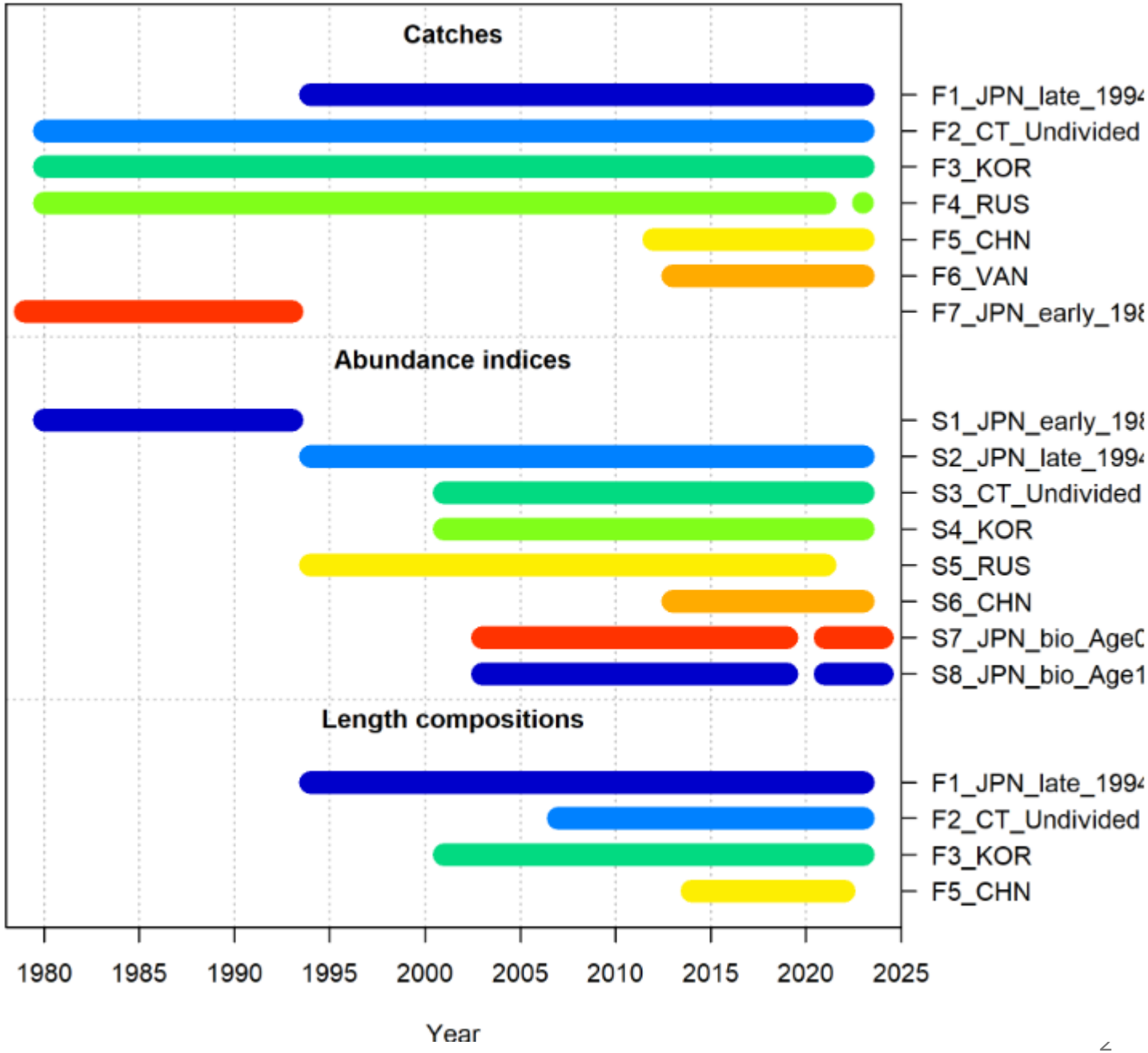
SS data confidentiality and access to NOAA technical help for SS3

Larry Jacobson (Consultant)

