**Scale-jumps: a new metric to describe historical retrospective patterns in ICES stock assessments**

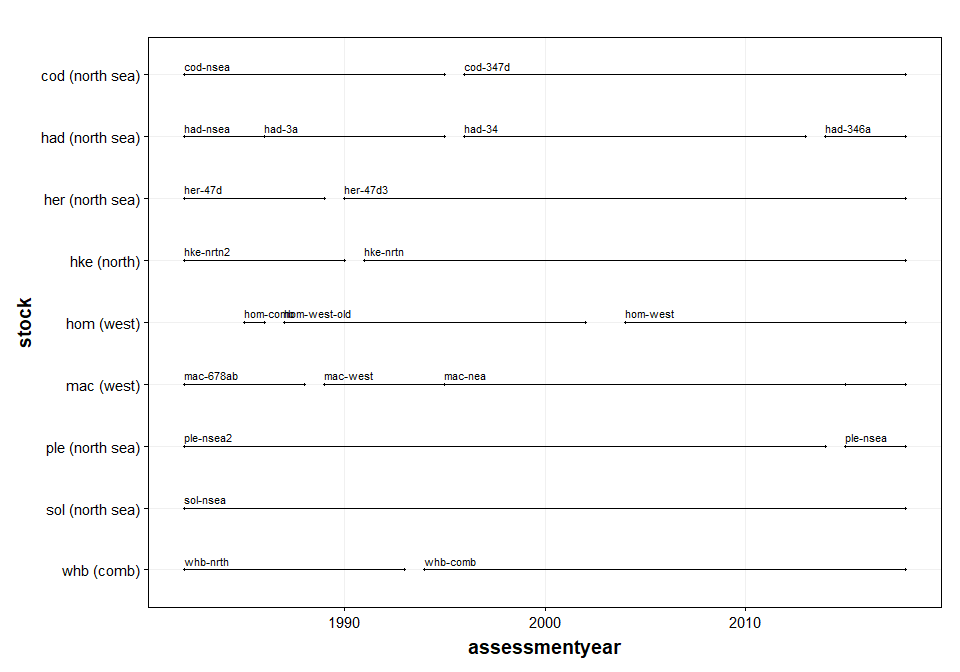
**Tables**

*Table 3.1: Percentage assessments with scale jumps of more than 10% expressed by decade. Between brackets: number of assessments within the decade.*

stock 1980 1990 2000 2010   
----------------- ---------- ---------- ---------- ----------  
cod (north sea) 33% (6) 10% (10) 33% (9) 33% (9)   
had (north sea) 29% (7) 0% (10) 40% (10) 33% (9)   
her (north sea) 33% (6) 10% (10) 10% (10) 22% (9)   
hke (north) . 33% (6) 0% (10) 100% (7)   
hom (west) 100% (1) 50% (6) 40% (5) 56% (9)   
mac (west) 17% (6) 25% (8) 10% (10) 11% (9)   
ple (north sea) 50% (6) 0% (10) 30% (10) 11% (9)   
sol (north sea) 50% (6) 10% (10) 0% (10) 0% (9)   
whb (comb) 67% (3) 29% (7) 30% (10) 33% (9)

