BAYESIAN NETWORK PREDICTIONS FOR PRESENCE / ABSENCE OF 20 mm SMELT IN WELL-DEFINED BAY DELTA ZONES

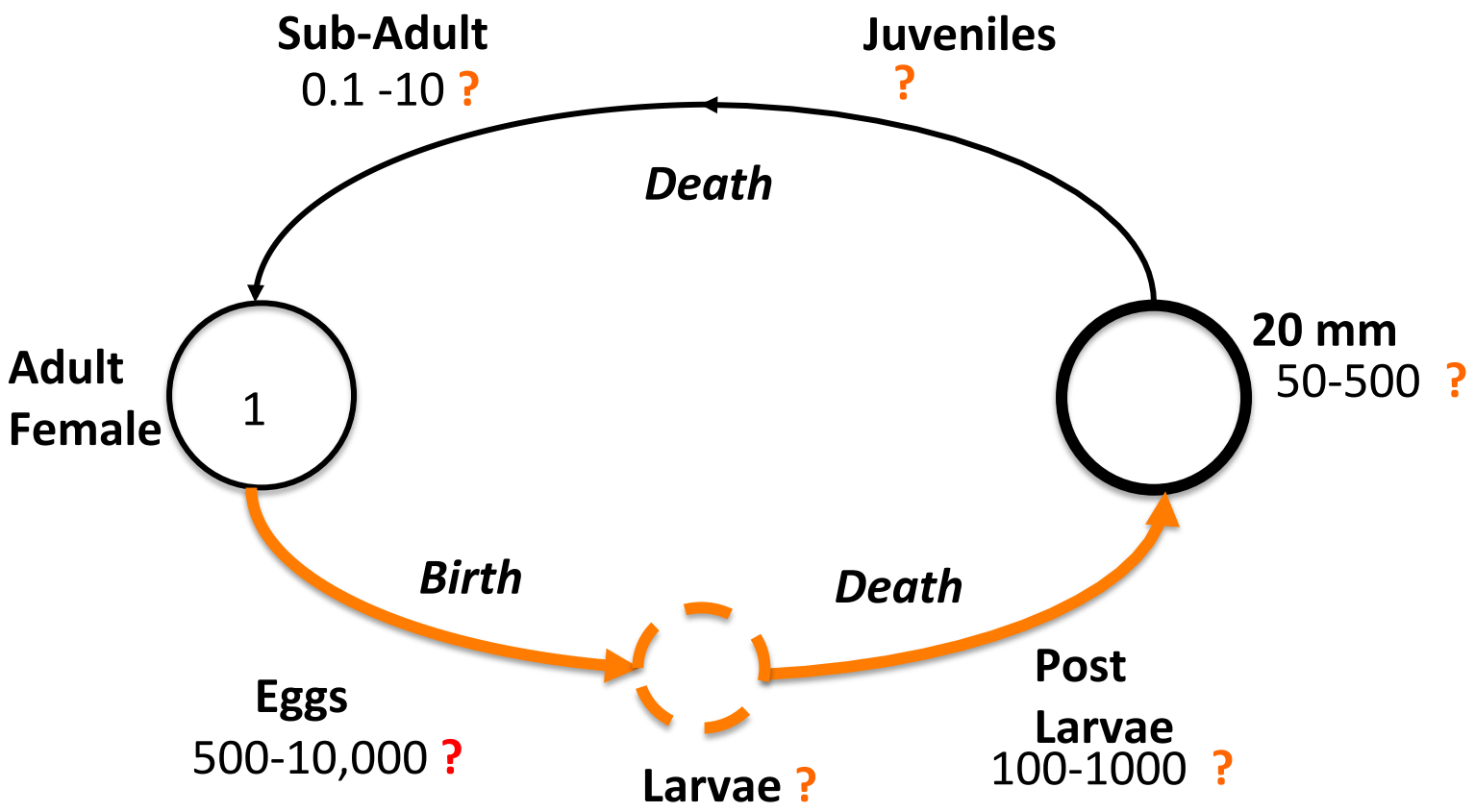
R M Oliver

B J Miller

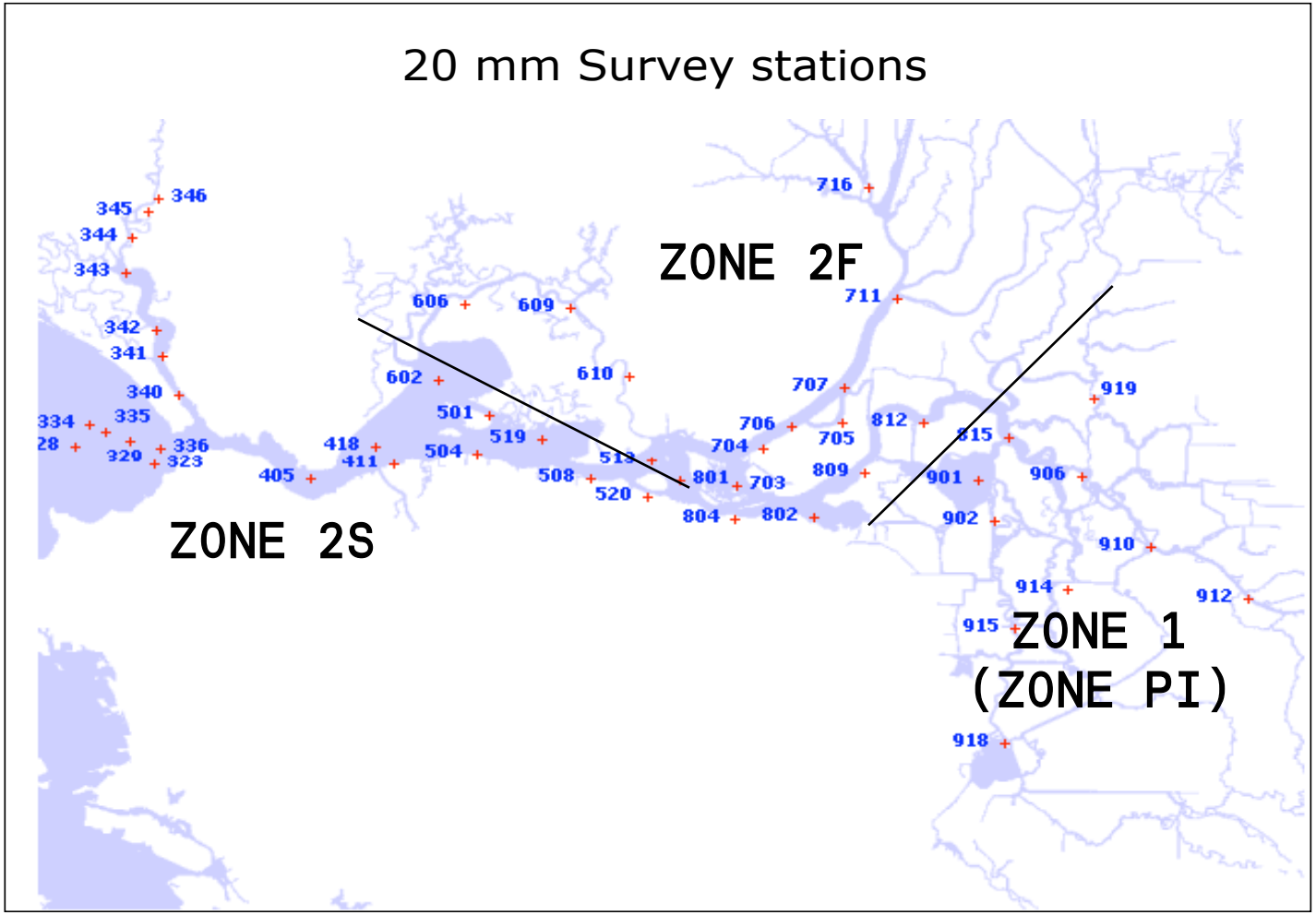
November 16, 2016

A Simplified Annual Smelt Life-Cycle Illustrating Magnitudes of Survival Uncertainty

(not Statistical Estimation Fitting Errors)

