification is based on existing studies (Montgomery and Buffington 1997, Beechie *et al*. 2006, Hall *et al*. 2007) and Linear Discriminant Analysis (LDA, Figure 5). Reaches were first divided into either confined or floodplain channel types based on bankfull width and confinement. The four confined channel types were classified based on slope. The four floodplain channel types were classified based on LDA of reach attributes. We tested all combinations of six attributes and selected the model with the highest overall accuracy. We also tested three other statistical techniques: prior probabilities, random forest, and bagging (bootstrapping).

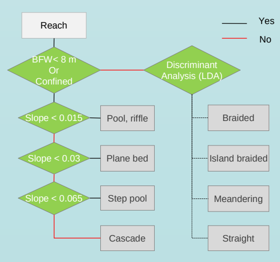


Figure 4. Channel typing schema. Reaches less than 8 m bankfull width were classified into 4 confined channel types based on Montgomery & Buffington (1977). Reaches with bankfull width > 8 m and unconfined were classified as migrating (Hall *et al*. 2007) and subjected to further classification with LDA.

Prediction accuracy and errors

We constructed a series of classification error matrices to examine prediction accuracy and errors. Then we further examined an error structure in our prediction by conducting an individual perturbation analysis.

	Slope	Accumulated Precipitation	Relative Shear Stress	Confine-ment	Fine Sediment
Mean accuracy (n=100)	65.8	76.2	72.3	74.2	75.2
SD	3.46	1.89	2.51	1.67	1.62

Table 3. Individual perturbation analysis on channel type classification errors. Errors were generated from known error distributions based on field measurements and their estimations. For the fine sediment, we used standard deviation of 0.25. An overall accuracy of the model before adding errors was 76.0 % for this analysis.

Results

We found that LDA with bagging showed the highest overall accuracy (80.8%, Table 1). However, the accuracy with a test dataset decreased to 60.3%. A model including all 6 parameters resulted in the highest accuracy among LDA (76.8%). We increased model accuracy about 20% compared to the traditional slope – discharge model (Model 8). The prior probability of existing channel types increased overall accuracy up to 8%. Random forest resulted in relatively low overall accuracy.

We found that major classification errors occurred in braided and straight channel types (Table 2). Model 1 often confused these two types with the island-braided type (omission error: 60.0 and 65.2, respectively). The most influential source on the classification error was slope (Table 3; 10.2% decrease) while other parameters showed only slight decrease in accuracy.

	Slope	Accum-ulated Precip.	Relative Shear Stress	Confine-ment	Fine Sediment	Alpine Sediment	Overall Accuracy (w/o prior probability)	Overall Accuracy: Test Dataset
Model 1	■	■	■	■	■	■	76.8 (75.0)	63.3
Model 2	■	■	■		■	■	75.9 (67.9)	66.7
Model 3	■	■	■	■	■		75.0 (71.4)	60.0
Model 4	■	■	■	■		■	74.1 (68.8)	56.7
Model 5	■	■		■	■	■	74.1 (72.3)	55.0
Model 6	■		■	■	■	■	74.1 (67.0)	51.7
Model 7	■	■	■				68.8 (61.6)	63.3
Model 8	■	■					61.6 (60.7)	60.0
Bagging (LDA)							80.8	60.3
Random Forest								60.0

Table 1. Classification models and their overall accuracies. Both with and without prior probability (frequency of existing channel types in the basin) models were tested with a training dataset. We also tested each model's accuracy with a test dataset.

Prediction	Training Data (Known Channel Types): Braided	Island-braided	Meandering	Straight	Total	Commission error
Braided	6	1	0	1	8	75.0
Island-braided	3	39	6	4	52	75.0
Meandering	0	3	26	3	32	81.3
Straight	1	3	1	15	20	75.0
Total	10	46	33	23	112	
Omission error	60.0	84.8	78.8	65.2		**76.8**

Table 2. Channel type classification error matrix of Model 1. We selected a training dataset to reflect a distribution of channel types in the basin.

Conclusion

We predicted 8 channel types using multiple approaches and found that the bagging with LDA is the most promising approach to estimate channel types at the reach scale across an entire basin. Bagging resulted in increased overall accuracy and calculated a voting distribution which can then be used as an indicator of prediction certainty. Even with medium resolution DEM and stream layers, bagging allowed us to predict channel types with relatively high accuracy across the Columbia River basin in spite of its diverse geological environment. We also found that relatively lower accuracies with the test dataset were in large part due to observer's channel typing error. Upon cross examination of channel type calls with two independent observers, we roughly estimated that 10 ~ 15% of error can come from observer differences. The applicability of our prediction method is large and can be applied in many regions because it uses pre-existing GIS data.

Citations:

Beechie, T.J., M. Liermann, M.M. Pollock, S. Baker, and J. Davies. 2006. Channel pattern and river-floodplain dynamics in forested mountain river systems. Geomorphology. 78(1-2): 124-141.
Hall, J.E., D.M. Holzer, T.J. Beechie. 2007. Predicting river floodplain and lateral channel migration for salmon habitat conservation. Journal of the American Water Resources Association. 43(3): 1-12.
Montgomery, D.R. and J.M. Buffington.1997. Channel-reach morphology in mountain drainage basins. GSA Bulletin. 109(5): 596-611.

GIS data sources:

	Canada	USA
Stream layer	The Watershed Atlas http://www.env.gov.bc.ca/fish/watershed_atlas_maps/	National Hydrography Dataset Plus (NHDplus) http://www.horizon-systems.com/nhdplus/
DEM	The Canadian Digital Elevation Data (CDED) http://www.geobase.ca/geobase	National Elevation Dataset (NED) http://ned.usgs.gov/
Precipitation	ClimateBC http://www.genetics.forestry.ubc.ca/cfcg/climate-models.html	PRISM http://www.prism.oregonstate.edu/
Geology	Digital Geology Map of British Columbia http://www.em.gov.bc.ca/Mining/Geolsurv/Publications/catalog/bcgeolmap.htm	Geological Survey from each state
Land use / land cover		National Land Cover Data (NLCD) http://www.epa.gov/mrlc/