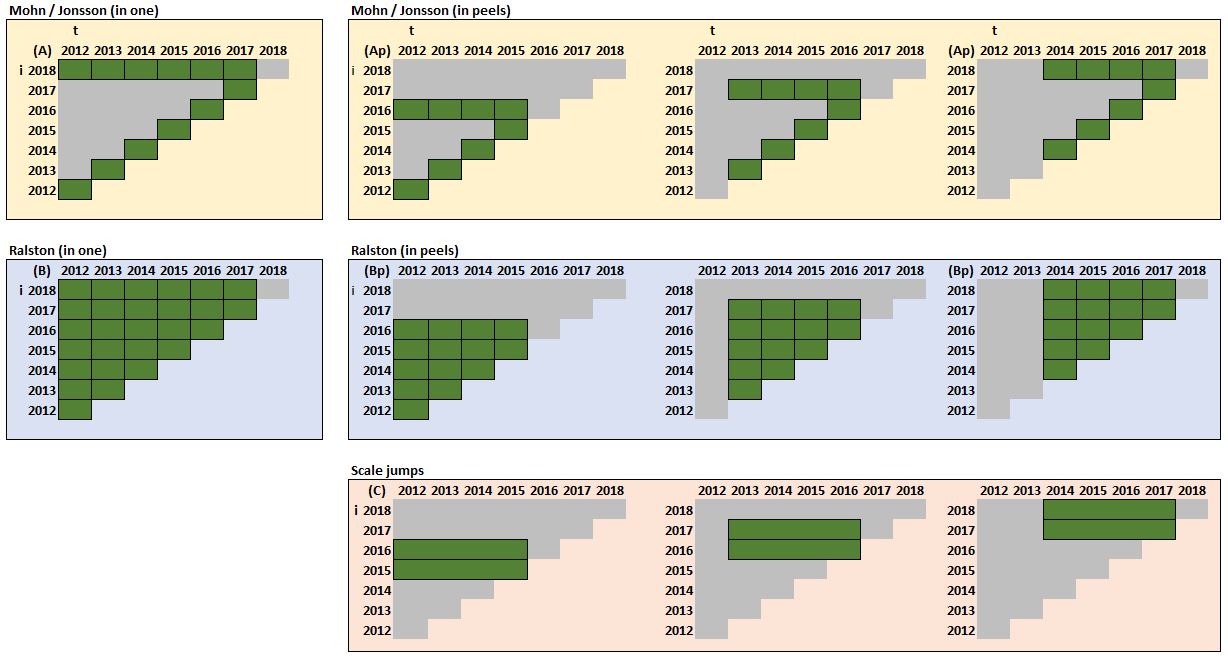
##### page break

**Figures**



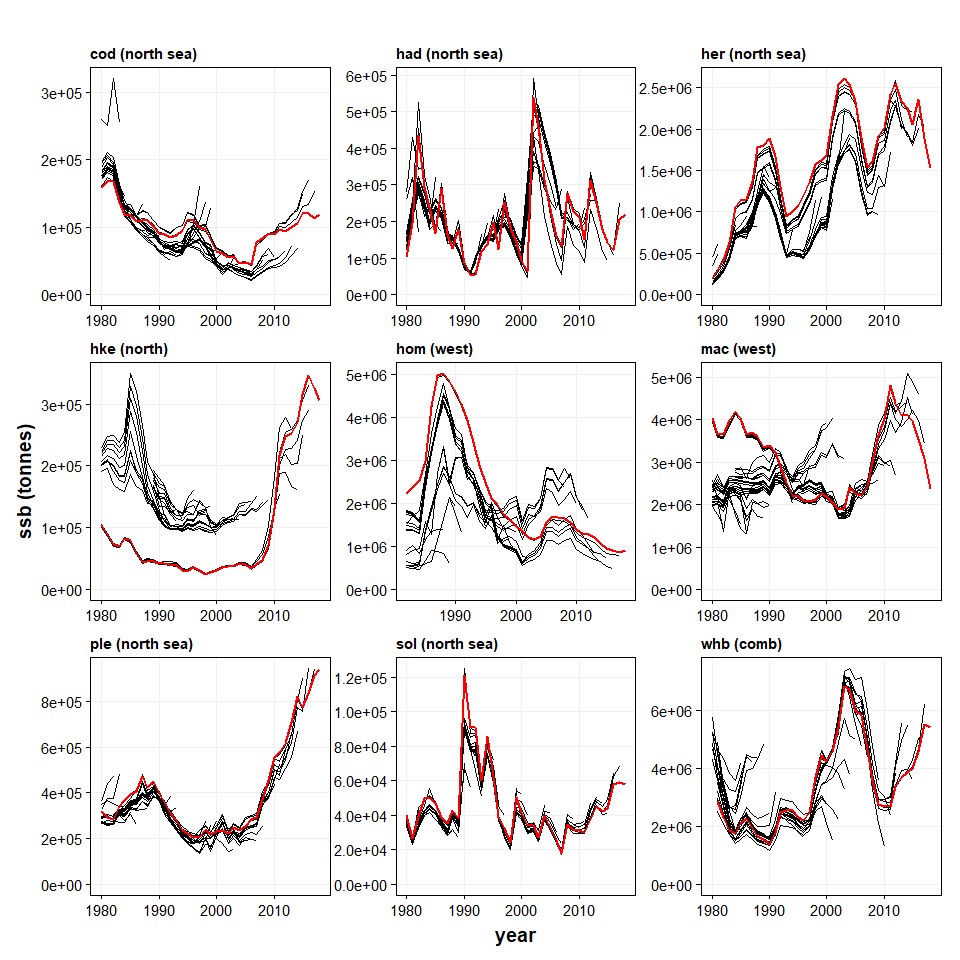
*Figure 1.1: Definitions of stocks (by year) that have been ‘lumped’ into meta-stocks for long time-series*

