

Not to be cited without prior reference to the first author

The determination of biomass reference points for North Sea plaice: The influence of assumptions about discards, weight, maturity and stock-recruitment relationships.

Sarah B. M. Kraak, Loes J. Bolle, and Adriaan D. Rijnsdorp

Netherlands Institute of Fisheries Research (RIVO), PO Box 68, 1970 AB IJmuiden, The Netherlands. Phone; +31 255 564783, fax: +31 255 564644, e-mail: sarah.kraak@wur.nl

Abstract

Many fisheries are managed with reference points. The limit biomass (B_{lim}) is defined as the spawning stock biomass (SSB) below which recruitment is impaired or stock dynamics are unknown. Management harvest rules are designed such that B_{lim} should be avoided with a high probability. In order to do so, management action must be taken at a higher biomass, such as the precautionary reference point B_{pa} . B_{lim} is usually determined through inspection of the historic relation between recruitment and the parental SSB. However, the perception on the historic number of recruits, through the stock assessment, is influenced by assumptions on true catch numbers, *e.g.* discards. Similarly, the perception of the historic SSB is influenced by assumptions on fish weight and fish maturity. Different assumptions may change the shape of the stock-recruit relationship. Furthermore, decadal changes in the abiotic and biotic environment (regime shifts) may influence the recruitment potential of the spawning stock. We investigated the influence of assumptions about discards, fish weights, maturity, and the choice of the reference period with regards to possible decadal changes in the environment on the determination of B_{lim} of North Sea plaice. The different assumptions indeed changed the shape of the stock-recruit relationship, and led to different B_{lim} reference points. In some cases B_{lim} was close to the lowest observed SSB, in other cases recruitment seemed to have been impaired at observed larger SSBs. We recommend that the calculation of reference points be based on stock assessments that incorporate improved biological realism and improved discard estimates and that the reference period for this calculation correspond to external insights about decadal changes in ecosystem productivity. We also note that in order to be internally consistent the actual values of the reference points should be re-calculated each time that the annual stock assessment results in a (major) change in perception of the historic stock status.

Key words: North Sea plaice, reference points, limit biomass, discards, maturity, weight, stock-recruitment, regime shifts.

Introduction

The Precautionary Approach to fishery management provides the framework for the fishery management advice provided by the ICES [International Council for Exploration of the Sea] Advisory Committee on Fishery Management [ACFM] (ICES, 2001). The ICES approach is that in order to have stocks and fisheries within safe biological limits, the probability should be high that

- the spawning stock biomass (SSB) is above a limit value (called B_{lim}) below which recruitment (R) becomes impaired or the dynamics of the stock are unknown, and
- the fishing mortality (F) is below a limit value (called F_{lim}) that will drive the spawning stock to that biomass limit B_{lim} .

With regards to the phrase ‘high probability’ it should be noted that the decision on the acceptable risk is not a science issue but should be decided by managers and stakeholders. No quantitative indication has been given how high a ‘high probability’ should be.

O’Brien and Maxwell (2002a) stated that this approach assumes a rather simple model of stock dynamics; *i.e.* that there is a specific value of spawning stock biomass below which recruitment is impaired, and that fishing mortality is the only external factor which influences the size of the spawning stock. Regarding their first point, it has to be kept in mind that the PA framework is not a scientific enterprise to unravel the biology of stock dynamics; instead the aim is the definition of operational limit reference points for management purposes. We will discuss their second point later in the context of ecosystem changes on a decadal time scale and climate change.

Because of uncertainty in the annual estimation of SSB and F, ICES defines the more conservative operational reference points, B_{pa} and F_{pa} (the subscript pa stands for precautionary approach). When a stock is estimated to be at B_{pa} the probability should be high that in reality it is above B_{lim} . Similarly, when F is estimated to be at F_{pa} the probability should be high that in reality it is below F_{lim} . In other words, if the assessed F is at or below F_{pa} , the risk is low – taking assessment uncertainty into account – that the real exploitation will lead to impaired recruitment. Thus B_{pa} and F_{pa} are operational values that ensure with high probability that exploitation is sustainable. However, this approach only takes the uncertainty in the annual estimation of SSB and F into account, but not the uncertainty in the estimation of B_{lim} and F_{lim} . B_{lim} and F_{lim} may be considered estimates of properties of nature, reflecting the reproductive capacity of a fish stock under the current natural regime. The distances between B_{lim} and B_{pa} and between F_{lim} and F_{pa} reflect our ability to measure the present SSB and F, and are thus related to data quality, estimation methodology, and the perception of acceptable risk.

Whenever F is above F_{pa} , the stock is considered to be harvested outside safe biological limits and ICES will advise that F should be reduced. Whenever a stock is estimated to be below B_{pa} , it is considered to be outside safe biological limits and ICES will advise that F should be reduced. Stocks that are both above B_{pa} and below F_{pa} are considered to be inside safe biological limits. ICES intends that B_{pa} and F_{pa} are boundaries to the safe domain, and not targets.

The values of B_{lim} , F_{lim} , B_{pa} , and F_{pa} are estimated based on the history of the stock and the fishery. B_{lim} has to be estimated first, because F_{lim} is defined with reference to B_{lim} , B_{pa} with reference to B_{lim} , and F_{pa} with reference to F_{lim} . As a method to determine B_{lim} from an R-SSB plot, ICES (2003) proposed to use the “segmented regression” or “hockey-stick method”

(O'Brien and Maxwell 2002a, 2002b). With this method a breakpoint SSB, labeled S^* , can be identified below which recruitment declines linearly to zero at $SSB = 0$, and above which recruitment is assumed to be independent of SSB. If the breakpoint is not significant (*i.e.* if the segmented regression does not explain significantly more variance than the arithmetic mean) it should be assumed that recruitment is not impaired at any of the observed SSBs, and the lowest observed value of SSB (B_{loss}) should be taken as the value of B_{lim} (as the SSB below which the dynamics of the stock are unknown). F_{lim} can be derived in a straightforward way from B_{lim} . The p_a values are determined from the lim values by fixed multipliers (ICES 1997, ICES 1998); it should, however, be born in mind that this is not a scientific way of taking account of assessment uncertainty, and that there is no guarantee that this procedure results in p_a values that ensure with high probability that lim values are actually avoided. As yet, no readily applicable method exists that takes the true assessment uncertainty into account. In this study we only deal with the estimation of B_{lim} .

As stated above, B_{lim} may be considered an estimate of a property of nature, reflecting the reproductive capacity of a fish stock under the current natural regime. As such, the estimate will depend on our perception of nature, more specifically, on our estimates of historic recruitment and historic SSB. These are themselves estimated through a stock assessment, with its own uncertainties and assumptions. Our perception on historic recruitment, for example, is influenced by whether or not we assume that variable discard mortality occurs. If mortality is underestimated (*e.g.* as a consequence of assuming no discard mortality) recruitment is underestimated as well. If discard mortality is constant over time, the underestimation will only result in an incorrect scaling of R and the estimated B_{lim} may still be useful. However, if discard mortality is variable over time -- a more likely hypothesis -- the true R -SSB pattern may have a different shape than the estimated one, resulting in a misleading estimate of B_{lim} . Similarly, if SSB is calculated with too high or too low estimates of maturity at age or weight at age and the bias is constant, SSB will be incorrectly scaled but the estimated B_{lim} may still be useful. However, if the bias is variable over time, again the true R -SSB pattern may have a different shape, and the estimate of B_{lim} will be misleading.