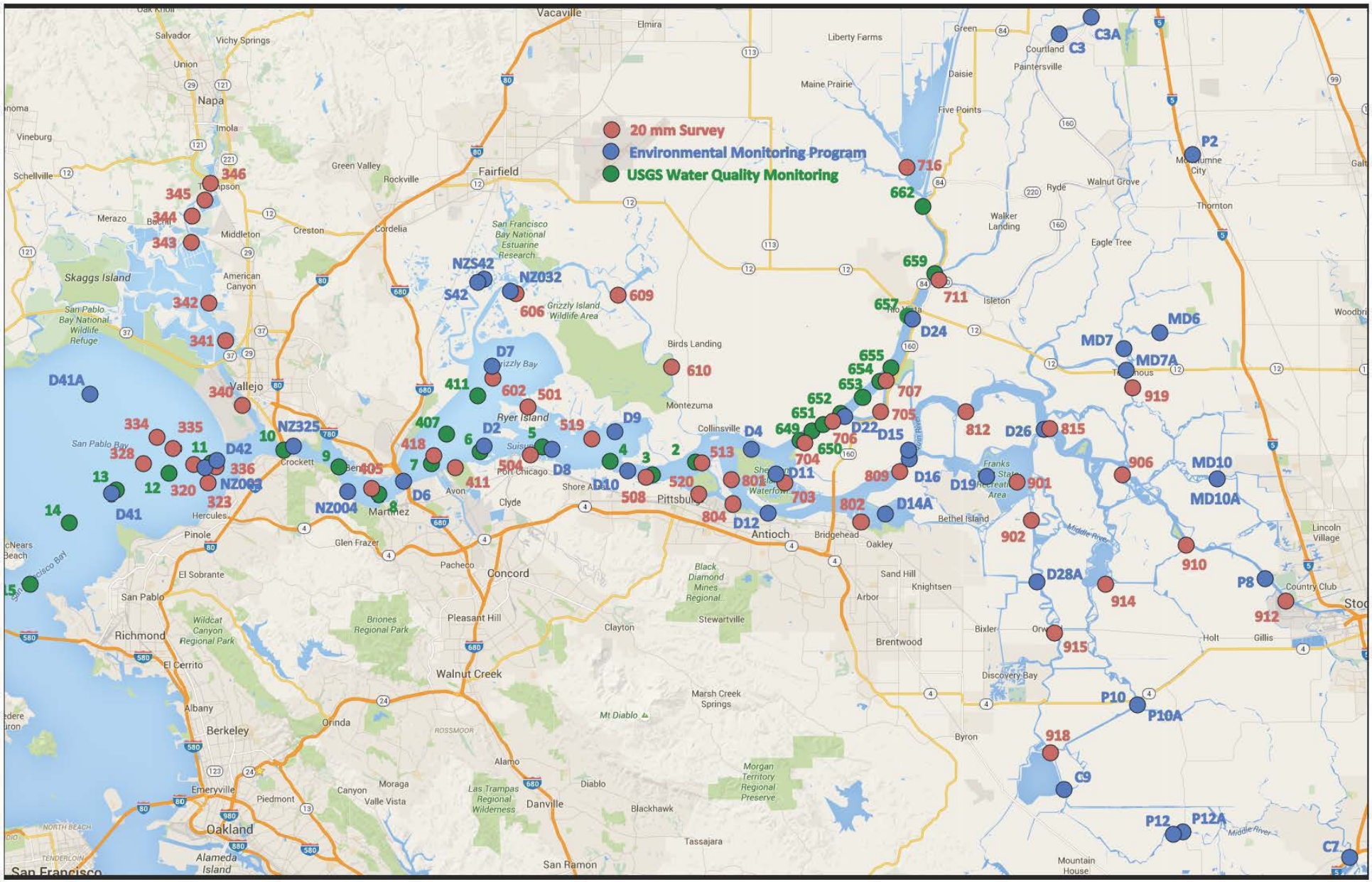


Original Definition of Zones from 20 mm Smelt Survey Data



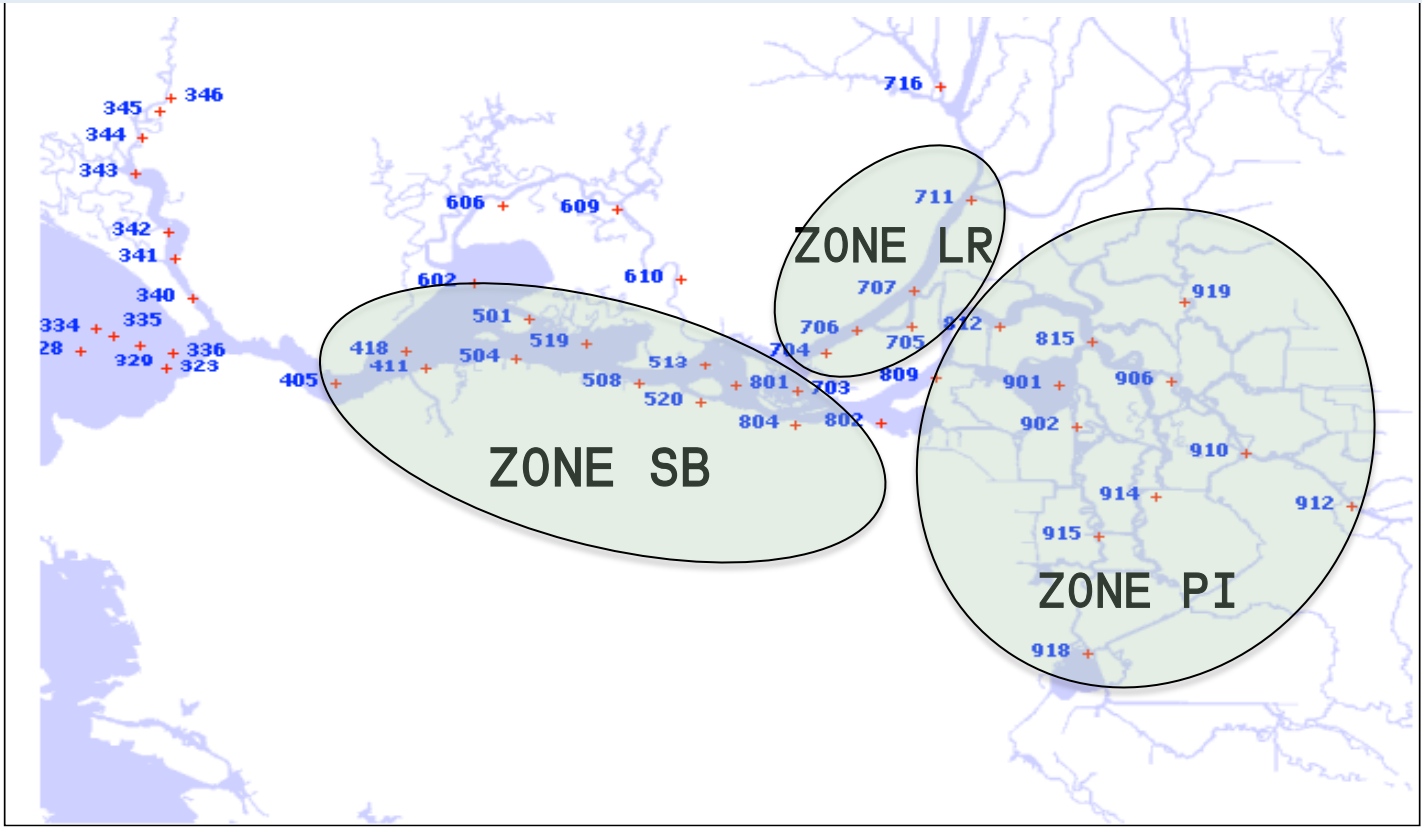
ZONE1: 901 902 906 910 912 914 915 918 919
ZONE2F: 508 513 520 602 606 609 610 703 704 705 706 707 711 716 801 802 804 809 812 815
ZONE2S: 320 323 328 334 335 336 340 341 342 343 344 345 346 405 411 418 501 504 519

MULTIPLE SURVEY AND MONITORING STATIONS