##### page break



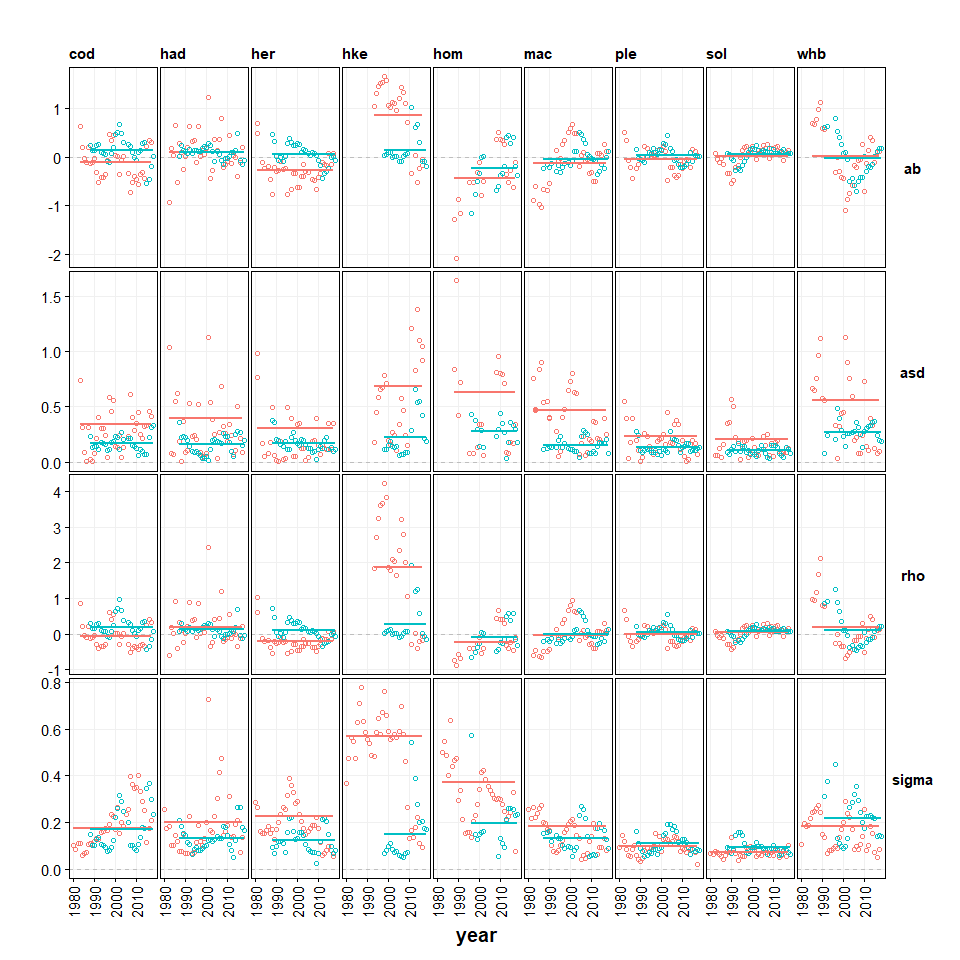
*Figure 1.2: Comparison of three different types of retrospective metrics. t refers to the year for a stock metric X, i refers to the assessment year. Diagram A is applicable to Mohn’s rho and Jonsson and Hjorleifsson’s retrospective metrics while Ap refers to the same metrics calculated in peels (in this example of 4 years), B refers to Ralston’s et al metrics and Bp to the same metrics in peels and C refers to the scale differences calculated in this example over 4 years (this paper)*

##### Page break



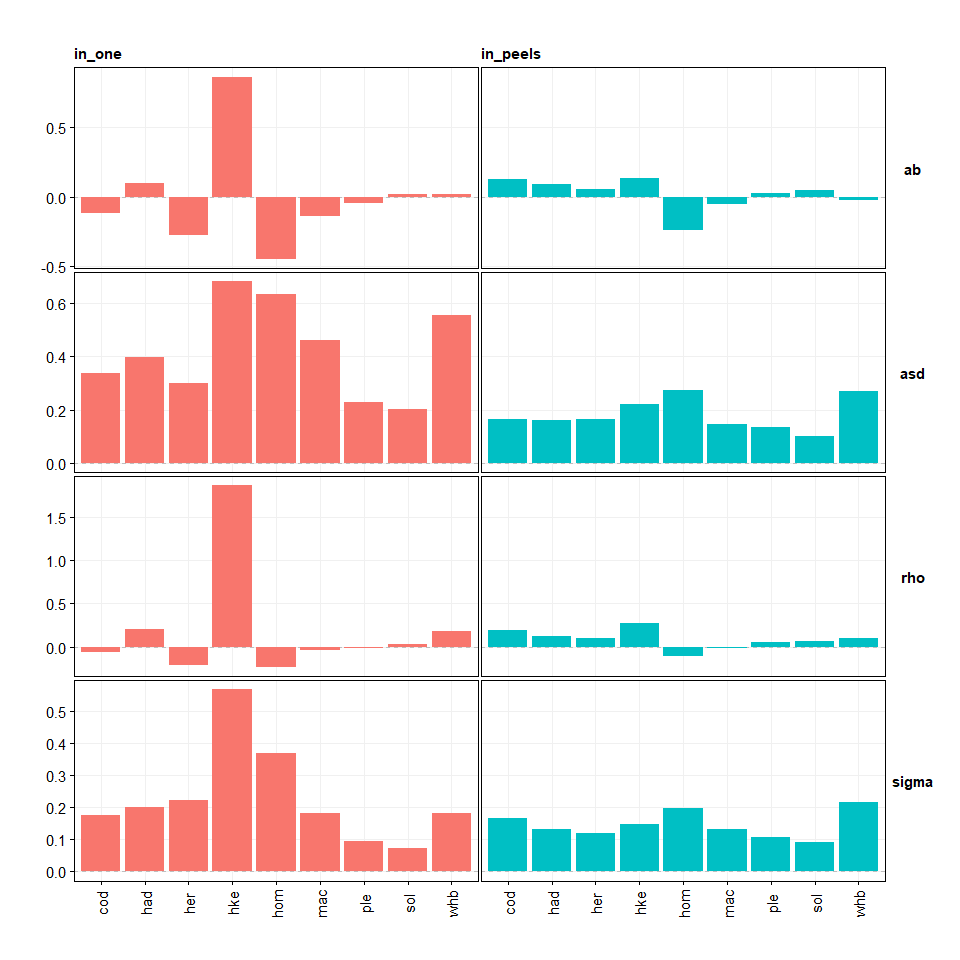
*Figure 1.3: Historic retrospective analysis of the nine stocks analysed. For each stock, the most recent assessment is in red and the previous assessments in black. The end point of each assessment is indicated by a dot. Stock definitions may have changed over time, but this is not included in the graphs*