There are indications that in North Sea plaice discard mortality (van Keeken *et al.* 2004), growth (Rijnsdorp *et al.* 2004), and maturity (Grift *et al.* 2003) have been variable over time (discard mortality appears to have increased, growth appears to have first increased and then decreased, maturity at age appears to have increased). This study will investigate the influence of assumptions regarding discard mortality, growth, and maturity, on the shape of the estimated R -SSB relationship and the estimated B_{lim} (see also Kell and Bromley 2004). Due to the lack of discard sampling data, the stock assessment of North Sea plaice did not include discard estimates until 2004, in which year new reference points were established as well. Maturity is assumed to be constant over time in the North Sea plaice assessment. Weights at age in the stock used for the North Sea plaice assessment are estimated from market samples; these will be overestimates for the younger age groups of which only the largest individuals are selected by the commercial gear (and this bias is unlikely to be constant over time). In this study we will compare R -SSB relationships derived from standard assessment practices (before 2004), with those derived when discard mortality is estimated through a method of discard reconstruction (van Keeken *et al.* 2004), and when weights at age in the stock and maturity are estimated from both market and research vessel samples.

The R -SSB relationship attempts to capture the functional relationship between the parent stock and the number of offspring produced by that parent stock that survive to the stage of recruitment to the fisheries. If there are ecosystem changes that affect egg production or the survival until recruitment, we need to take account of these ecosystem changes in our analysis

of the functional relationship between stock and recruitment. The recent biological analysis of changes in growth of plaice has indicated that the productivity of the North Sea ecosystem showed decadal changes (Rijnsdorp *et al.* 2004): growth rate of 1-4 year old plaice and sole was relatively high in the 1970s and early 1980s. There is an accumulating body of evidence that shows the importance of ocean climate for the marine ecosystems and that documents decadal changes in marine ecosystems. Amongst others, Reid *et al.* (1998) showed a step-wise change in phytoplankton colour around 1987/1988 and suggested that these changes were related to changes in the atmospheric forcing (NAO-index). The step-wise change around 1988 was found by Weijerman and Lindeboom (2005) in a number of time series of abiotic as well as biotic variables. Other step-wise changes were shown to have occurred around 1979 and (perhaps) 1998 (Weijerman and Lindeboom 2005). The causal pathway of these changes, however, remains obscure. The variations in growth seem to coincide with the step-wise changes reported by Weijerman and Lindeboom 2005, but they may also be related to the changes in nutrient enrichment of the coastal waters (Boddeke and Hagel 1995). These step-wise changes may also have influenced egg production and survival to the stage of recruitment, although factors that affect eggs and larvae are not necessarily the same as factors that affect growth in the demersal stages. Furthermore, it is generally accepted that the sea temperatures in the North Sea have been exceptionally high since the late 1980s, in particular during the winter. The higher temperature conditions could play a role in the causal mechanisms of survival of eggs and larvae, as it is well established that the mortality rate of plaice eggs is positively related with temperature (van der Land 1991) and that the timing of spawning is affected by temperature (Rijnsdorp and Witthames 2005) which may affect the match-mismatch of the larvae and their pelagic food (Cushing 1990). For plaice stocks around the UK, including North Sea plaice, negative relationships exist between sea surface temperature and recruitment (Fox *et al.* 2000).

As stated earlier, B_{lim} should reflect the reproductive capacity of a fish stock under the current natural regime but it is generally estimated over a range of productivity regimes. In the evaluation and revision of the biological reference points it is therefore relevant to take account of the (possibility of) changes in ecosystem productivity. If evidence indicates a change in terms of recruitment it seems justifiable to assume that the R-SSB relationship observed since the latest change is more likely to be representative for the forthcoming decade, than the R-SSB relationship of a longer time period (now: 1957-2003). This study investigates the evidence of changes in recruitment coinciding with the step-wise changes reported by Weijerman and Lindeboom (2005) and Reid *et al.* (1998). A B_{lim} estimate will then be calculated based on the R-SSB relationship observed since the latest change, and its appropriateness will be discussed.

This study will not come forth with a proposal for a new B_{lim} reference point; instead it proposes a direction. Work on better estimates of discards, maturity, and reproductive potential (based on female SSB) is on its way. Furthermore, more external evidence is needed to establish whether and when decadal changes have taken place. This study explores the impact of the various assumptions on the determination of B_{lim} , and its consequences.

Methods

The time series of R and SSB used in these analyses were generated through application of the XSA stock assessment to the data that were available at the ICES Working Group on the Assessment of Demersal Stocks in the North Sea and Skagerrak (WGNSSK) in 2004 (ICES 2005). Several modifications were made to the data in succession, and each time XSA was

run anew to generate a different set of R and SSB time series for each modified data set. The eight data sets were:

Data set	WGNSSK (ICES 2005) without discards	Modified weights at age in the stock	Modified maturity at age	Addition of discard estimates method 1	Addition of discard estimates method 2
1	X				
2		X			
3			X		
4		X	X		
5		X		X	
6		X	X	X	
7		X			X
8		X	X		X

(modifications and additions are with respect to data set 1).

Modification of weights at age in the stock

Weights at age in the stock used by the WGNSSK for the North Sea plaice assessment are estimated from market samples; these will be overestimates for the younger age groups because only the largest individuals are selected by the commercial gear (and this bias is unlikely to be constant over time). Therefore, we estimated weights at age in the stock from research vessel survey samples for the younger age groups (ages 1-4). The selectivity of the gear (20mm or 40mm stretched mesh) of the research vessel survey results in smaller fish being caught, and therefore better estimates of the weights at age of the younger age groups can be derived from these samples.

Modification of maturity at age

The WGNSSK calculates yearly SSB as the sum of the estimated numbers at age multiplied by the estimated stock weights at age multiplied by a fixed maturity ogive (proportion mature at age: 0, 0.5, and 0.5 for the ages 1, 2, and 3 respectively, and 1 for all older ages). However, several studies have shown that in North Sea plaice female maturity at age is not constant over time. Maturity at age not only fluctuates, but it has increased over the past half century, meaning that in recent times females mature at younger ages than in the past (Grift *et al.* 2003). Here we derive an annually varying maturity ogive, which can be used to calculate SSB while taking this change into account.

The approach taken is that, for each year, female maturity at age and male maturity at age are estimated, as well as the sex ratio at age. The new yearly maturity ogive is taken to be female maturity at age times the proportion of females at that age plus male maturity at age times the proportion of males at that age. Hence, the new yearly maturity ogive is the average of female and male maturity at age weighted by the proportions females and males at age respectively.

Ideally, male and female stock weights at age of the mature component should be used in the calculation of a more realistic SSB as well, instead of a combined weight at age. However, we did not estimate these.

In the discussion we will consider that in fact female SSB only is a better estimate for reproductive potential of the stock.

Male maturity

Male maturity at age was assumed to be constant over time and as observed in a dedicated maturity survey conducted during the spawning period in 1985 and 1986 (Rijnsdorp 1989) (see text table below).

Age	1	2	3	4	5+
Male maturity	0.005	0.51	0.72	0.95	1

Female maturity

In principle, yearly female maturity at age was calculated by raising the Dutch market samples (note that over the past years the Dutch landings contributed about 45% of the total international landings of North Sea plaice). However, this will introduce a bias because at the younger ages, only the larger individuals, which are more likely to be mature, are recruited to the fishery. Therefore, the estimates of female maturity at age were corrected for variations in the size distribution by age and variations in the maturity-length relationship.

The size distribution by age was estimated from the time series of otolith back-calculation of otoliths of female plaice collected in the southern North Sea (around 53°N) during the spawning months (Rijnsdorp & Van Leeuwen, 1992; 1996). Additional samples were processed for the years since 1999. In general, otolith samples were length stratified with up to 4 females sampled of each cm-group in a year. The mean back-calculated length at each age was calculated taking account of the length distribution of the female population during the spawning period.