NPFC-2024-SSC PS14

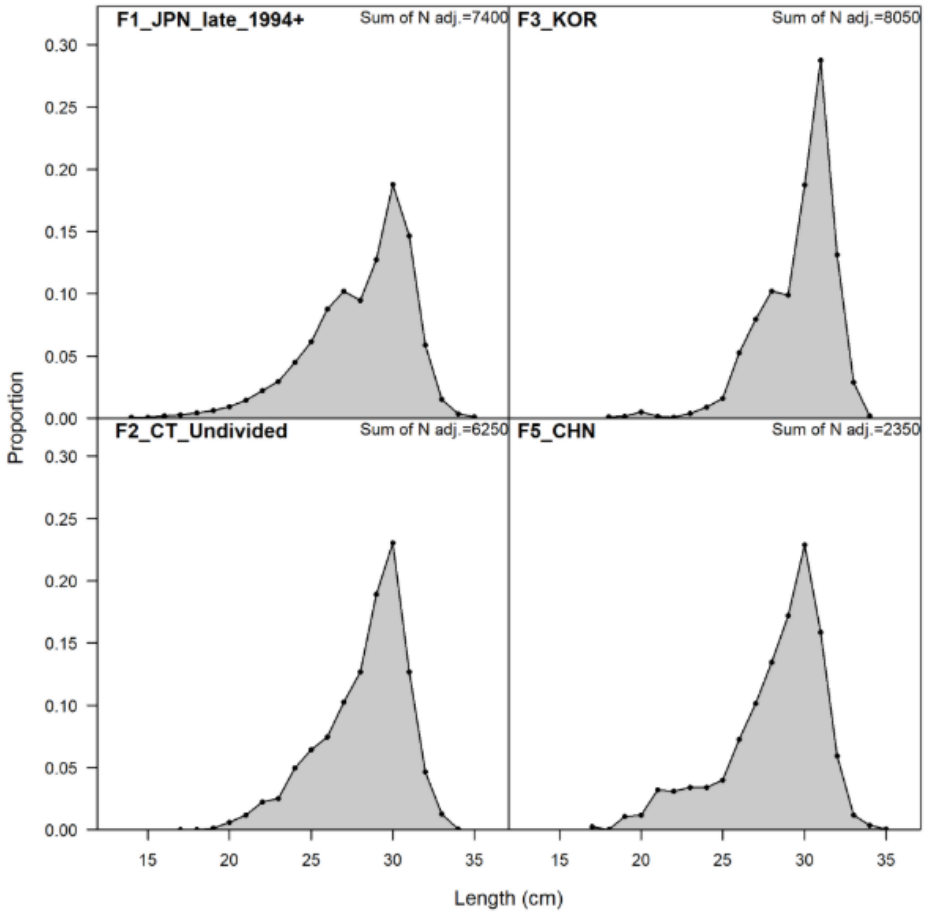
11–13, 16 December 2024

All of the data in SS3



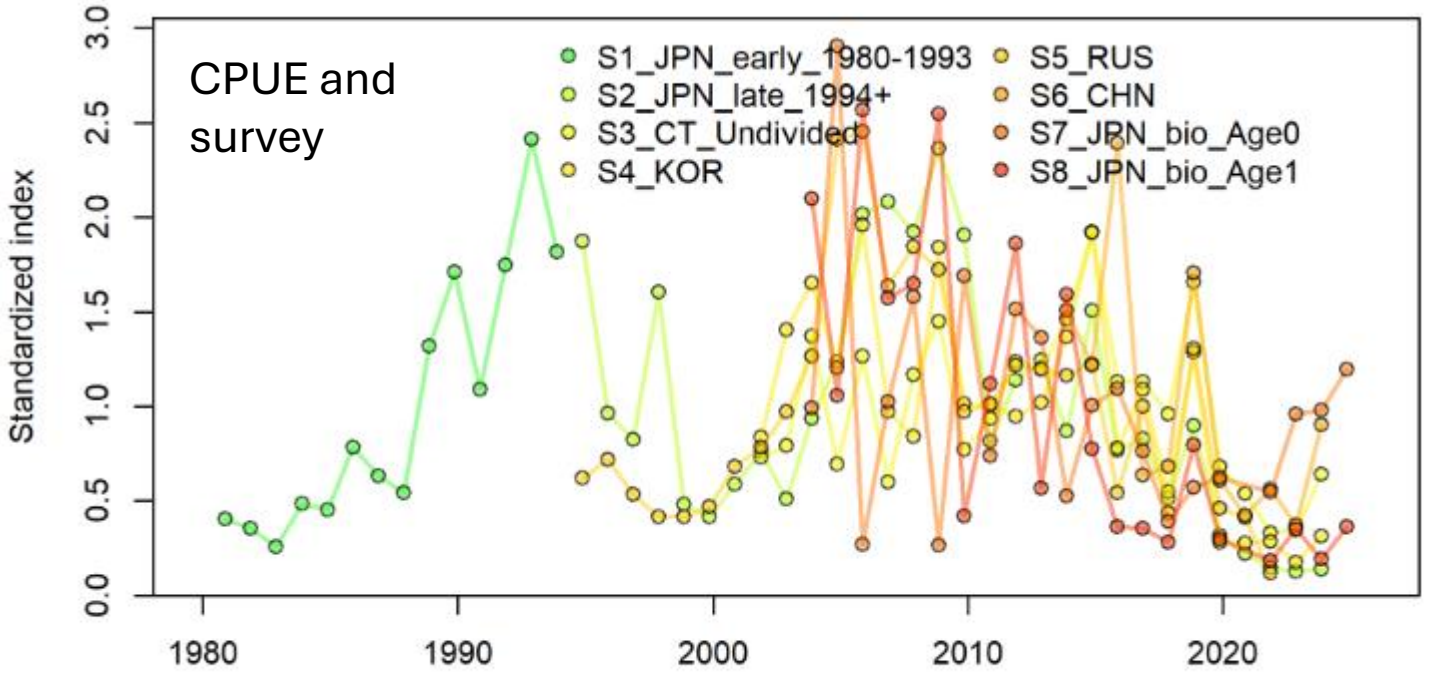
Data in SS3 for Saury

Length data (average over years)

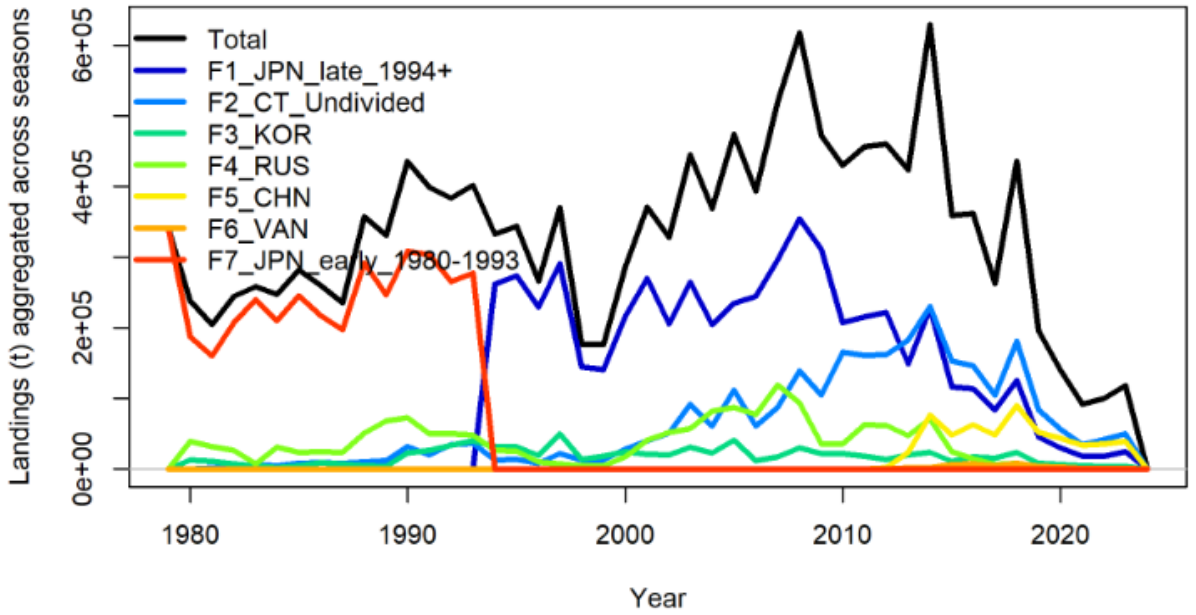


Length comp data, aggregated across time by fleet.

CPUE and survey



Length data (annual)



Pity the poor SS modeler (problems inevitable)…

- SS3 models are difficult, time consuming even for experienced analysts
  - Harder for beginners like NPFC and unusual stocks like Pacific saury
  - Easier after first model is completed
- C++ and R code problems sometimes found when dealing with features not widely used (e.g. nonlinear CPUE, Gomertz growth curve for saury) or unusual fish/fisheries (e.g. Pacific saury)
  - Testing requires or better with real data
- Days to weeks can be wasted trying to resolve problems without help
  - Solutions may not be possible without expert help from the SS3 Team

Regulations for Management of Scientific Data and Information

Data *and* code???