Location and Definition of Zones for 20 mm Smelt

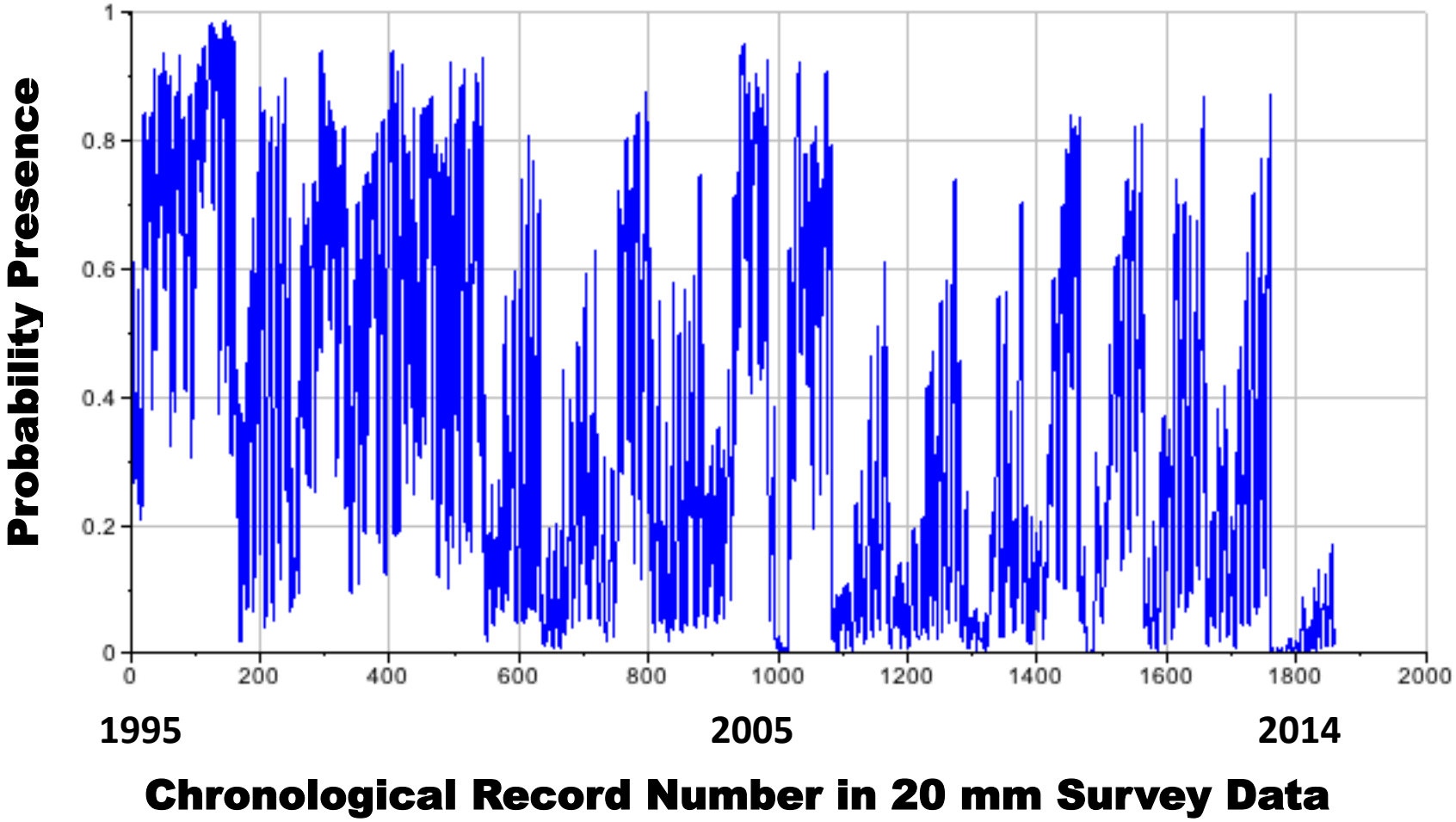
Zone Definitions not limited to Geographical Features; may also include salinity, stations in sheltered waters, relevant habitat for spawning, nutrient rich zones




```
ZONEPI:  809 901 902 910 912 914 915 918 812 815 906
ZONELR:  705 707 711 703 704 706
ZONESB:  513 804 405 418 520 602 411 501 504 508 519
```