The maturity-length relationship was then estimated using the model:

$$P = \text{Length} + \text{Length} * \text{Age} + \text{Year} * \text{Age}$$

P is a binomial variable that indicates whether a female was immature (0) or mature (1). A logit link function and binomial error was used to model a sigmoid relationship. The slope of the maturity ogive was allowed to differ between age groups, whereas the intercept of the ogive was allowed to differ between years and ages. With the model, the proportions mature females were fitted by cm-size class and applied to the length distributions by age and year to derive the proportion mature females for each age group.

In order to smooth out yearly variation due to measurement error, for each year the average maturity at age (arcsine square root transformed) was taken from 5 years, *i.e.* the 2 years before, the year itself, and the 2 years after (Figure 1).

Sex ratios

In general, both fishing and natural mortality is higher for males than for females. Therefore, the proportion of females in the population increases with age. Fishing mortality has increased over the past half century, resulting in a higher proportion of females at older ages in recent years compared to earlier years. In this study, both these effects (the age effect and the year effect) should be accounted for.

For the older ages (>4) that are fully recruited to the fishery, the raised Dutch market sampling data were used to estimate the proportions of females at age for each year. In order to smooth out measurement error, for each year the average proportion (arcsine square root transformed) was taken from 5 years, *i.e.* the 2 years before, the year itself, and the 2 years after.

For the younger ages (<4) that are not fully recruited to the fishery, survey data were used to arrive at estimates for the proportion of females at age. The survey data were raised as usual for the tuning indices, but for the sexes separate. The three surveys are (1) SNS from 1970 onwards, (2) BTS Tridens from 1996 onwards, (3) BTS Isis from 1985 onwards. For the younger ages no year effect was assumed. For each of the younger ages, in order to smooth out measurement error, the average proportion (arcsine square root transformed) was taken from all survey-year observations.

For age 4, the proportion of females in each year was estimated as the average (arcsine square root transformed) of the estimates for age 3 and age 5.

Figure 2 shows the resulting proportions female for the years 1960, 1970, 1980, 1990 and 2000. The Figure shows clearly that in more recent years the proportion of females in the population was higher than in the more distant past.

New maturity ogive

As explained before, the new yearly maturity ogive is calculated as female maturity at age times the proportion of females at that age plus male maturity at age times the proportion of males at that age. The resulting combined maturity at varying over time is displayed in Figure 3. When comparing Figure 1 with Figure 3, it can be seen that the increase in female maturity over time is partly cancelled out by the decrease in the proportion of males (Figure 2), which have a higher maturity than females. Overall, maturity has fluctuated and slightly increased over the past half century.

Reconstruction of discards

Until 2004 the WGNSSK did not include quantitative estimates of discarding in its plaice assessment and assumed that the landings were equal to the catch. In 2004 van Keeken *et al.* (2004) presented a working document to the WGNSSK with a method to reconstruct numbers discarded at age. The approach builds on the notion that during its life a cohort will grow through the discard size range. Dependent on the mesh size, minimum landing size, growth rate and the availability of the fish to the fishery, the cohort size distribution may be broken up in different components: fish that are unavailable or escape through the meshes; undersized-fish that are retained in the cod-end; marketable fish that are retained in the cod-end. In brief the method is as follows (for a full description see van Keeken *et al.* 2004). The method assumes that the population consists of a part that is available to the fisheries and a part that is not available to the fisheries. In this case it is assumed that only fish outside the coastal zone (12 nm zone and since 1989 the plaice box) are available to the fisheries. Younger fish tend to live in the coastal zone. For each cm-class population numbers per ICES rectangle are estimated as the sum of the catch rates times the surface area. Length based availability ogives are estimated for individual years, from the proportions of fish outside the coastal zone per cm-class. These ogives are rescaled to an availability of 1 for a fish of 30 cm.

Subsequently, length based mesh selection ogives are estimated (constant over time), based on a selection factor, a selection range and a mesh size. Then length based sorting ogives are estimated, based on a minimum landing size. Because growth was assumed to vary over time, yearly length-at-age is modelled from the BTS survey data. Size distributions at age are modelled as a normal distribution and a CV of 9% (observed from the SNS survey data). From the length-at-age distributions, and the availability and mesh selection ogives, proportions caught are calculated by age and year. Similarly, by subsequently applying the sorting ogive, proportions landed are calculated. For stock assessment purposes new Fs are calculated for the ages 1 to 4, assuming that the fish are fully recruited at age 5 and 6. First the average of F (from the original assessment) at ages 5 and 6 is taken, corrected for the small proportion discards. The resulting F is multiplied with the proportion caught calculated previously, for the ages 1 to 4. With the new Fs the population numbers are calculated, and subsequently the catch and discard numbers.

Discard estimates method 1

The above method was applied to the data available to the WGNSSK in 2004 (ICES 2005). We will refer to the resulting estimates as derived with method 1.

Discard estimates method 2

The final plaice assessment of the WGNSSK (ICES 2005) was carried out with a slight modification of the above. The above procedure was used to reconstruct discards for the period 1957-1998. The actual observations from the Dutch discarding program (van Keeken in preparation) were used for the period 1999-2003. In Figure 4 the resulting numbers of landed and discarded fish are shown. The resulting time series will be referred to as derived with method 2.

Segmented regression

Segmented regressions were carried out on the various R-SSB data sets according to O'Brien and Maxwell (2002a) in S-plus. The most recent three years of recruitment estimates and corresponding SSB estimates were excluded, except if this resulted in very short series; in that case only two years were excluded (see legend of Figure 6).

Results

All calculations of SSB show that in all years from 1993 onwards SSB has been lower than in any year before (Figure 5, Figure 6).

The use of biological estimates of weights, maturity, or both, in the calculation of SSB results in lower estimates of SSB (compare Figures 5a to 5d). The same effect of the use of biological estimates of maturity is seen in the calculations with reconstructed discards included (compare Figure 5e with 5f and 5g with 5h). This occurs because the biological estimates of weights as well as maturity are lower than the values assumed by the WG. As expected, the inclusion of reconstructed discards results in an almost four fold increase in the perception of the number of recruits (compare Figures 5a-d with 5e-h; note the different scale on the y-axis), since a higher number of young fish has to account for the higher catches of young fish.

The stock-recruit relation calculated when reconstructed discards are not included (Figure 5a-d) appears to show signs of impaired recruitment at low SSB values, especially when biological estimates of weights are not included (Figure 5a and 5c). However, when discard reconstructions are included (Figure 5e-h) these signs seem to disappear, especially when the second method of discard reconstruction is used (Figure 5g-h).

The segmented regressions show (Figure 5a-h) that the distance between the breakpoint S^* and the lowest observed SSB value (B_{loss}) varies when different assumptions on weight, maturity, and discards are used in the calculations. With the default data (Figure 5a) the distance between S^* and B_{loss} is relatively big, implying that recruitment has been decreasing with decreasing SSB at the lower end of observed SSBs. The distance between S^* and B_{loss} is even bigger with biological estimates of maturity (Figure 5c), implying that the decrease in recruitment already starts at higher SSB values. Apparently the stock-recruit relationship is not just differently scaled under these different assumptions, but the exact shape of the relation differs as well (not all SSB values are scaled down by the same proportion), whereby S^* ends up in a different place relative to B_{loss} . When biological estimates of stock weights are included, the distance between S^* and B_{loss} is much smaller, regardless of the inclusion of biological maturity estimates (Figure 5b and d). Hence, under this assumption of stock weights almost no impaired recruitment is observed. Also under all assumptions of discard estimates, with or without biological estimates of maturity, the distance between S^* and B_{loss} is quite small (Figure 5e-h) and almost no impaired recruitment is observed.

