Center for Independent Experts (CIE) Independent Peer Review Report

Review of Technical Documentation for the Vertical Line Analysis Model Supporting an Amendment to the Atlantic Large Whale Take Reduction Plan

November 2012

Executive Summary

This review covers the technical documentation of the vertical line model developed to support the continued development of the Atlantic Large Whale Take Reduction Plan.

The report of the technical documentation provides a clear description of the model and its elaboration. Some clarification of parts of the process is suggested through the use of further flow diagrams.

Widely varying levels of detail were available at the Federal and each of the relevant State levels to describe fishing fleets and their activities. Best use appears to have been made of these disparate data sets.

Data on the distribution of whales are clearly explained, though the data and their interpretation conceal a number of assumptions that should deserve a fuller exploration.

Some elaboration of how data are aggregated and disaggregated at different spatial scales is suggested, along with an exploration of the consequences of assumptions concerning the distribution of fishing effort within each management zone.

Vertical Line scores are generated from a series of model vessel descriptors. State and Federal fishing records are used to describe a series of 'model' vessels, essentially different fleet strata, each of which is ascribed values for the number of trawl-lines, traps or nets, and vertical lines, based mainly either on logbook / interview returns or on expert opinion. These are then given monthly weightings and the effort attributed to them is allocated across 1 minute grid cells in the fishery areas. This procedure in general seems sound, though it is not always easy to follow.

A number of concerns are highlighted relating to the procedure. These include the lack of detailed stratification within the gillnet fleet(s); the possibility of effort by small scale operators having been overlooked; the lack of validation procedures to test the predicted number of lines per model vessel; the seemingly very low numbers of nets used per string by gillnet vessels in some places; the seemingly very high numbers of trap-lines hauled by some lobster trap vessels in some areas; the fact that effort is apparently arbitrarily allocated evenly across all 1-minute grid cells within a reporting area.

Whale distribution data collected over several decades using a variety of platform types and data collection protocols have been collated to provide a crude overview of likely seasonal whale densities.

This assumes there have been no long term trends in whale distribution or density, and assumes that there are no systematic biases within the data. Both assumptions should be checked.

There is no explicit consideration of uncertainty in the whale distribution data analysis or in the vertical line analysis – and the model would benefit from such considerations.

The co-occurrence indicator may not be a perfect analogue for risk of entanglement- but in principle at least it is probably the best that can be done. In detail however, there is a risk that imprecise and extreme values may give unwarranted confidence that specific fishery management measures might have a greater effect on reducing the overall risk of entanglement than is really the case. Some modifications to the indicator are suggested.

Finally a series of specific improvements and developments of the model and its documentation are suggested.

Background

Right whales, humpback whales, and fin whales are listed as endangered species under the Endangered Species Act (ESA). Pursuant to the ESA and the Marine Mammal Protection Act (MMPA), the National Marine Fisheries Service (NMFS) – with guidance from the Atlantic Large Whale Take Reduction Team (ALWTRT) – is responsible for the development and implementation of the Atlantic Large Whale Take Reduction Plan (ALWTRP). The plan seeks to reduce the risks of entanglement through a set of gear modifications and other requirements that affect commercial fishing vessels operating in Atlantic waters.

At the 2003 ALWTRT meeting, the ALWTRT agreed to two overarching principles associated with reducing large whale entanglement risks: reducing entanglement risks associated with groundlines (lines between trap/pots) in commercial trap/pot gear; and reducing entanglement risks associated with vertical lines (endlines or buoy lines) in commercial trap/pot and gillnet gear. NMFS addressed the first principle - reducing entanglement risk from groundlines - in October 2007 with the implementation of a sinking groundline requirement for all trap/pot fisheries throughout the entire East coast (72 FR 57104, October 5, 2007).

NMFS is currently addressing the second principle, reducing entanglement risks associated with vertical lines in commercial trap/pot and gillnet gear.