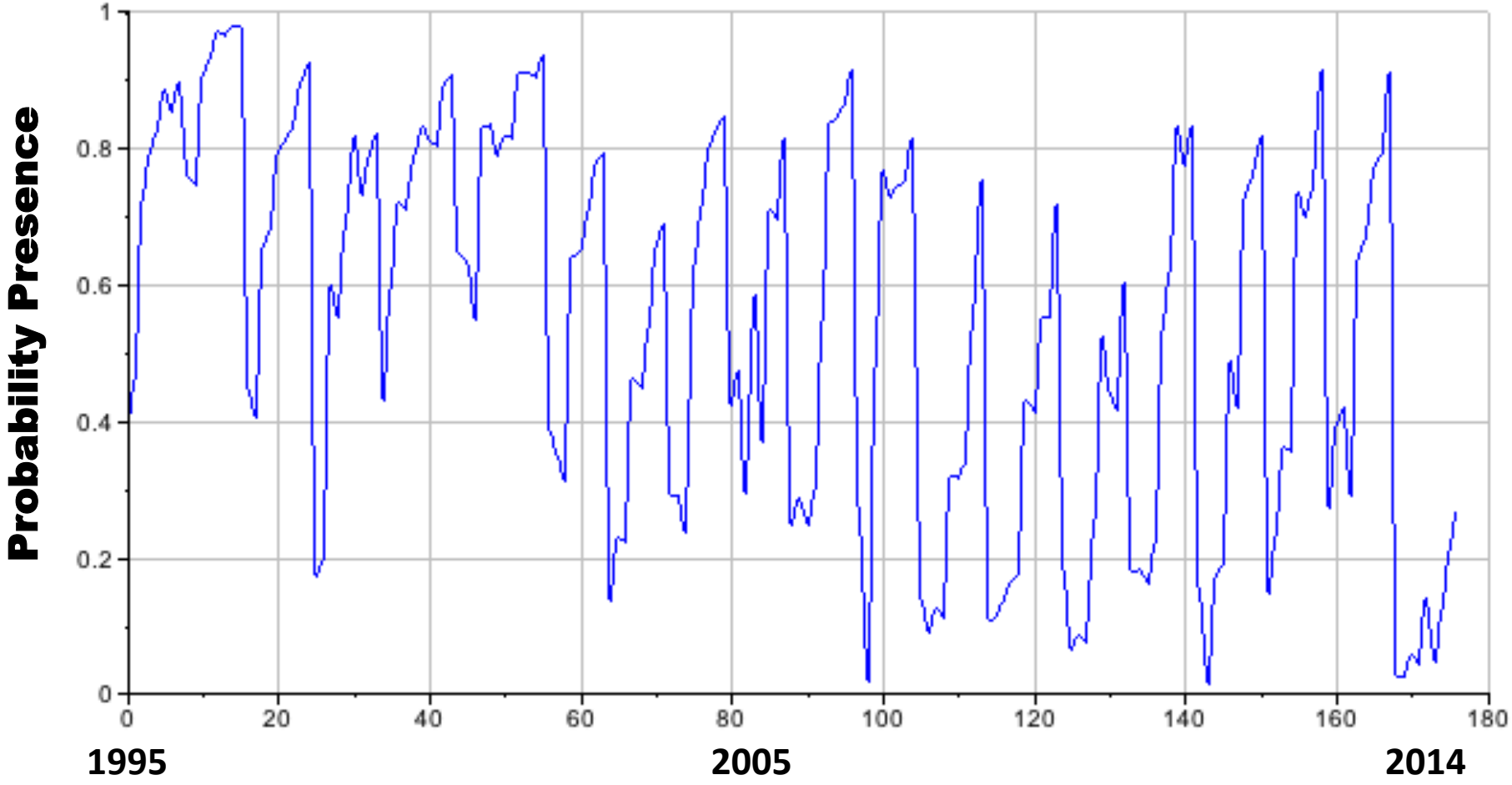
10/17/16 4:12 PM: RMO

20 mm Smelt Predictions in Zone SB from Relevant Flow, Food, Nutrient, Secchi … Data