##### Page break



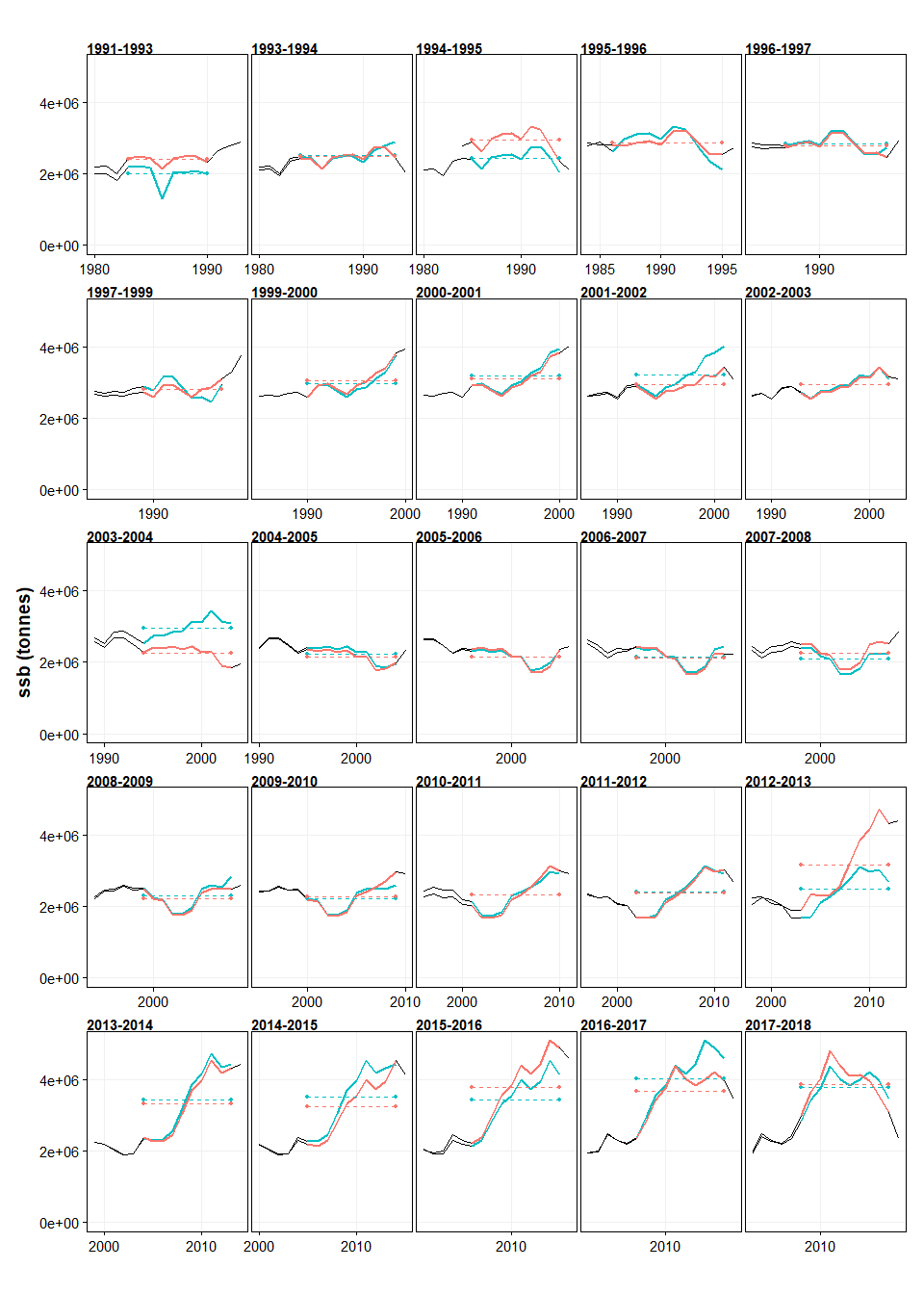
*Figure 2.1: Retrospective metrics for 9 species derived from historical assessment records. Metrics ab (average bias) and asd (average variation) from Jonsson & Hjorleifsson 2000, rho from Mohn 1999 and sigma from Ralston et al 2011. In red: calculated values based on the whole time-series of historical data. In blue: based on retrospective peels of 5 years. Metrics ab, asd and sigma on a log scale, rho on a ratio scale.*