In 2009, the ALWTRT agreed on a schedule to develop conservation measures for reducing the risk of serious injury and mortality of large whales that become entangled in vertical lines. NMFS committed to publishing a final rule to address vertical line entanglement by 2014. Unlike the broad-scale management approach taken to address entanglement risks associated with groundlines, the approach for the vertical line rulemaking will focus on reducing the risk of vertical line entanglements in finer-scale high impact areas.

Using fishing gear characterization data and whale sightings per unit effort (SPUE) data, NMFS developed a model to determine the co-occurrence of fishing gear density and whale density to serve as a guide in the identification of these high risk areas. The ALWTRT agreed that NMFS should use the model to develop suites of conservation measures that would ultimately serve as options for the ALWTRT to consider when identifying management alternatives. The conservation measures would address vertical line fishery interactions with large whales by reducing the potential for entanglements and minimizing adverse effects if entanglements occur.

The NMFS model has been elaborated by Industrial Economics, incorporated (IEc), and this work is described in the document TECHNICAL DOCUMENTATION FOR THE VERTICAL LINE ANALYSIS MODEL (June 2012).

Description of the Reviewer's role in the review Activities

The reviewer has conducted an independent peer review of the Technical Documentation under the Terms of Reference described in Appendix 2. No meetings or travel were involved, and the reviewer relied mainly upon documents provided by the CIE as described in Appendix 1.

Summary of Findings on each of the Terms of reference:

1. The description of the model's purpose and scope and of the data and methods it employs.

The report provides a clear description of the model's purpose and scope, and a clear exposition of most of the data and methods used. In the main, it is a model of clarity in explaining a very involved and at times tortuous process. The use of numerous footnotes was helpful, and indeed critical to understanding some of the details. Notwithstanding this, the process described is complex and not easy to retain fully without repeated reading. Some clarification, by use of one or more flow diagrams, would aid in the description of the allocation of effective numbers of vertical lines to model vessel categories.

a) Vessel activity is inferred or modelled based on a mixture of officially collected data (Vessel Trip Reports, Landings records, questionnaires and call back surveys) and expert opinions. No new surveys were conducted and no new data were collected. It appears that best use has been made of the very wide variety of data available at the State and Federal levels.

b) Information on the distribution of gear is mainly well explained and clear. However, I was left uncertain about the exact spatial scale being used in some applications. A very similar process of vessel stratification by amount of gear used, number of vertical lines used and monthly activity was used in each State's waters to model the distribution of gear. Further comments are given under ToRs 2 & 3 below.

c) The information on whale distribution was clearly explained, as far as it went. Assuming one takes the combined sightings of the NARWC at face value, the data and methods are clearly explained. I have some reservations about the underlying data which are elaborated under ToR 3 further below.

2. Characterisation of Fishing Activity in Federal and State Waters.

a) Estimates of the <u>number</u> of boats in Federal and State waters in each fishery are not fully addressed, because the metric of interest is the amount of fishing effort. Fishing permits, Vessel Trip Reports and Logbooks provide estimates of effort on a vessel by vessel basis at the Federal level, so it must be assumed that the model could identify the number of vessels involved in each fishery, and the method seems appropriate, at least at the Federal level. No estimates are provided of the numbers of boats in each fishery (footnotes excepted), or of how many operate in more than one fishery, but these data are not strictly needed to address vertical line density. At the State level too, the number of vessels is not explicitly estimated, but rather the use of logbook data, surveys and 'best professional judgement' lead to estimates of numbers of vessels active by month in each fishery and the number of strings each fishes. Again – the method generally seems appropriate, given the apparent limitations of the data.

I am slightly concerned that the number of vessels estimated to be fishing in any month in federal waters – where vessels licenced to fish for lobster do not need to supply VTRs – may be biased, because the model assumes that those vessel that are licenced to fish for lobster as well as for other resources will behave in the same way as those that are licenced to fish for lobster alone. Footnote 13 indicates that there are large numbers of vessels that do not provide VTR data, and acknowledges this may be a problem for the model. This would be worth exploring some more with some sensitivity analysis based on different assumptions about how those vessels may behave.

