Markov Models for Coho Variable Selection

Heather Sweeney

EISI REU Summer 2014 Final Presentation
August 21, 2014

Markov Models

- Current state depends on only last n states
- 1st order Markov: P($S_t \mid S_{t-1}$)
- 2nd order Markov: P($S_t \mid S_{t-1}, S_{t-2}$)

Purpose

- Develop and test Markov models to predict Coho selection of hydraulic variables over time
  - NOT directly based on location
- Patterns to understand system
- Insight to develop better models

Questions

- Do fish have a tendency to remain in the same variable range?
- How accurate are first and second order Markov models?
- Is a first order Markov process better than a second order Markov process
  - Does S_t depend on S_{t-2}?

Method

- Analysis based on 5 variables:
  - TKE, Strain, Vxyavg, Depth, DistWood
- For each variable, analysis conducted for 2-4 states
- Data split into training and test sets
  - Model trained on training set
  - Tests conducted using test set

Method

Question	Test
Tendency to change environment	Transition Probabilities
Accuracy and Reliability of Models	Confusion Matrix
1st vs. 2nd order Markov	Transition Probabilities & Confusion Matrix

Transition Probability Matrix

- 1st order: tends to remain in current state
- 2nd order: tends to remain in current state, appears to vary depending on t-2

Current

Previous	Low	Medium	High
Low	0.73	0.21	0.06
Medium	0.22	0.55	0.23
High	0.03	0.15	0.82

1st Order

Current

Previous t-2, t-1	Low	Medium	High
Low, Low	0.83	0.14	0.03
Medium, Low	0.53	0.33	0.13
High, Low	0.22	0.67	0.11
Low, Medium	0.44	0.37	0.19
Medium, Medium	0.17	0.66	0.16
High, Medium	0.08	0.49	0.43
Low, High	0	0.33	0.67
Medium, High	0.04	0.29	0.67
High, High	0.03	0.12	0.85

2nd Order

- *Example-DistWood, 3 compartments*

Confusion Matrix

- Reliability = fraction of times values predicted to be in section A are actually in section A
- Accuracy = fraction of times values in section A are predicted to be in section A

Average Measures per Variable							
		# of States	Depth	DistWood	TKE	VxyAvg	Strain
Reliability	1st order	2	0.84	0.71	0.60	0.62	0.55
		3	0.59	0.55	0.50	0.48	0.41
		4	0.51	0.50	0.38	0.39	0.33
	2nd order	2	0.88	0.73	0.67	0.64	0.58
		3	0.64	0.59	0.55	0.52	0.45
		4	0.56	0.52	0.40	0.44	0.36
Accuracy	1st order	2	0.84	0.71	0.60	0.62	0.55
		3	0.57	0.57	0.51	0.48	0.41
		4	0.51	0.50	0.40	0.40	0.34
	2nd order	2	0.87	0.73	0.67	0.64	0.58
		3	0.63	0.61	0.55	0.52	0.44
		4	0.55	0.52	0.42	0.44	0.36

Discussion

Question	
Tendency to change environment	Tend to remain in same variable range, moreso for 1st than 2nd order model
Accuracy and Reliability of Models	Better than uniform random distribution, but improvement possible
1st vs. 2nd order Markov	2nd order better

- 1st: tendency to remain in same variable range
- 2nd: if moving, have a greater tendency to continue moving

Moving Forward

- More rigorous statistical tests
- If a fish is moving, which variable ranges is it more likely to move to?
- Incorporate location into the model

Acknowledgments

- Desiree Tullos (Principal Investigator)
- Jorge Ramirez (Mathematics Advisor)
- Rebecca Hutchinson (Postdoc)
- Cara Walter (Graduate Student)
- Tom Dietterich (Professor)
- Julia Jones (Professor)
- Alan Stanton (Project Mentor)
- Isabela Brown, Skyler Doak, Shannon Rummage (REU students)

Questions?

Appendix

Confusion Matrix

- Predicts better than random uniform distribution
- No accuracy or reliability above 90%
- 2nd order performs better than first order
- *Example–DistWood, 3 compartments*

Predicted

Actual	1	2	3	Accuracy
1	280	104	49	0.646651
2	134	171	112	0.410072
3	41	51	166	0.643411
Reliability	0.615385	0.52454	0.507645	

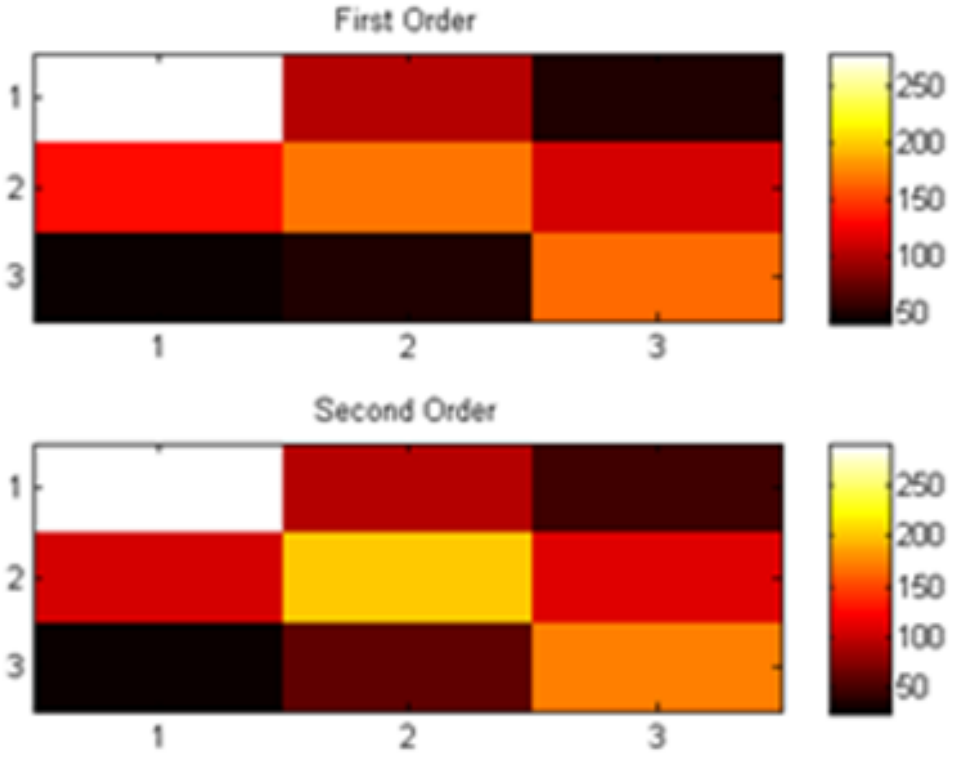
1st Order

Predicted

Actual	1	2	3	Accuracy
1	291	94	48	0.672055
2	105	201	111	0.482014
3	25	62	171	0.662791
Reliability	0.691211	0.563025	0.518182	

2nd Order

Confusion Matrix



- *Example-DistWood, 3 compartments*

	Low	Medium	High
Low	0.732719	0.207373	0.059908
Medium	0.222222	0.550926	0.226852
High	0.029412	0.152406	0.818182

	Low	Medium	High
1, 1	0.828025	0.140127	0.031847
2, 1	0.533333	0.333333	0.133333
3, 1	0.222222	0.666667	0.111111
1, 2	0.44186	0.372093	0.186047
2, 2	0.172727	0.663636	0.163636
3, 2	0.081633	0.489796	0.428571
1, 3	0	0.333333	0.666667
2, 3	0.041667	0.291667	0.666667
3, 3	0.027211	0.122449	0.85034