##### Page break



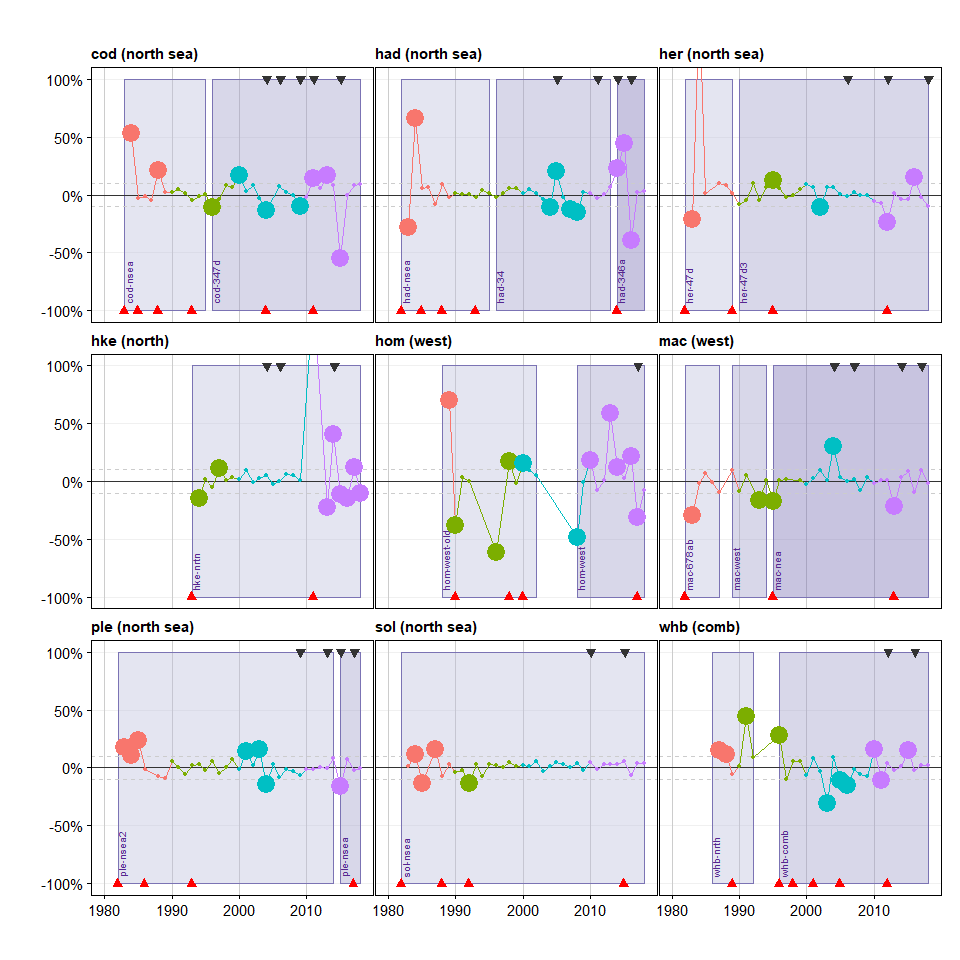
*Figure 2.2: Summary of retrospective metrics for 9 species derived from historical assessment records. Metrics ab (average bias) and asd (average variation) from Jonsson & Hjorleifsson 2000, rho from Mohn 1999 and sigma from Ralston et al 2011. In red: calculated values based on the whole time-series of historical data. In blue: based on retrospective peels of 5 years. Metrics ab, asd and sigma on a log scale, rho on a ratio scale.*