In State waters, however, I am more concerned that there may be a hidden vessel stratum that falls outside the remit of State fishery managers – specifically very small scale operators. We learn in the model description that "the [Maine Department of Marine Resources] provided an analysis of its "100 percent dealer reporting" (not explained) "data that shows the percentage of vessels ... that were active in each month", and "where active vessels are defined as those that landed <u>at least 100 pounds of lobster</u>" (per month) (p 27). I do not know the catch rates of lobster per trap in Maine, but it is feasible that there could be a large number of vessels that fish yet fall below the 100 pounds per month limit, and are therefore not considered active. In New Hampshire we learn (p 87) that "the New Hampshire Fish and Game Department (FGD) requires that fishermen who land up to 1,000 pounds [...use] the Annual Lobster Harvester Report". This indicates a more than negligible level of effort in this category of vessel (sub 1000 lbs per year) in New Hampshire, yet similar vessels seem to have been ignored in Maine. In other states (like Massachusetts) it is left unclear whether there is any substantial fishing effort by vessels that fall below some reporting threshold level.

b) In terms of effort distribution, I am unclear about how the various spatial scales have been used and morphed from one to another. The documentation refers to grid cells of 1-minute. The statistical reporting areas vary widely in shape and size from State to State and within Federal waters too. I assume that effort within each reporting zone is allocated evenly to each 1 minute grid cell within the zone. Yet in the Vertical Line Model Charts provided under background documents, the effort appears to be shown by ten-minute grid cells – it seems these are simply amalgamations of the 100 1-minute cells. It is not clear how reported vessel fishing effort within each state is allocated between state areas and thence to the 1-minute cells. Do all vessels only fish in one reporting area? A flow diagram showing how effort has been allocated at each stage of the process would be helpful.

Throughout the model description it seems that where geographical information is lacking, an assumption of even-effort distribution is made. In Federal waters, VTR data provide an average location of fishing for a whole trip, and this is allocated to a grid cell to represent effort by that trip. It is not clearly stated (p 18) what size these grid cells are, but I assume these are the 1-

minute cells previously referred to. For licenced fishing vessels that do not supply VTR data, "the model assumed that the activity ... is distributed evenly across the LMA, and apportions activity to each grid cell with in the LMA accordingly". It seems unlikely that these vessels (possibly smaller ones- but they are not characterised in the report) are likely to fish evenly across the LMA (see also comments under 3e below). I also wonder how well the VTR data represent actual fishing effort by 1 minute grid cell. I doubt that the an average fishing location for a trip will be adequate to represent fishing effort distribution at the 1 minute cell level, though it may work at the LMA level; some further exploration and discussion of this, and the potential violation of the implicit assumptions here, would help.

In State waters- the same assumptions about even distribution of effort are made – for example on page 27, in relation to Maine, we learn that "The model assumes that the activity of these vessels is distributed evenly throughout the state-waters portion of each lobster zone." Some investigation into how sensitive the model results are to this assumption would be useful.

As it stands, it would seem that any proposed management measures that involve spatial effort limitations should only be postulated at the level of the 'lobster zone' rather than at any higher level of spatial detail such as at 1 or 10 minutes of latitude and longitude, because the effort data going into the model are mostly based at the 'lobster zone' level; but this may be an unduly conservative view. Further assessment and discussion of the most appropriate spatial scales would therefore be useful (see also further comments below under 3e and 3g).

c) Monthly variations in fishing effort are generally well thought through in State waters, but less so in federal waters, at least in relation to the amounts of gear used. On page 20 we learn that "Based on discussions with NMFS gear specialists, the model's default values assume no seasonal adjustments for the number of traps or strings fished in Federal waters". It seems unlikely to me that there is no seasonal fluctuation in gillnet or lobster trap numbers fished per boat in Federal waters. I would guess that poor weather alone should ensure that seasonal fluctuations in gear usage are likely, but this could easily be checked by talking to a few people involved in the fishery.