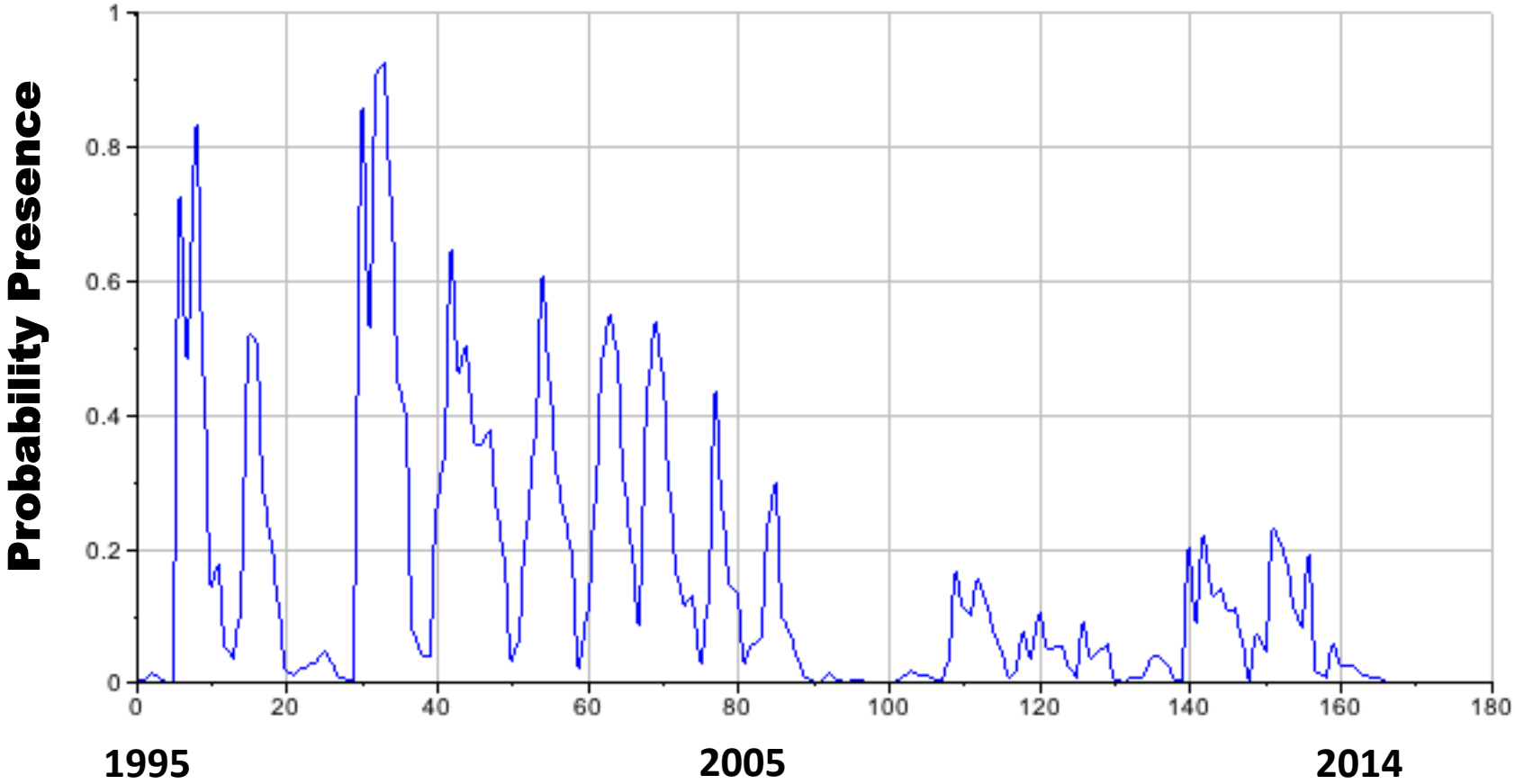

```
RR← (¯14 ¯1 ¯6 ¯30 11 13 16 17 21 24 27 31 32 33 ) (¯14 ¯1 ¯6 ¯30 11 21 24 27) LOGREGFACCHART 3 0↓SMELTSB927
```

Time-dependent Probability of 20 mm Smelt Presence, Station 804 in SB Zone (1995-2014)




```
RR← (¯12 ¯1 ¯5 ¯30 9 10 11 13 14 24 27 31 32 33)  LOGREGCOMPDEV SMELTSB927
  ⎕CHART (804 SMELT20mmSTN 16⊃RR)[;36]
```

Time-dependent Probability of 20 mm Smelt Presence, Station 918 in PI Zone (1995-2014)




```
RR←(¯12 ¯1 ¯5 ¯30 9 11 13 14 21 24 27 31 32 33) (¯12 ¯1 ¯5 ¯30 9 11 21 24 27) LOGREGFACCHART 1 0↓SMELTPI927
 ⎕CHART (918 SMELT20mmSTN 16⊃3⊃RR)[;36]
```

BAYES FACTORS

$$\Pr\{\textbf{Presence of Smelt} \mid \textbf{relevant Data}\} \triangleq p(1 \mid \mathbf{x}, \mathbf{y})$$