##### Page break

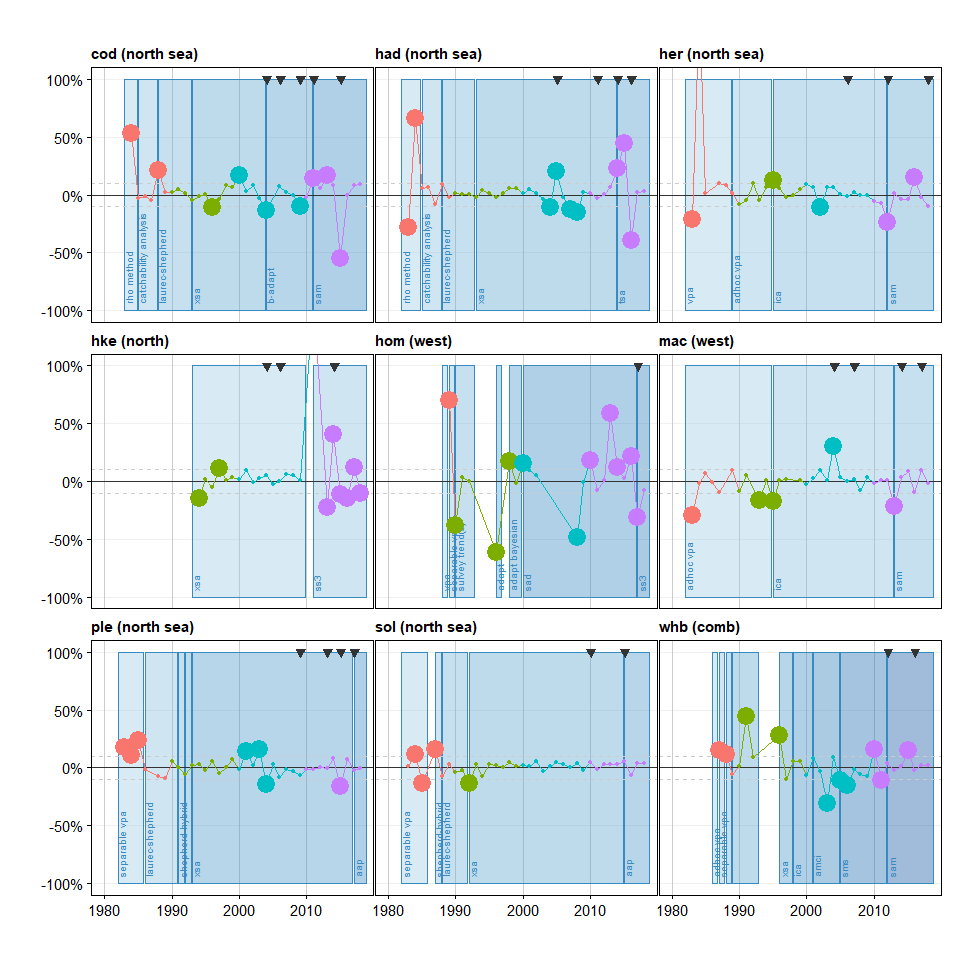


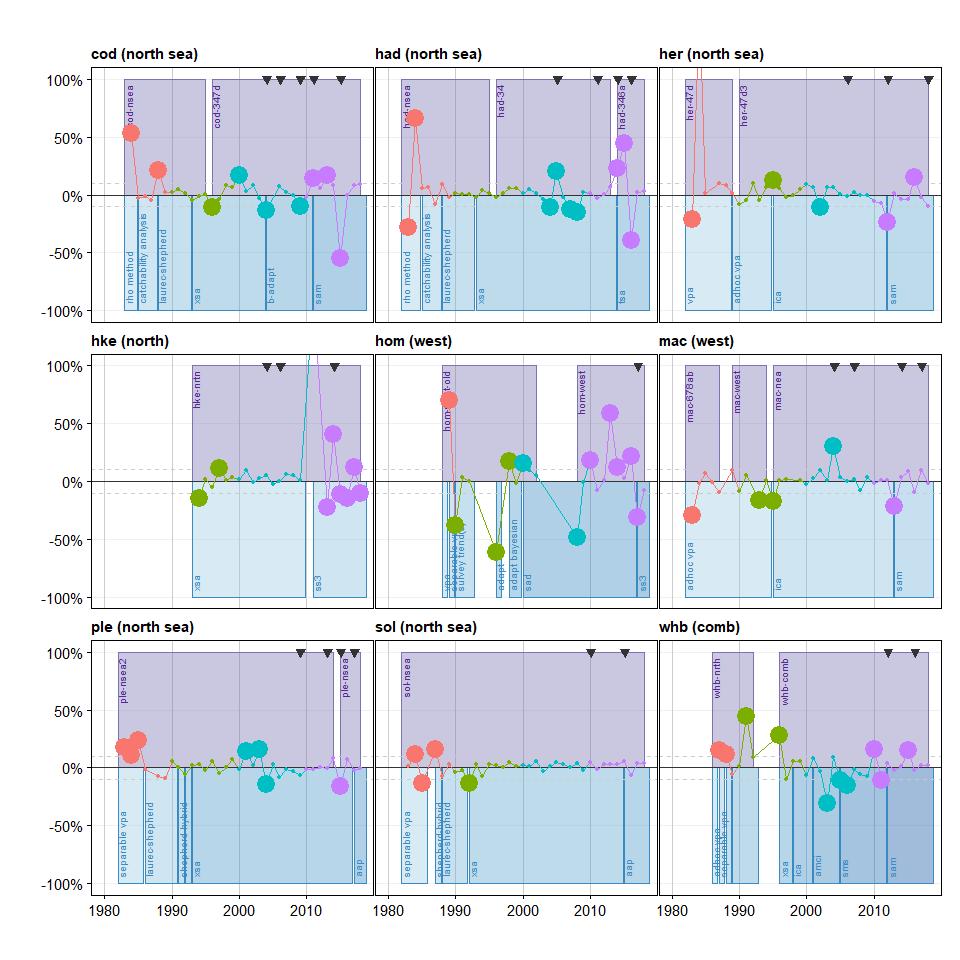
*Figure 3.1: Example of pairwise assessment comparisons for Northeast Atlantic mackerel over a period of 10 years in common. The dotted lines indicate the average values of the different assessments which are used to calculate scale differences between assessments.*