The time series of SSB and recruitment without reconstructed discards seems to consist of three periods (Figure 6a). The first period, up to 1972, is characterised by relative low recruitment (generally between 200 and 400 million) and large SSB (300 – 430 thousands tonnes). In the following 20 years (from 1973 up to around 1989-1992) recruitment is much higher (generally between 400 and 650 million) and SSB is slightly lower. In the last period, from 1993 onwards, recruitment is low (generally around 200-300 million) and SSB as well. Because the shape of the functional relationship between recruits and SSB is not known, it is not possible to determine whether the decreased recruitment in the last period is due to a decadal change in productivity or whether recruitment is impaired at such low SSBs. When reconstructed discards are added to derive the time series (Figure 6b-c) the decrease in recruitment from 1993 onwards is less apparent, especially when the second method of discard reconstruction is used (Figure 6c). This happens because discards in recent years are estimated to be higher than before, and therefore higher numbers of recruits are needed to account for the observed catches. The increased recruitment from the early 1970s onwards (with slightly lower SSB than before) suggests a regime of higher productivity than in the 1960s. However, the period of increased recruitment does not coincide very well with the period of increased plaice growth that found by Rijnsdorp *et al.* (2004): the increased growth seems to start a few years earlier and ends almost a full decade before the recruitment goes down again (if at all).

If a segmented regression is carried out with the default data but only from 1990 onwards, the distance between S^* and B_{loss} is much smaller than if the whole time series is used (compare Figure 7a with Figure 5a). That means that under the assumption that the period from 1990 onwards is characterized as a period of low productivity, almost no impaired recruitment at low SSB values is observed. The same is true when reconstructed discards are included (Figure 7b) (note that this was also the case when the whole time series was used, as in Figure 5e). When the period from 1972 onwards is assumed to be one characterized by high productivity, it depends on the method of discards reconstruction whether recruitment seems

to be impaired at low observed SSB values (compare Figure 7c with Figure 7d, and note that S^* in Figure 7d is the same as in Figure 5g): only with the first method does recruitment seem to be impaired at a wide range of observed SSB values (however, these low values are all from the most recent years, and might of course be due to the most recent period being one of lower productivity) Apparently, the location of S^* depends strongly on the recruitment estimates in recent years, which are in turn sensitive to whether discards in recent years are reconstructed or observed (method 1 versus method 2).

Discussion

Our explorations clearly show that the reference points for fisheries management are conditional on the stock assessment methodology and the biological realism that is included. As the biological parameters such as growth, maturity and sex ratio show substantial changes over time, the stock assessment should include this biological realism. Also, the variations in discarding have a profound influence on the perception of the stock as well as on the reference points. Since B_{lim} may be considered an estimate of a property of nature, reflecting the reproductive capacity of a fish stock, we recommend that the biological and discard estimates be used, under one condition that remains to be investigated: that the inclusion of biological realism and discards adds information about the real world rather than noise that increases uncertainty. Furthermore, there is evidence that both the abiotic conditions as well as the ecosystem changed since the late 1980s, which could have influenced the productivity of the flatfish stocks. This would support a choice for the time period 1990-present as a basis for the calculation of the reference points. However, we can only recommend this choice under the condition that not too much uncertainty is introduced owing to the use of fewer data points for the estimate of B_{lim} , which remains to be investigated.

It should be noted that any chosen B_{lim} value has been estimated under a particular set of assumptions regarding stock weights, maturity, discards, *etc.* This means that this reference point should only be used to evaluate the state of the stock estimated under the same assumptions. For example, if future stock assessments continue to use the standard values for maturity at age, it makes no sense to evaluate the resulting SSB with respect to a reference point that was estimated with biological estimates of maturity. Even if the shape of the stock-recruit relationship is the same under these different assumptions it is scaled differently, and therefore values such as S^* or B_{loss} will be different. Similarly, if future stock assessments will incorporate discard estimates, it makes no sense to evaluate the resulting SSB with respect to a reference point that was estimated without discards. Even the addition of new data points may change the perception of historic R and SSB. As a consequence, we recommend that each time that the WGNSSK updates the perception on the historic status of the stock (through the annual stock assessment), the estimates of B_{lim} and the derived reference points will have to be updated as well – if not annually, then at least each time that the converged part of the time series undergoes a major change. This is *not* the same as saying that the reference points should be reviewed each year. The same technical definition for the calculation of the reference points could be applied for several years in a row, but to be internally consistent this means that with each new perception on the historic stock status, tightly associated with the assessment of current stock status and the short term forecast, the actual values of the reference points need to be recalculated consistent with the new perception.

This may seem problematic in terms of presentation to the customers of ICES; they might not like an annually fluctuating reference point. A solution to this problem may be to present SSB

not as an absolute value, but as relative to the biomass reference points, and, similarly, to present F as relative to the F reference points. For example, the SSB in a particular year could be estimated at 1.1 times B_{pa} (within safe biological limits) or at 0.9 times B_{lim} (outside safe biological limits). The graphs of SSB and F could then be presented with a y-axis scaled relative to the reference points instead of with absolute values.

The implication of our recommendation to base the future determination of reference points on improved biological realism and discard estimates, is that the annual stock assessment should incorporate this improved biological realism and these discard estimates as well. Research on improved biological realism and improved discard estimates of North Sea plaice is ongoing (RIVO, CEFAS). For example, it can be argued that, since SSB is used as a proxy for the productivity of the stock, it would be better to calculate separate female and male SSBs. Usually, only female SSB is of interest because males are assumed to be always in sufficient numbers (there are exceptions where the stock is sperm limited). Nevertheless, male SSB may yet be important, because the adult sex ratio and the average male size can have important implications for the productivity of the stock via their influence on the mating system (Rowe and Hutchings 2003). In our study we calculated a sexes combined SSB, and we saw that the increase in female maturity over time is partly cancelled out by the decrease in the proportion of males. Separate values of SSB of the sexes could have been estimated by applying the proportions of males and females as estimated in this study to the estimated numbers at age, followed by multiplication with their respective maturities at age and stock weights at age. Better would be to conduct separate assessments for each sex, to account for the differing fishing mortality on the sexes, and arrive at separate estimates of numbers at age by sex followed by multiplication with their respective maturities at age and stock weights at age (this approach was taken by Kell and Bromley 2004). It would be better still to use estimates of stock weights of the mature fraction at each age for each sex. We recommend that separate assessments for each sex be undertaken, based on weight and maturity estimates not only from Dutch data but also from data of other countries, and with improved discard estimates.

For most of the stock-recruit relationships investigated in this study, S^* lies about 20 thousand tonnes above B_{loss} , implying that recruitment seems to have been impaired at only the lowest observed SSB values. However, without discard estimates and with default stock weights at age, S^* was at a much higher value of SSB than B_{loss} (Figures 5a and 5c), especially when biological maturity estimates were used (Figure 5c), implying that under those assumptions recruitment has been impaired at a much wider range of observed SSB values. It seems that B_{lim} should currently lie at a value of approximately 20 thousand tonnes above B_{loss} or at a much higher value. Apparently there is uncertainty in the estimation of this 'property of nature'. Such uncertainty is not taken into account in current ICES practices. The precautionary reference point B_{pa} takes only the uncertainty into account of each current assessment, and ensures that if the estimated SSB in the final year is at B_{pa} the chance is very low that true SSB in that year is in fact below B_{lim} , whose value is assumed to be certain. Furthermore, the uncertainty that is taken into account through the use of B_{pa} is a fixed amount of uncertainty, which does not reflect the uncertainty of each particular year's stock assessment.