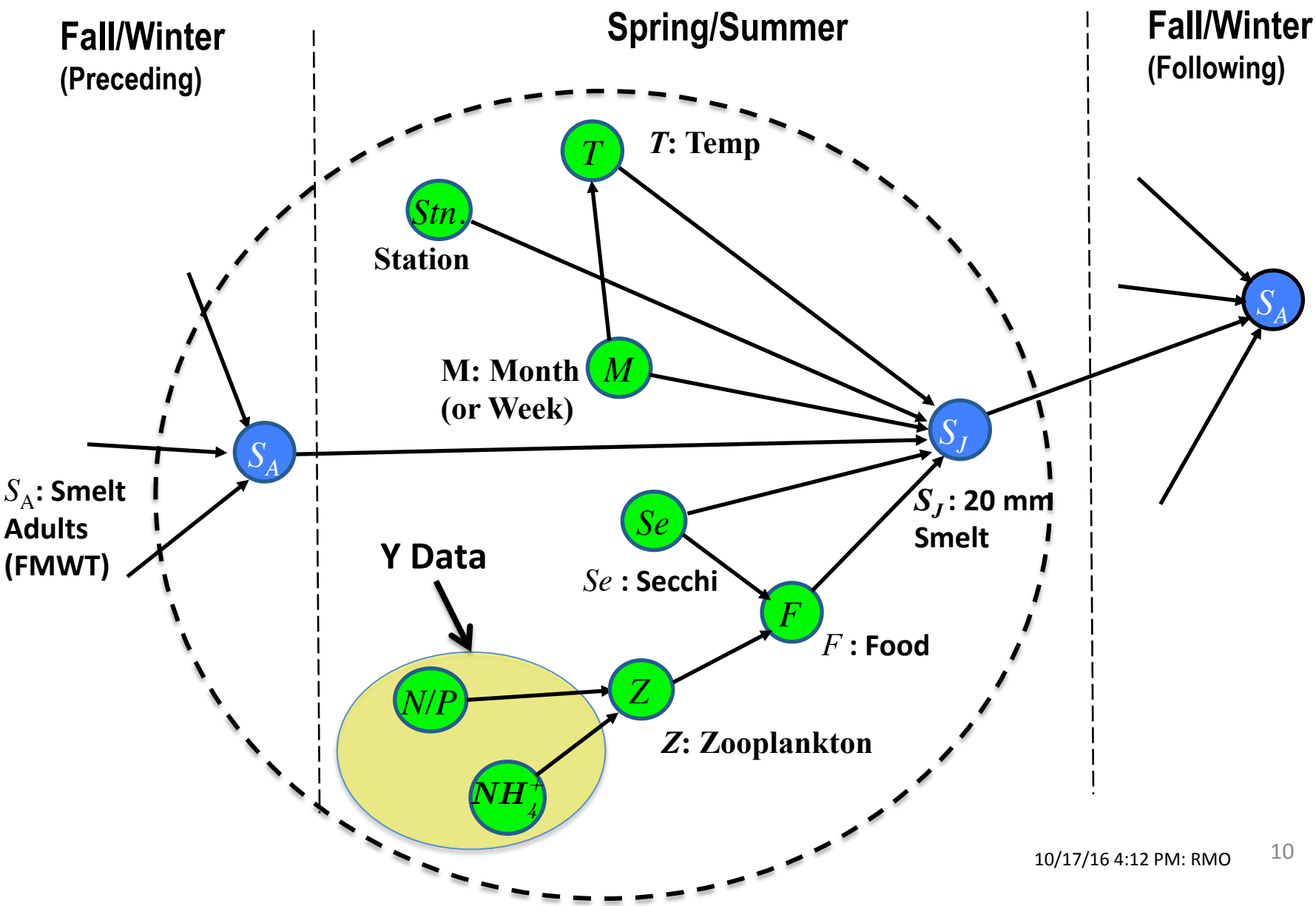
Bayes' Factor Contribution from new Clique (y) of Variables :

$$\frac{\Pr\{\textbf{Presence} \mid \textbf{ Data } \mathbf{x}, \mathbf{y})}{\Pr\{\textbf{Presence} \mid \textbf{ Data } \mathbf{x})} = \frac{p(1 \mid \mathbf{x}, \mathbf{y})}{p(1 \mid \mathbf{x})} \quad \textbf{stations within a Zone}$$

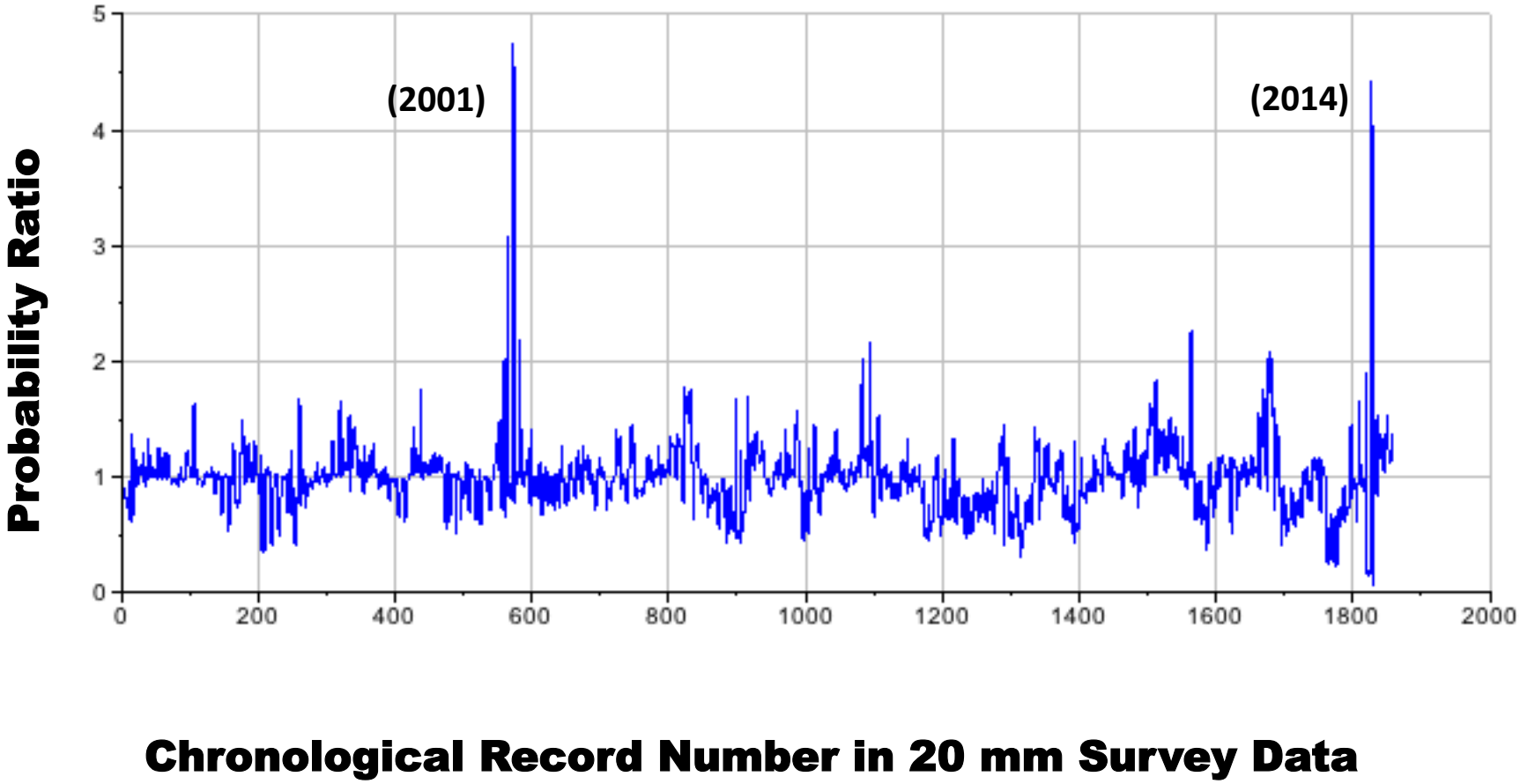
Examples : Conditional Independence in non - PI Zone to export policies in PI Zone or Influence of Nutrients within a Zone :