d) Key data limitations and uncertainties have in general been highlighted, though there are some areas which could have been more fully explored, as has been suggested above. Specifically these are: the fact that a large number of vessels are able to fish for lobsters without supplying VTR data (mentioned but could be more fully explored); the possibility that substantial effort may be hidden by boats that land less than 100 lbs. of lobster in Maine at least (similar issues may extend elsewhere); the even distribution of effort throughout management areas where there is a lack of detailed data on actual effort distribution (needs further exploration of what the consequences of this series of assumptions might mean); how the use of average latitude and longitude in the VTR data to represent effort for each trip may or may not influence the model's overall predictions of effort distribution; assumption of constant gear characteristics in all months in Federal waters.

e) Within the limits of the data described here, I think that tabulating the numbers of vessels active in each fishery stratum would be a helpful way to clarify and provide an overview of static gear fishing in the wider region. Some consideration as to how much inter-annual variability in fishing vessel activity would also be helpful (throughout it appears to be an assumption that the most recent years data will be the best descriptor of subsequent years: this is probably true but some idea of the amount of change over the past few years might help put this in context). Some consideration of the effects of violations of some of the key assumptions mentioned above would also help improve the model's characterisation of the fleet.

   Beyond that, some ground-truthing of certain aspects of the models predictions would also be helpful, perhaps drawing on additional existing data. As mentioned under 3j below – a comparison of the model's predictions of fishing effort distribution within each management area based on other data sources, including fishery observer records of where people actually fish, and sightings of fishing activity recorded by the NARWC (and other? Coastguard?) aerial and shipboard surveys – would help confirm the model's validity.

3. Characterization of gear use in the fisheries of interest

a) The use of model vessels, or of fleet stratification by vessel characteristics, is essentially sound. For the majority of the overall US Atlantic fleet it has been possible to obtain mean values for the number of vertical lines used, based on the mean number of traps deployed per boat and the number of traps per trawl and a series of geographically varying assumptions concerning the number of vertical lines per trawl. I am slightly concerned that the stratification of the gillnet vessel fleet takes no account of target species (see c below).

b) The *parameters* employed to describe the amounts of vertical (and horizontal) lines employed are sound and are sufficient for the purposes of the model. The *values* ascribed to those parameters are not all equally clearly elaborated. In some States it appears that vessels report the number of traps used and the number of end lines used. The model requires the data to be broken down by the number of traps by vessel, and the number of traps per trawl-line. To achieve this end assumptions are made about the likely number of trawl-lines used and the number of traps per trawl-line. Specifically in each State an assumption is made about the number of end lines used on trawl-lines with different numbers of traps: typically any trawl lines with two or three traps or fewer are ascribed one end-line, while those with more than 2 or 3 are ascribed 2 end lines. In this way, the original data, which appear to include the very parameter that is needed – i.e. the number of endlines being used, are manipulated via a series of assumptions to generate ... the number of endlines being used. While I can see that the model is attempting to provide a more detailed picture of the fishery – including the amount of groundline being used, it might have been useful to try to use the original data, or at least to formally compare the results of the model 'predictions' with the raw data.

I do not suppose that the assumptions that are made about the number of endlines per trawl-line will have a very great impact on the overall vertical line scores, nor on the co-occurrence index, but it would have been nice to have seen some sensitivity testing of the predicted number of endlines dependent on the assumptions made.

I note that the model does not consider the use of gangions or snoods on trap lines; I assume that these are not used anywhere in the region under consideration.

c) I am not so confident about the assumptions underlying the gillnet part of the model. In my experience there is a great deal of variation in net string lengths deployed by vessels that is often driven by the target species catch rates in particular fisheries. Where a target species is taken very sporadically, long nets are the norm, while for schooling target species, shorter net strings are used. I was therefore a little concerned to find that there is little discussion as to the length or the number of gillnet strings used by vessel category. On page 25 we learn that gillnet vessels can be characterised by a "total number of strings typically fished", and on page 24 we find that gillnet gear configurations in Federal waters assume 33 nets fished in 3 strings or 25 nets fished in 3.6 strings are the norms for the model vessels in three Federal fishing zones. I have no detailed knowledge of these fisheries, but it seems likely that the "typical number " of strings will be quite different between fisheries targeting different species of fish, and that a vessel fishing outside State waters (presumably the larger vessels) would be unlikely to fish just three strings of nets. I may be wrong but this strikes me as very limited fishing effort. We also learn on page 89 that "the model assumes that gillnetters fish one net per string" in New Hampshire. Again, I have no expertise on the fishing techniques of New Hampshire gillnetters, but I find this highly implausible. Gillnetters typically rig net strings with between 3 and 60 or more net panels. In federal waters we learn on page 25 that the typical values are 11 and 7 net panels per string. I wonder if there is some confusion in terminology here between net panels, nets and net strings? On the other hand, single net-panel strings continue to be mentioned in several other States further south, so perhaps this is a correct interpretation.