The stock assessment in 2004 (ICES 2005) was carried out with discard reconstruction according to method 2, with modified stock weights at age, and with standard maturity at age (as in Figure 5g and Figure 7d). In that year, ICES (2005) chose to set B_{lim} at 160 thousand tonnes, which was estimated to be approximately equal to B_{loss} (161 thousand tonnes) and higher than S^* (134-159 thousand tonnes, depending on the number of most recent estimates

included). Our present estimate of S^* , under the same assumptions is a bit higher (180 thousand tonnes, Figure 5g and Figure 7d), because we used a different method of segmented regression (O'Brien and Maxwell 2002a, 2002b), and we excluded the last three years' recruitment estimates. Therefore, it can be argued that B_{lim} should have been set at 180 thousand tonnes, or at an even higher value if the uncertainty discussed above should be taken into account.

It should be noted that the conclusions on B_{lim} are strongly influenced by the choice of the discard estimation method (methods 1 and 2). For example, in the case that the period from 1972 onwards was taken as reference period (Figures 7c and 7d), the location of S^* depended strongly on the recruitment estimates in recent years, which were in turn sensitive to whether discards in recent years are reconstructed or observed (method 1 versus method 2). Also, the hypothesis of a recent regime shift towards a lower recruitment potential receives stronger support from the data if method 1 is used than if method 2 is used (Figures 6b and 6c). These methods for discard estimation were devised because sufficiently long time series of discard observations were lacking; this necessitated reconstruction of discards (van Keeken *et al.* 2004). However, because the differences between reconstructed and observed values were large for the most recent years, we decided to investigate both time series: method 1 with only reconstructed values, and method 2 with reconstructed values for most of the time series and the observed values for the most recent years (method 2 was used by ICES 2005). In this paper we will not discuss which of the values are more realistic nor give an evaluation of the method of discard reconstruction (van Keeken *et al.* 2004); the method is preliminary and is currently being improved (van Keeken in preparation). We just want to point out that the determination of B_{lim} is sensitive to the discard estimation method used. This fact gives an extra reason for calculating the reference points anew when the assessment makes use of a new method for discard reconstruction.

The assumption that the period from the 1990s onwards is characterized by a lower productivity than before, led to a different estimate of B_{lim} (namely closer to B_{loss}) only when the default data were used, but not when discard estimates were included. Our analysis remains inconclusive about the issue whether decadal shifts in productivity in terms of recruitment have occurred or not. The evidence is especially weak for a reduction in productivity in terms of recruitment starting around 1990. Moreover, the influx of settling larvae into the nursery grounds of the western Wadden Sea was found to be unchanged (Bolle *et al.* in preparation). The indication from this study of the start of a period of a higher productivity around 1972 may correspond to the timing of the increase in plaice growth found in the earlier study (Rijnsdorp *et al.* 2004), but the end of this period of higher recruitment around 1990 – if it ended at all – does not correspond with the timing of the reduction in growth, which happened almost a full decade earlier (Rijnsdorp *et al.* 2004). But it is of course not necessarily the case that factors affecting the survival of eggs and larvae also affect growth in the demersal stages. The tentative decline of productivity in terms of recruitment around 1990 does, however, correspond relatively closely to the step-wise change in phytoplankton colour around 1987/1988 detected by Reid *et al.* (1998), who suggested that these changes were related to changes in the atmospheric forcing (NAO-index), and the step-wise change around 1988, found by Weijerman and Lindeboom (2005), in a number of time series of abiotic as well as biotic variables. Furthermore, the higher temperature conditions since the late 1980s could play a role in the causal mechanisms of survival of eggs and larvae, as it is well established that the mortality rate of plaice eggs is positively related with temperature (van der Land 1991). For plaice stocks around the UK, including North Sea plaice, negative relationships exist between sea surface temperature and recruitment (Fox *et al.* 2000).

Our approach does not deal with the possibility of gradual changes in productivity that are due to gradual changes in the ecosystem as opposed to regime shifts. It will certainly be difficult to detect such gradual changes if the reduction in recruitment is accompanied by a reduction in SSB, especially as long as the functional relationship between SSB and recruitment remains unknown. In such a case it is not possible to say whether a series of low recruitments is caused by recruitment being impaired at the lower observed values of SSB, or by productivity being low owing to a gradual shift. A study on pooled data from all cod stocks of the European Shelf south of 62°N showed that environmental variability (as represented by the NAO-index) has a negative effect on recruitment, but only when the spawning stock is low (Brander 2005). The author recommends that short- and long-term assessments should take account of the current level and trends in the NAO and/or related environmental indices.

The time period that should be chosen as representative for the near future cannot be decided on the population dynamic data of the species itself. As it is likely that the ecosystem shows decadal scale changes in productivity, the reference time period should be based on external insight in such ecosystem dynamics. In our southern North Sea flatfish case, there is evidence that both the abiotic conditions as well as the ecosystem changed since the late 1980s, which could influence the productivity of the flatfish stocks. This would support a choice for the time period 1990-present as a basis for the calculation of the reference points.

Acknowledgements

This study was carried out as part of a 5-year research project funded by the Dutch Ministry of Agriculture, Nature Management, and Food Quality. The project is aimed at improving the basis for the stock assessments of North Sea plaice and sole and to achieve better communication between science, policy and fishermen. We like to thank David Maxwell and Coby Needle for help with the segmented regressions.

References

- Boddeke, R. and Hagel, P. 1995. Eutrophication, fisheries, and productivity of the North Sea continental zone. Condition of the world's aquatic habitats. Proceedings of the world fisheries congress, theme 1. Armantrout, N.B. and Wolotira, R.J. New Delhi, Oxford & IBH Publishing Co. Pvt. Ltd.: 290-315.
- Cushing, D. 1990. Plankton production and year-class strength in fish populations: an update of the match/mismatch hypothesis. *Advances in Marine Biology*, 26: 249-293.
- Fox, C.J., Planque, B.P. and Darby, C. D. 2000. Synchrony in the recruitment time-series of plaice (*Pleuronectes platessa* L.) around the United Kingdom and the influence of sea temperature. *Journal of Sea Research* 44: 159-168.
- Grift, R.E., Rijnsdorp, A.D., Barot, S., Heino, M. and Dieckmann, U. 2003. Fisheries-induced trends in reaction norms for maturation in North Sea plaice. *Marine Ecology Progress Series* 257: 247-257.

- ICES 1997. Report of the Study Group on the Precautionary Approach to Fisheries Management. Feb. 1997. ICES CM 1997/ASSESS:7.
- ICES 1998. Report of the Study Group on the Precautionary Approach to Fisheries Management. Feb. 1998. ICES CM 1998/ASSESS:10.
- ICES 2001. Report of the ICES Advisory Committee on Fishery Management, 2000. *ICES Coop. Res. Rep.* 242.
- ICES 2003. Report of the Study Group on the Further Development of the Precautionary Approach to Fisheries Management. ICES Headquarters, 2-6 December 2002. ICES CM 2003/ACFM:09.
- ICES 2005. Report on the Assessment of Demersal Stocks in the North Sea and Skagerrak. Bergen, Norway, 7-16 September 2004. ICES CM 2005/ACFM:07.
- Van Keeken, O.A., Pastoors, M.A. and Rijnsdorp, A.D. 2004 Reconstructing the numbers of plaice discarded in the demersal fisheries since 1957. Working document for the ICES Working Group on the Assessments of Demersal stocks in the North Sea and Skagerrak.
- Kell, L.T. and Bromley, P.J. 2004. Implications for current management advice for North Sea plaice (*Pleuronectes platessa* L.): Part II. Increased biological realism in recruitment, growth, density-dependent sexual maturation and the impact of sexual dimorphism and fishery discards. *Journal of Sea Research* 51: 301-312.
- Van der Land, M.A. 1991. Distribution of flatfish eggs in the 1989 egg surveys in the southeastern North Sea, and mortality of plaice and sole eggs. *Netherlands Journal of Sea Research*, 27: 277-286.
- O'Brien, C.M. and Maxwell, D.L. 2002a. Towards an operational implementation of the Precautionary Approach within ICES – biomass reference points. WD 8. Report of the Study Group on the Further development of the Precautionary Approach to Fisheries Management. Lisbon, Portugal, 4-8 March 2002. ICES CM 2002/ACFM:10 (Annex 2).
- O'Brien, C.M. and Maxwell, D.L. 2002b. A segmented regression approach to the Precautionary Approach - the case of Northeast Arctic saithe (Sub areas I and II). WD 10. Report of the Study Group on the Further Development of the Precautionary Approach to Fisheries Management. Lisbon, Portugal, 4-8 March 2002. ICES CM 2002/ACFM:10 (Annex 3).
- Reid, P.C., Edwards, M., Hunt, H.G. and Warner, A.J. 1998. Phytoplankton change in the North Atlantic. *Nature*, 391: 546.
- Rijnsdorp, A. D. 1989. Maturation of male and female North Sea plaice (*Pleuronectes platessa* L.). *J. Cons. Int. Exp. Mer*, 46: 35-51.
- Rijnsdorp, A.D. and Leeuwen, P.I. van. 1992. Density-dependent and independent somatic growth in female North Sea plaice, *Pleuronectes platessa* L., between 1930-1985, as revealed by back-calculation of otoliths. *Mar. Ecol. Prog. Ser.* 88: 19-32.