"If a Member or cooperating non-contracting Party (CNCP) wishes to cite and/or use these data or code for work that is intended to be conducted or shared outside of the NPFC, such Member or CNCP should consult with the provider(s) of the data or code through the Secretariat, stating 1) the data or code subject to the request, and 2) the purpose for which the data or code is intended to be used. The Secretariat should immediately notify the provider(s) of the request. The provider(s) should inform the Secretariat within 30 calendar days whether to accept or reject the request. "

This is my formal request

Two roads to SS help

- 1) Public list server SS3-Forum https://vlab.noaa.gov/web/stock-synthesis/public-forums/-/message_boards/message/37861965
  - Anyone in the world can send questions, problems, suggestions etc. to SS3 Team at NOAA and forum members
  - Response includes answers, advice, r code, examples etc. from interested list members
  - ***I recommend colleagues interested WG NSAM colleagues join***
  - See my saury interactions past and future
  - I ***hate*** list servers but this is a good one
- 2) Private communication with the NOAA SS3 Team (nmfs.stock.synthesis@noaa.gov)
  - Team = 3 programmers
  - Rick Methot (SS3 author, NOAA Senior Scientist for Stock Assessments, affiliate professor at the University of Washington) with about 40 years of SS experience
  - Ian Taylor (R4SS author, at least 20 years of SS experience)
  - One other I don't know

List servers *vs.* SS3 Team for Saury

- I used the list server for saury problems in SS and R4SS (Gompertz growth and nonlinear CPUE)
- Help with Gompertz came fast from SS3 users around the world.
- SS3 Team, programming and testing on actual data required for nonlinear CPUE
- I did not share NPFC data but the solution was delayed
  - Had to replicate the problem in data files for another species
  - Finding data for testing was time consuming and frustrating
- SS3 Team responded quickly and are releasing a new SS3 version

NPFC needs to find a way to share test data and make scientific progress

- WG NSAM needs a way to share data for testing by SS3 Team
  - Help is faster and more certain if models and data are available
  - Model development by new users cheaper, easier and more likely to be “correct” with access
  - ***Lack of support increases risk of failure***
- The SS3 team often works with confidential data, see message from Rick Methot to me on Dec. 6, 2024:
  - “*Larry, Our core group of SS3 developers is very accustomed to dealing with confidential data. It is secure with us whenever you provide explicit instructions that it must not be shared. Rick*”
- Methot is senior in NOAA, experienced, reliable and trustworthy. So is the team.

Basis for suggestion

- Complicated software (e.g. cell phones) require technical support for effective use
- US (NOAA) is a member of NPFC but scientific contributions modest to date
  - The SS Team would be valuable support to NPFC at little cost to US
- SS3 team helps scientists from any country and encourages use of SS3 for any species in the world
- Team is is used to dealing with confidential data and has a non-disclosure policy
- Risk of accidental data disclosure by the SSC Team same or less than risk for any NPFC member or contractor
- Serious damage unlikely if accidental disclosure occurs, particularly since comprehensive assessment reports are freely available
- Most data are published in tables and figures inside of annual stock assessment reports or can be obtained by digitizing figures
- SS3 Team Decision to share data for testing will pay off
  - NPFC is starting to invest in SS3 and it may be useful for other stocks
  - SS3 useful for assessment modeling and MSE/harvest control work—all important topics for NPFC at present

Possible approaches if full disclosure is a problem

- Hide fleet names in data file (e.g. “CHN” becomes “A”)
- Remove the last five years of data (all types)
- Find or build another model that can be supported by NPFC stock assessment staff
- Reach a formal or informal agreement with NOAA, like the agreement with contractors
- Modify my contract as follows: *“Parties understand that NPFC disclosure policies affect model development because they limit or prevent access to technical support. These policies may prevent or substantially delay completion of the SS3 modeling project for Pacific saury. The Consultant may terminate this contract without prejudice and receive payment for work already completed if technical problems arise which, in the Consultants view, will substantially delay or prevent completion of the project. In that case, the Consultant will provide NPFC with all data files, R code and other intellectual property in the Consultant’s possession that is related to the project.”*

Thank you!