The equations used to calculate the number of vertical lines – and the length of the groundlines – are simple and conceptually correct for the pot lines, except that gangions are not included (as mentioned previously and presumably because they are not in use). However, for the gillnet calculations, exhibit 6 on page 12 states that: Length of Groundline = Total strings fished x Anchors per string* Length of anchor lines. Footnote 5 on page 11 states that: "For use in potential revisions to the model, IEc also collected information on the number of net panels per string, the height and length of the net panels, and the length of the line between the net panels. Currently, these values are not used in the calculations described above." It is quite unclear to me why the length of the anchor lines is deemed important while the length of the leadline or floatrope of a gillnet string is not included. Furthermore, length of trap and pot groundlines are clearly important in the model even if not the primary subject of interest at present. This seems anomalous.

d) The methods and assumptions used to define model vessels in the Federal lobster fishery are adequate given the apparent lack of data. The fact that there are no VTR requirements for lobster vessels fishing in Federal waters severely restricts the authors' ability to allocate model vessels. I am surprised that there appears to be no other data source that can be relied upon to help determine model lobster vessels in federal waters. I had understood for example that a weigh-out system enabled landings to be tracked by vessels, which would at least give an indication of the scale of operation by vessel, as well as seasonal trends. Compared with the detail available in most State waters, the simple allocation of all vessels into one of three areas, each with its own 'typical' gear configuration, seems laconic. Furthermore, despite the presentation of methods to allocate vessel effort on a seasonal basis, for federal waters we find that "Based on discussions with NMFS gear specialists, the model's default values assume no seasonal adjustments for the number of traps or strings fished in Federal waters". I am also concerned that "In the absence of more detailed information on the location of fishing activity, the model assumes that the activity of these vessels is distributed evenly across the LMA, and apportions activity to each grid cell within the LMA accordingly". This appears to mean that expected effort levels in each 10 minute grid cell are identical across each LMA for the 'non-VTR' fleet. It seems to me very unlikely that effort in the outer limits of LMA 3 would be expected to be identical to those of Georges Bank for example. Given that "Information on the location of trips taken by vessels that hold Federal lobster permits is limited to those that also hold permits for other fisheries that impose VTR requirements; these vessels must report all fishing activity to NERO" (p18-19) – might it not have been possible to weight the distribution of effort for the non-VTR fleet by the data supplied by those vessels that hold permits for other fisheries, specifically, determining which parts of each federal LMA are most used for lobster fishing?

e) More or less the same comments apply to the blue crab and other pot fisheries. For blue crab fisheries there are limited federal data available, so model averages from nearby state waters are used. One must wonder whether vessels fishing inside State waters truly reflect those fishing further offshore. For the other pot and trap fisheries, expert opinion was used, which is hard to argue with in the absence of any more substantial data. Some telephone calls to a selection of relevant fishermen might have helped substantiate these assumptions.