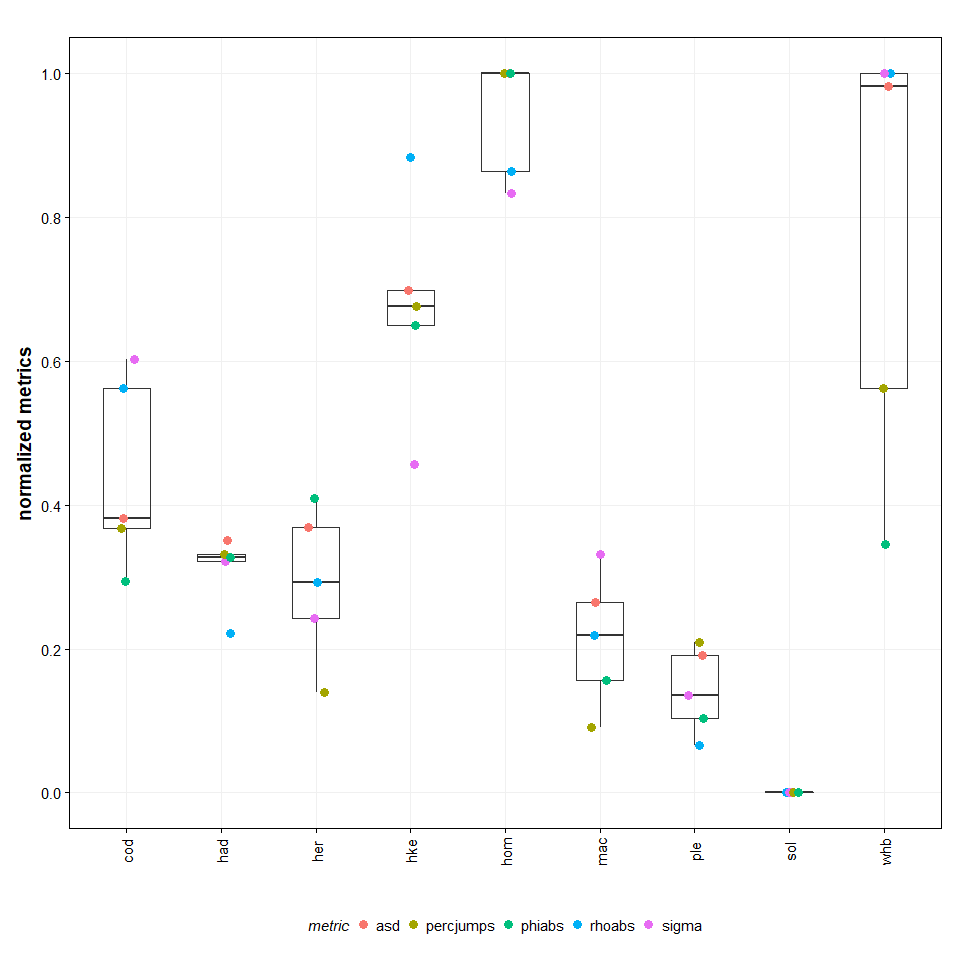
##### Page break



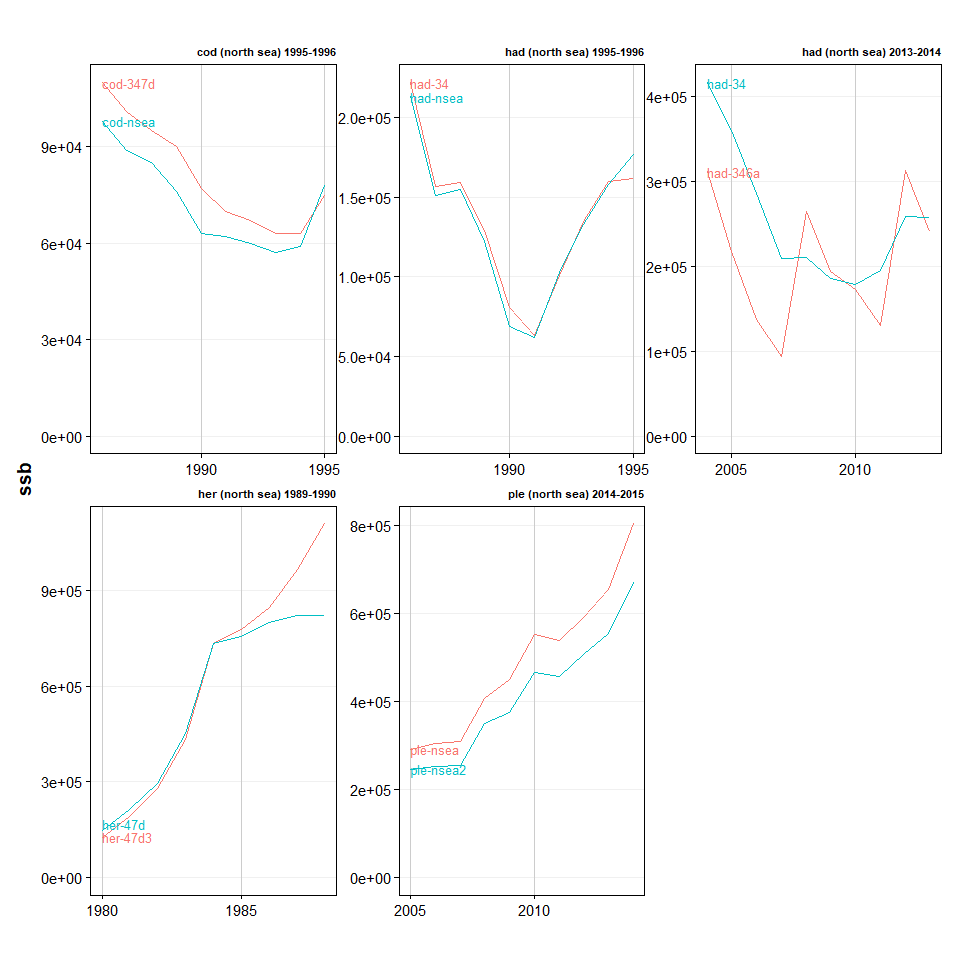
*Figure 3.2: Retrospective scale difference (sigma) for 9 species derived from historical assessment records. Results from pairwise comparisons over periods of 10 years. Large dots refer to years when the scale difference is larger than 10%. Colours indicate the decades. Black triangles indicate benchmark assessment years. The bottom part of the panels shows the assessment methods being deployed, the upper part of the panels the stock definitions.*



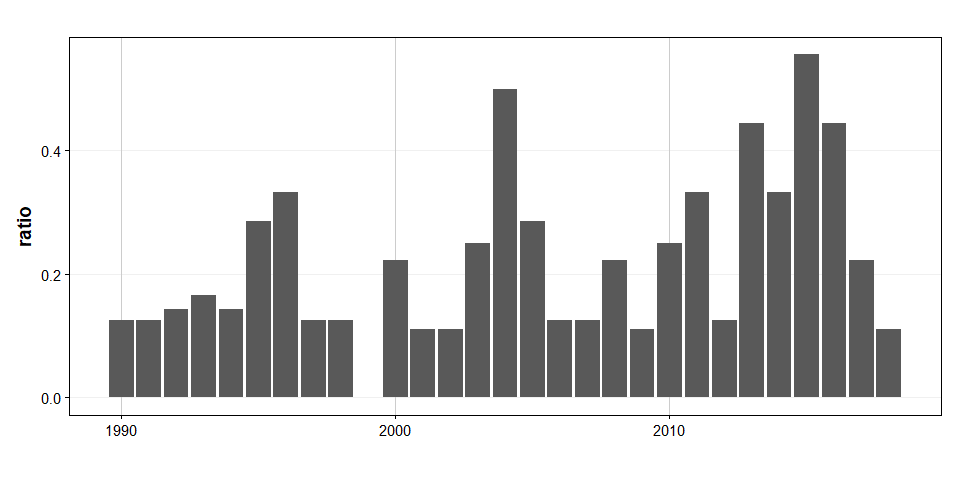




*Figure 3.3: Comparison of retrospective uncertainty metrics from historical assessment records, calculated in peels, normalized between 0 and 1. Asd from Jonsson & Hjorleifsson 2000, rhoabs as absolute value of rho from Mohn 1999, sigma from Ralston et al 2011, percjumps as the percentage of jumps (this paper) and phiabs as the absolute value of phi (this paper).*



*Figure 4.1: Comparisons of assessments (SSB) where stock definitions changed. Green is old stock definition, red is new stock definition*



*Figure 4.2: Ratio of scale-jumps in assessment relative to the number of assessment by year.*