$$\frac{p(1 \mid \mathbf{x}, \mathbf{y})}{p(1 \mid \mathbf{x})} = \frac{p(1 \mid \mathbf{x}, \textbf{OMR Flow})}{p(1 \mid \mathbf{x})} \quad \textbf{or} \quad \ln \frac{Odds(1 \mid \mathbf{x}, \textbf{Nutrients})}{Odds(1 \mid \mathbf{x})}$$

Life-cycle Model Linking 20 mm Smelt Juveniles in Spring to Winter Adults (non-PI zone)

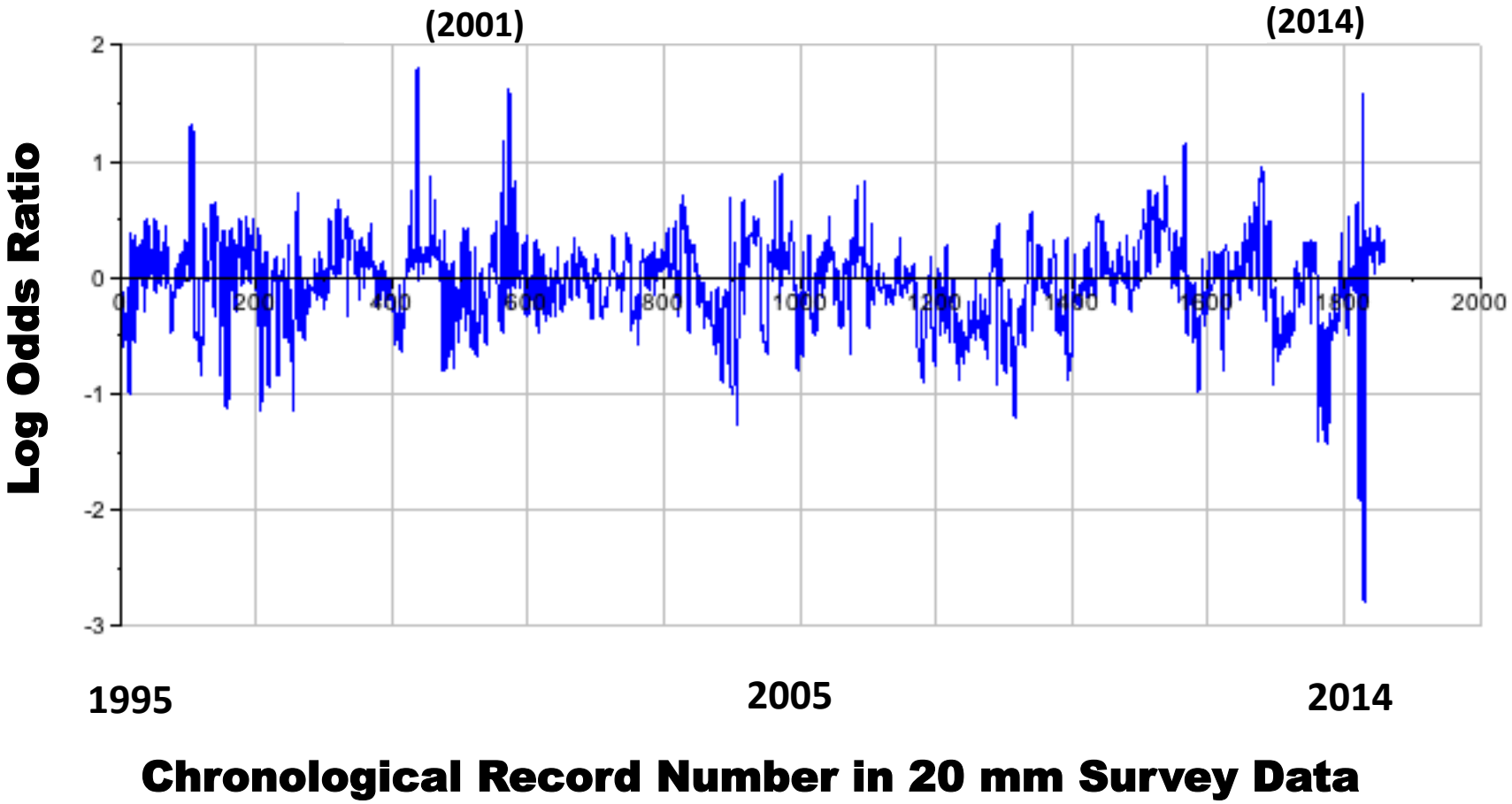


CONTRIBUTION OF NUTRIENTS TO PROBABILITY RATIO OF SMELT PRESENCE IN ZONE SB



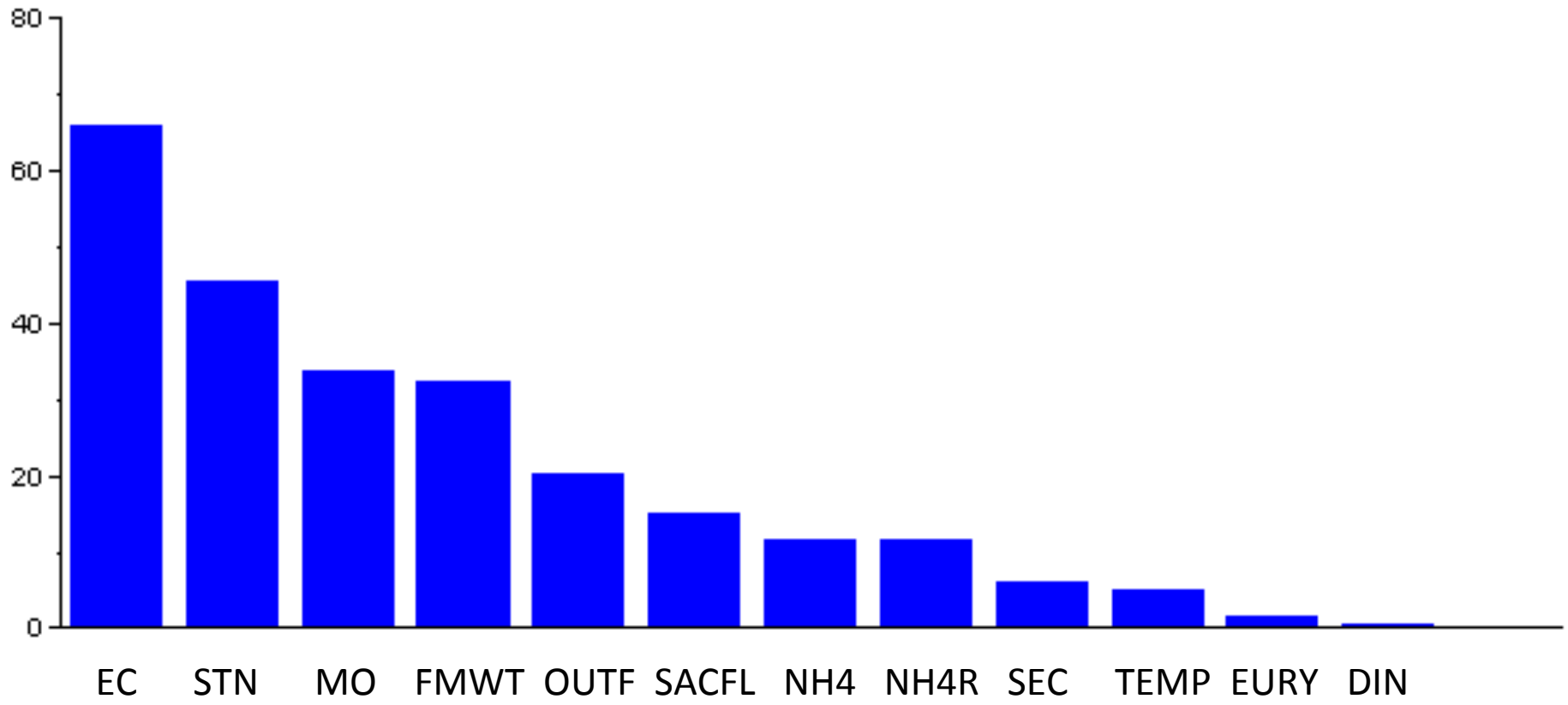

```
RR← (¯12 ¯1 ¯5 ¯30 9 11 13 14 21 24 27 31 32 33) (¯12 ¯1 ¯5 ¯30 9 11 13 14 21 24 27) LOGREGFACCHART 1 0 ↓ SMELTSB927
```

CONTRIBUTION OF NUTRIENTS TO LOG ODDS OF SMELT PRESENCE in ZONE SB




```
RR← (¯12 ¯1 ¯5 ¯30 9 11 13 14 21 24 27 31 32 33) (¯12 ¯1 ¯5 ¯30 9 11 13 14 21 24 27) LOGREGFACCHART 1 0 ↓ SMELTSB927
```

RANKED DIVERGENCE OF INDIVIDUAL FACTORS IN ZONE SB (1995-2014)




```
RR← (¯12 ¯1 ¯5 30 9 10 11 13 14 21 24 27 31 32 33)  LOGREGCOMPDEV SMELTSB927
 STATIONS: 513 804 405 418 520 602 411 501 504 508 519   ln odds proportional to previous FMWT
```

Summary Thoughts and Recommendations

- Well-defined Zones are critical to predictions of Smelt. To understand mortality, look for Zones where Smelt thrive.

- Candidates for “Good” prediction/decision models should include hierarchical food-web structure, expert knowledge and testable influences. Bayesian Belief Networks and Influence Diagrams help display important conditional independencies.

- A prediction model can help us identify signals and cliques of decision and random variables that reveal “probable causes” for fecundity, mortality and survival of Smelt.