f) Comments regarding the specification of model federal gillnet vessels have already been made above in c). In terms of gear usage- it seems that some very simplistic assumption may have been made about both seasonality (based on NMFS expert opinion) and on the model vessel types' use of gear – where there does not appear to be any stratification by fishery / target species - yet where model specification was based on observer records where thousands of detailed records should enable a much finer scale stratification. Furthermore, the modelling of spatial distribution of effort is very crude, and based on VTR / logbook returns: "Specifically, fishermen [in NE federal waters] provide longitude and latitude coordinates that represent their average location for each fishing trip. The Southeast Logbook, which covers Federal waters south of Cape Hatteras, North Carolina, similarly requires trip level reporting; however, fishermen are required to identify the location of their fishing effort on a 1-degree grid, as opposed to a specific location". From this, effort is apparently estimated by 1 minute grid cell. It might have been more sensible to model the distribution of fishing effort based on the observer logbook records to see which areas are most heavily fished. Or are there coastguard data that might be used to map effort distribution more finely? It is not possible to say how much of a difference it might have made to the end product, but some kind of sensitivity analysis on how the geographical misallocation of effort through generalised allocation across wide areas (for example allocating an even spread of effort among 360 one minute grid cells from a single 1 degree effort record) might affect the predictions of the co-occurrence indicator would help. There does appear to be a mismatch of the level of detail collected and analysed among some of the State fisheries (specifically the State lobster trap fisheries) and fisheries in Federal waters where more broad assumptions have been made.

g) The methods and assumptions used to define model vessels in the State lobster fisheries appear adequate given the available data. Although it impossible for this reviewer to be certain that all avenues of available data within each individual State's fishery management division have been sourced, the data sources are all laid out clearly. There is a wide variety of routine data collection methods which are then all 'shoe-horned' into a similar process for generating model vessel categories based principally on the numbers of traps being used, and the numbers of traps per trawl. I do wonder, however, how *reliable* the vessel allocation methods have been and it would have been useful to have some discussion on this point. For example, in Maine we find a vessel category with two vessels fishing in Federal waters in January using over 600 traps (600+ category) with 2 traps per trawl (p35). This implies over 300 trawl lines / vertical lines being fished by a single vessel; in January when weather is worst. This seems like a lot to me, and it could have been helpful if there had been some validation of the model categories – especially the extreme ones like this – for example by asking State experts or others working in the fishery – if such extreme values are reasonable. However, in general I think that the best use of the available data seems to have been made in allocating effort to model vessel categories.

h) The authors of the model documentation have done a good job in mentioning many or most of the data limitations and uncertainties, including some that are in the footnotes and that may not have been obvious to the naïve reader. However, I would have liked to have seen a little more focus on the key assumptions – which lead to those uncertainties – tabulated and clearly available, together with some qualitative assessment at least of how important each assumption might be in determining the final number of vertical line score per grid cell. I have mentioned several of these assumptions above (e.g., lumping gillnet vessels into so few strata / model vessels; assumptions about the number of end lines relative to the number of traps per trawl-line; spatial reallocation of fishing effort to 1 minute grid cells ...).

i) The most obvious way in which IEc could improve the model's characterization of gear use would be to test the validity of the assumptions and characterisations thus far. I have mentioned the possibility of getting local experts to examine the matrices of model vessel categories to try to weed out any that are obviously erroneous. The predictions about vertical line distribution should also be validated against observations of fishing gear distribution that are routinely collected for example under the NARWC aerial and shipboard sightings surveys. Another method would be to compare on board observer data with model predictions. I assume Federal observers from the NEFOP would make observations inside State waters, or that other observer schemes may have operated there. These might be used to help validate model vessel categorisation and effort distribution.

4. The seasonal distribution of endangered species of large whales in waters subject to the ALWTRP

a) The whale sightings data that are used to characterise the variation in monthly distribution are practically speaking the only data that are available, so they are the most appropriate data to use. This is not to say they have been used in the most appropriate way.

b) I do not think that key data limitations and uncertainties have been adequately explored in relation to the whale sightings data. Firstly, there is an implicit assumption here that whales have a standard and predictable migration pattern, and that by pooling data over as many years as possible, that the best picture of distribution will be obtained. I note that in at least one of the elaborations of the model that are provided in the Background documents whale distribution maps are based only on a few recent years of sightings data, so this concern is clearly felt by others. I also note that on examination of the NARWC maps available elsewhere, there have been some major changes in distribution (shifts in and out of Cape Cod Bay for example) even within the last few years. Secondly, I am concerned that a mix of different platforms and protocols has been used to summarise sightings densities. Aerial platforms for example have operated at very different heights within the overall NARWC data series, and are combined with shipboard sightings. Some consideration of the potential for bias in the resulting compiled dataset would be appropriate. For example, are early years' data – or data collected from the outer shelf region, systematically less likely to detect whales for a km of trackline effort? Thirdly, the maps presented give no indication as to how likely any zero sightings rate in a cell may represent an actual absence of animals. Much depends on the amount of effort allocated to each cell. Likewise, high sightings rates can be derived from a single observation in a cell with very little effort. Some way of characterising precision of the assumed SPUE value would be helpful.