Rijnsdorp, A.D. and van Leeuwen, P.I. 1996. Changes in growth of North Sea plaice since 1950 in relation to density, eutrophication, beam-trawl effort, and temperature. *ICES Journal of Marine Science* 53: 1199-1213.

Rijnsdorp, A.D., van Keeken, O.A. and Bolle, L.J. 2004. Changes in the productivity of the southeastern North Sea as reflected in the growth of plaice and sole. *ICES CM 2004/K:13*.

Rijnsdorp, A.D. and Witthames, P.R. 2005. The ecology of reproduction. In Gibson, R.N. (ed), *The biology and exploitation of flatfish*. Blackwell Scientific Press. Pp. 68-93.

Rowe, S. and Hutchings, J.A. 2003. Mating systems and the conservation of commercially exploited marine fish. *Trends in Ecology and Evolution* 18: 567-572.

Weijerman, M. and Lindeboom, H. 2005. Regime shifts in marine ecosystems in Northwestern Europe. *Marine Ecology Progress Series*, **in press**.

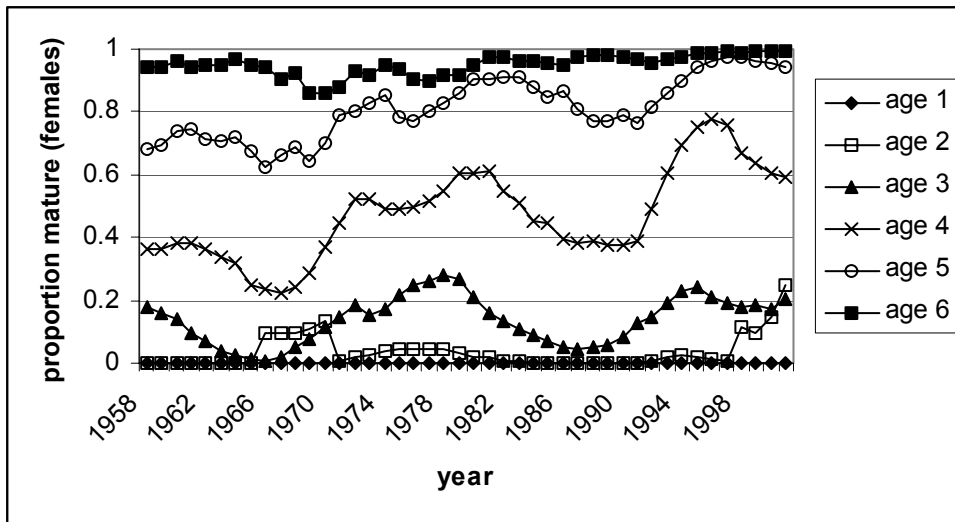


Figure 1. Smoothed female maturity by age versus time. Smoothing was done by taking the 5-year running average.

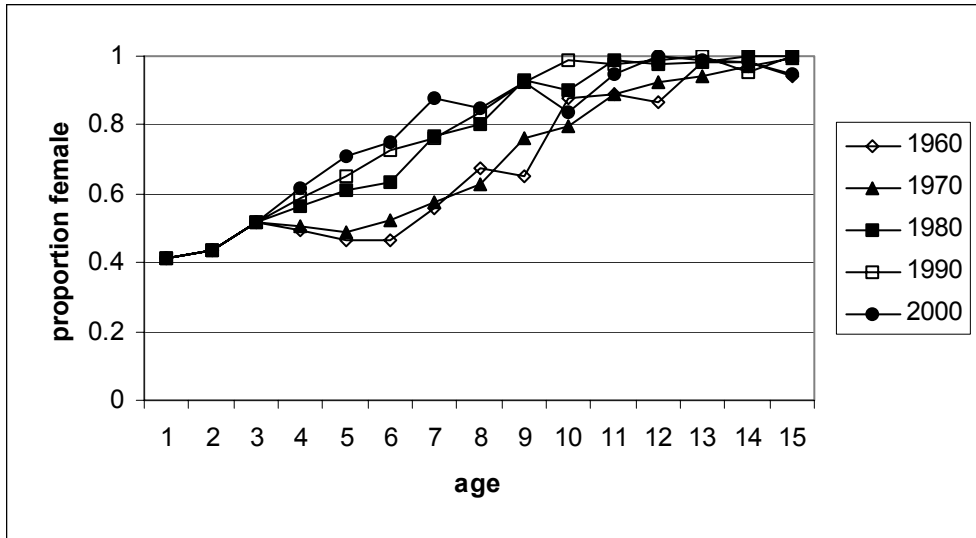


Figure 2. Smoothed proportions female versus age in different years. For the procedure of smoothing see main text.

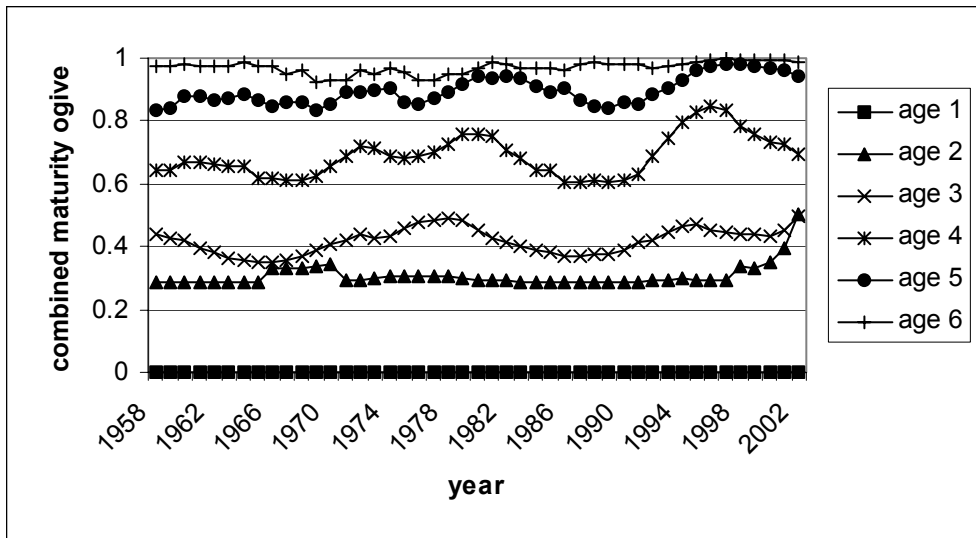


Figure 3. Combined maturity at age versus time. For each year, maturity at age equals female maturity at age times the proportion of females at that age plus male maturity at age times the proportion of males at that age.