- Bayes Factors extract significant influences and patterns. Can yield information on clusters of independent variables

- Be skeptical of global, static “Fish Stew” regression models

Bayesian Networks, Belief Nets, Influence Diagrams, Decision Analysis

(Selected References, Books, You-Tubes)

Predicting Location and Retrieval of Air France Flight 447 Airbus Recorder (6/1/2009); Article by Lawrence Stone in OR/MS Today August 2011

Neuse Estuary Water Quality Study, Duke University and NOAA Great Lakes Environmental Research Laboratory, 2004

Howard, R. A., (1988) Decision Analysis, Practice and Promise" Mgmt. Science 34(6) p.679-675

Lauritzen, S. L. and. D. J. Spiegelhalter (1988), "Local Computations with Probabilities on Graphical Model Structures and their Application in Expert Systems" JRSS Series B, v.50

Max Henrion "Rigs to Reefs: Decommissioning California's Offshore Oil Platforms", OR/MS Today, February 2015, https://www.youtube.com/watch?v=S7ws9V5sH5o

Oliver, R. M. and J. Q. Smith (1989), "Influence Diagrams, Belief Nets and Decision Analysis", Proceedings of the Berkeley May 1988 Conference, John Wiley, New York

Pearl, Judea (2000) "Causality: Models, Reasoning and Inference" Cambridge University Press

Koller, D. and N. Friedman (2009), "Probabilistic Graphical Models" The MIT Press

Marshall, K. T. and R. M. Oliver, (1995) "Decision Making and Forecasting", McGraw Hill, New York (See Chapter 4 on Model Building)

Software/Consulting: Prevision (Proprietary), OpenBUGS, Netica (Canadian), BayesiaLab, Lumina, …