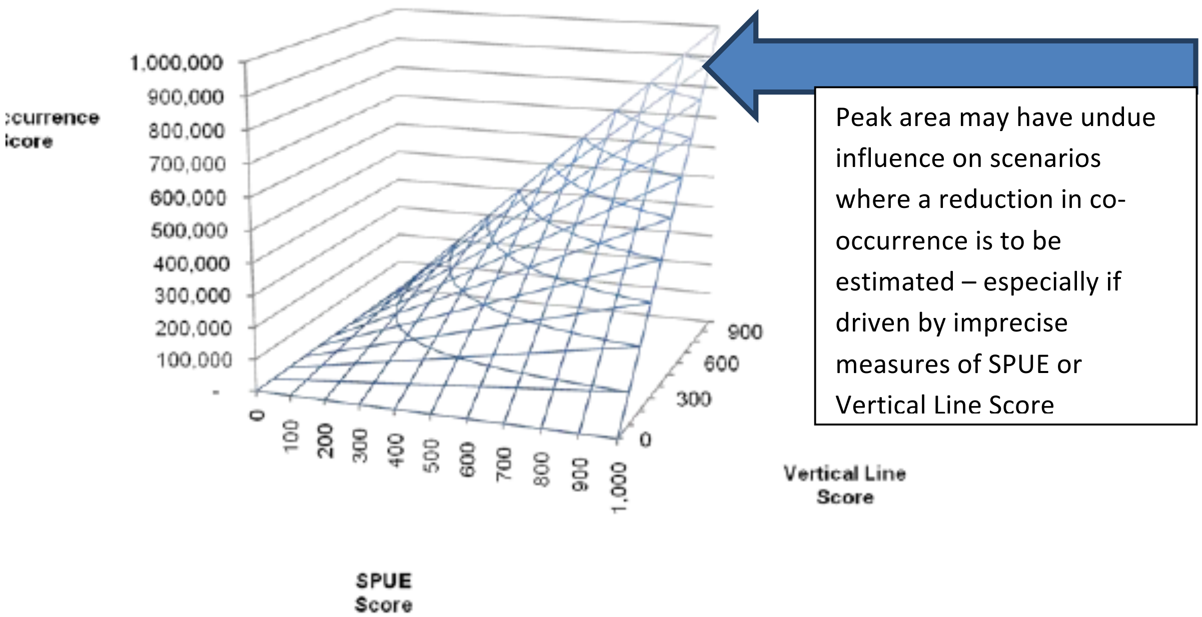
c) There are several ways in which characterisation of seasonal distribution could be improved. Firstly – some thought should be given to how much search effort is enough to reasonably assert that whales are most likely absent from an area. A single 10km survey trackline through an area in one year is unlikely to be sufficient. As stated above, some measure of precision of the SPUE values would be appropriate, as there may be serious consequences for assuming zero presence or a high density in an area that has been subject only to limited survey effort (see 5b below too). It may be appropriate to exclude cells with less than some other cut-off effort value (other than the arbitrary 10km that is used presently), but this would need to be justified. Secondly, I am not convinced that mixing aerial and shipboard sightings is without bias. Some analysis of the sightings data would be appropriate to ensure that any potential bias is clearly understood. Thirdly, some analysis of the inter-annual stability of sightings rates is needed. In some areas and at some spatial scales, inter-annual seasonal differences may not be significant, whereas in other areas, and perhaps at smaller spatial frames, such as within Cape Cod Bay, there are clearly major differences in seasonal density from year to year. I gather that attempts are underway to try to model whale distribution in relation to prey availability. A spatial modeling framework may provide a more useful means of examining whale distribution than a simple series of observations. At the very least, the variability of distribution needs to be considered during exploration of potential management closures or restrictions to fishing.