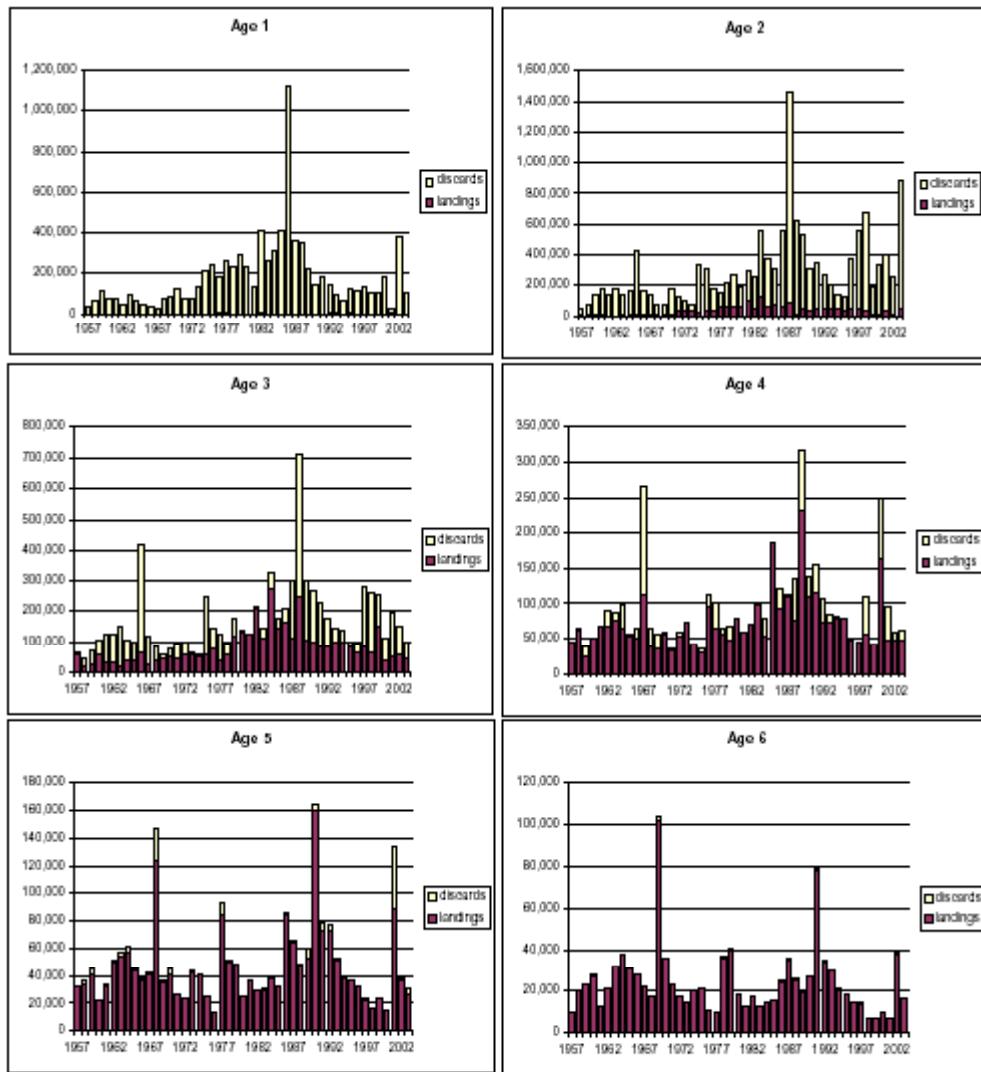


Figure 4. Landed and discarded (observed + reconstructed, method 2) numbers by age.

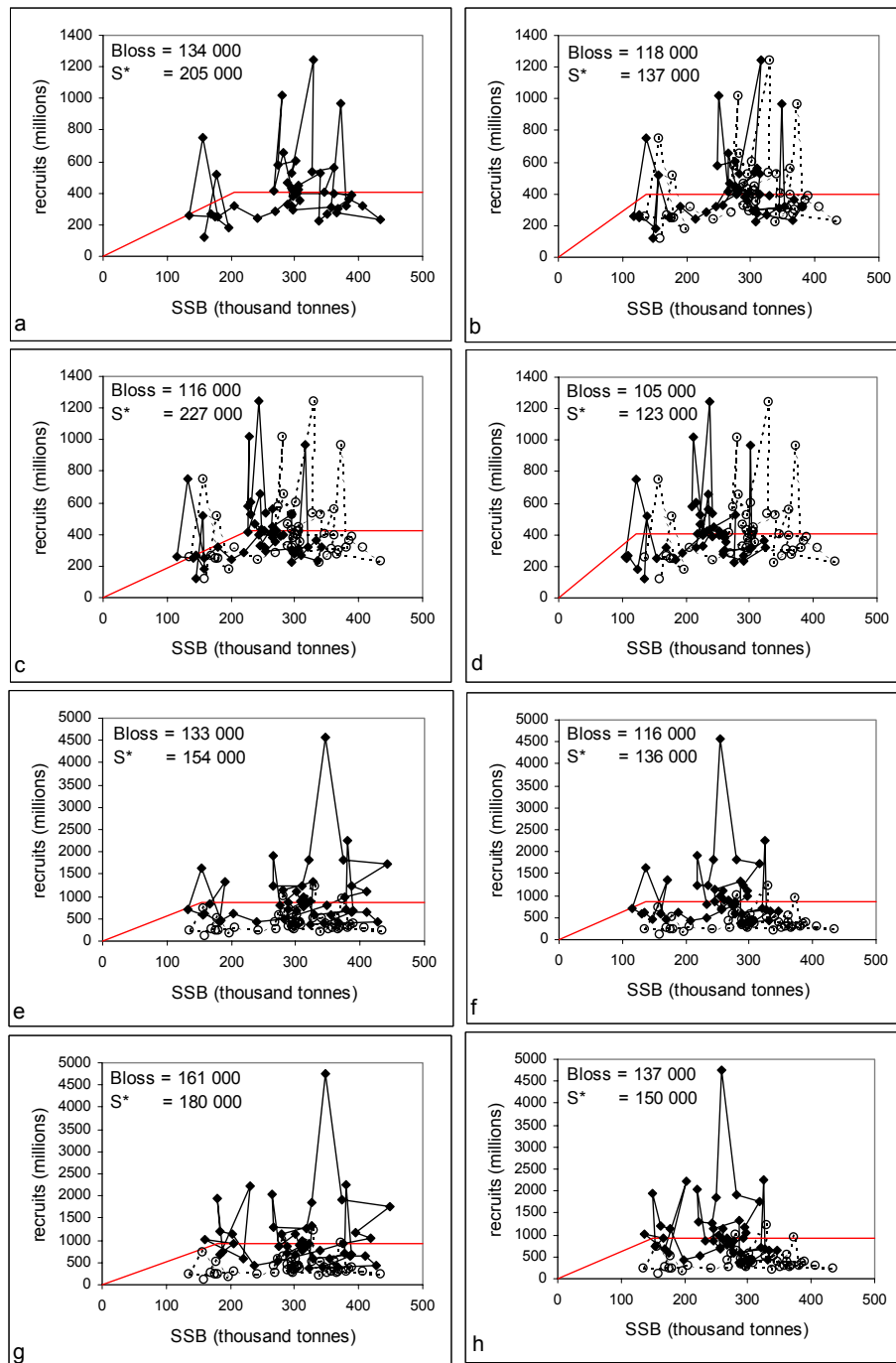


Figure 5. Stock-recruit relation 1957-2002. Solid symbols and solid lines represent (a) update assessment, (b) with biological estimates of stock weights, (c) with biological estimates of maturity, (d) with biological estimates of stock weights and maturity, (e) with reconstructed discards (method 1) and biological estimates of stock weights, (f), with reconstructed discards (method 1) and biological estimates of stock weights and maturity (g) with reconstructed discards (method 2) and biological estimates of stock weights, (h) with reconstructed discards (method 2) and biological estimates of stock weights and maturity. Open symbols and hatched lines in b-h represent the update assessment as in a and are shown for comparison. The lines connect the symbols in chronological order. The red line represents the segmented regression, based on 1957-1999. B_{loss} is the lowest observed SSB and S^* is the SSB at the breakpoint.

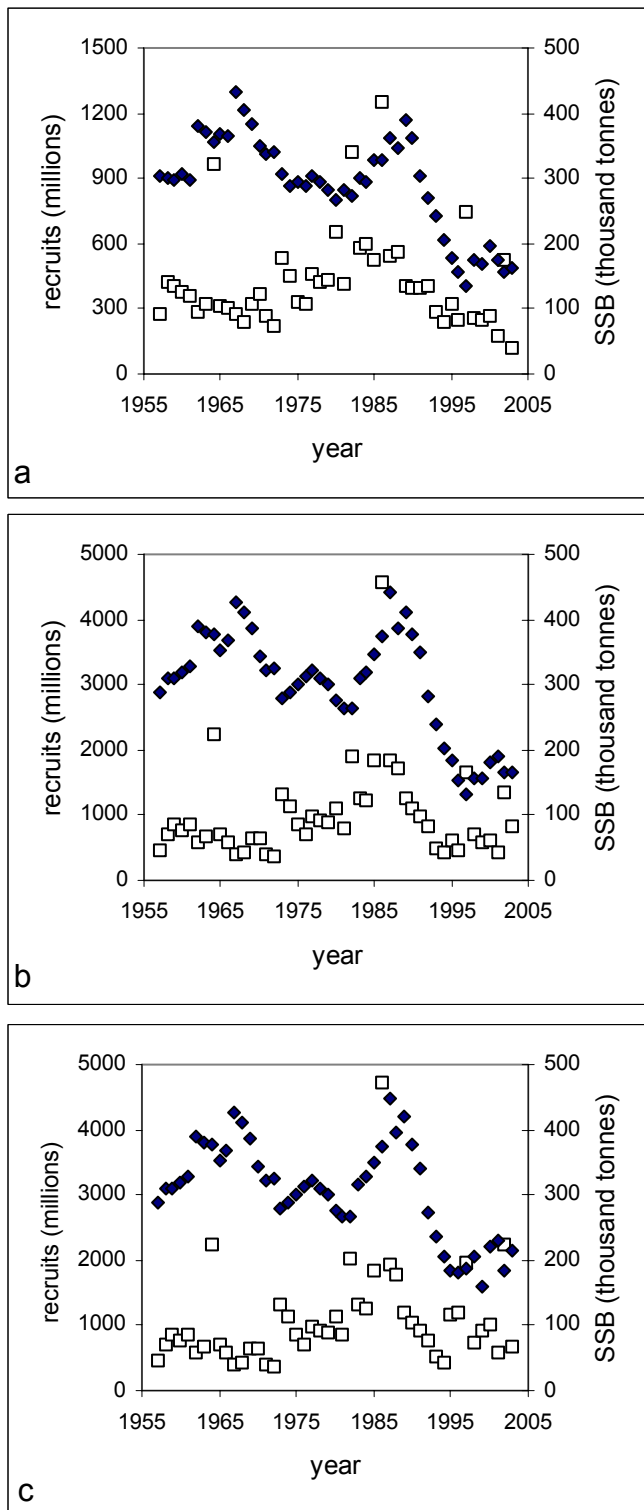


Figure 6. Time series of recruits (open squares) and SSB (closed diamonds); (a) the update assessment, (b) with reconstructed discards (method 1) and biological estimates of stock weights, (c) with reconstructed discards (method 2) and biological estimates of stock weights. Note the different scales on the left y-axis.

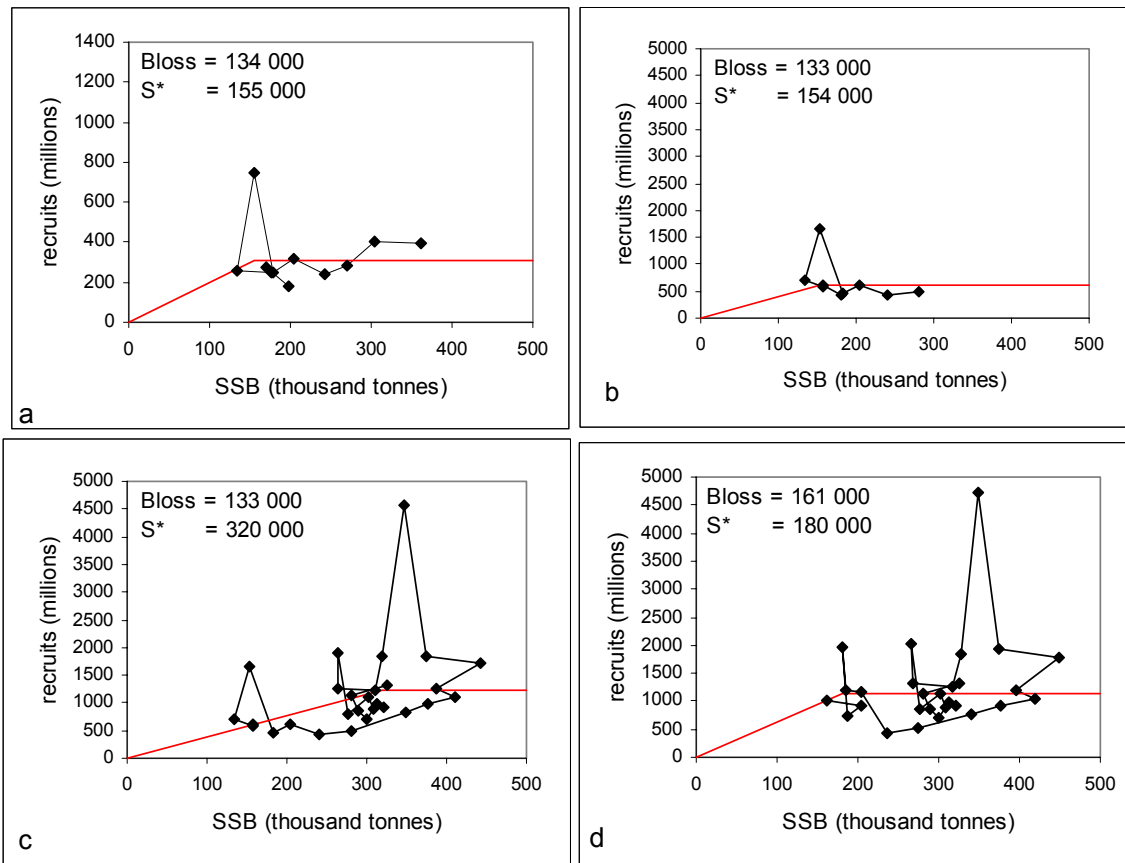


Figure 7. Stock-recruit relation for varying time periods. Black symbols and lines represent (a) update assessment, 1990-2000, (b) with reconstructed discards (method 1) and biological estimates of stock weights, 1992-2000, (c), with reconstructed discards (method 1) and biological estimates of stock weights, 1972-1999, (d) with reconstructed discards (method 2) and biological estimates of stock weights, 1972-1999. The lines connect the symbols in chronological order. The red line represents the segmented regression. B_{loss} is the lowest observed SSB and S^* is the SSB at the breakpoint.