5. The model's primary outputs:

a) The data, methods and assumptions to develop the measures of fishing effort distribution and whale distribution have been addressed above. Given the various limitations and concerns expressed previously, the data, methods and assumptions in generating a co-occurrence index are appropriate, assuming that the whale and fishery distribution are indeed adequately characterised.

b) Whether or not the co-occurrence indicator provides a reasonable basis for evaluating relative differences in the likelihood of whale entanglement is another matter. There are two areas of concern that I have. The first is that the indicator itself does not necessarily relate to the probability of entanglement. To the extent that whale behaviour, or indeed end line behaviour and specific density at the appropriate 'whale scale', may influence the probability of entanglement, the two metrics being compared may not adequately characterise risk of entanglement. However, this issue is explicitly acknowledged in the model description and there seems little that can be done at present to address this issue. In other words, a simple measure or co-occurrence is probably the best we can do at present.

   The second concern, however, relates to the scaling of the co-occurrence index and here it may be possible to improve the metric being proposed or used. I note that there are some extreme values in SPUE and vertical line density and hence in the co-occurrence indicator. For example, most positive co-occurrence scores on a monthly basis lie in the hundreds, yet one value in the Gulf of Maine in July reaches 44,747 in one of the background documents. Without access to the original data it is difficult to be sure, but it seems that the total co-occurrence index is somewhere around 100,000-200,000 for this area and this month. Simply eliminating the single high value cell would then have a massive impact on the overall co-occurrence indicator value. This is a concern because a single high value could easily have been achieved for this one cell through a very small amount of observation effort and the chance encounter of a small number of whales there. As this is such an obvious concern, I feel sure that this issue has been addressed somewhere - perhaps in the workshops for which presentations were provided – but could not find any text relating to it. One way to deal with this issue would be to apply a square root or log transform to the data, thereby effectively flattening off the right hand end of the distribution shown below, taken from the presentation on co-occurrence document presented at the April 2011 Mid-Atlantic/Southeast Sub-Group meeting.



c) I found little discussion of the effects of data limitations and uncertainty on the final metric being calculated here – i.e. the co-occurrence indicator. The uncertainties lie mainly in the estimation of vertical line density and in whale SPUE, and aside from the two issues raised above (under 5b), the issue of data limitations are not strictly relevant to the calculation of the co-occurrence index. In other word, if the SPUE and fishing effort maps were known to be correct and without error, then the metric as described would be correctly calculated. The issue remains whether this is the best way to characterise risk of entanglement, and I would agree that at present it is (subject to possible rescaling as described in 5b above).

d) Within the limits of available data, any potential improvements to the final product – the co-occurrence indicator, are limited. I have suggested that rescaling the indicator might help eliminate the possible over-dominant effect of extreme and potentially unreliable values. The introduction of some measure of uncertainty to the metric – based on measures of the underlying uncertainty for the SPUE and fishing effort density values - might also help, but would require substantial more analytical work, for possibly limited management gain. I think that the greatest improvement that could be made would be through some external validation of the model results. This could be done in two ways – first determine how reliably the vertical line density and SPUE values predict seasonal distribution from year to year. This might be done using existing data and re-running the analysis using different years of data (I note that the fishing effort data were often available for several years – albeit within a limited range). Second- and much more speculatively- would be to seek out available data on whale entanglement events, and try to determine to what extent the model predictions in terms of co-occurrence indicators matched the known distribution of entanglement events. Given the fact that whales can move long distances with lines attached to them, this might only be possible at a very coarse scale, but such an exercise might help bring some confidence to the model's predictions.

6. Overall, what steps should IEc take to improve the model and/or its documentation?

IEc could and probably should determine whether the known locations of entanglements are clustered in any way that is similar to the predicted areas of highest distribution. If entanglement probability is not directly related to sighting per unit effort rate, then this would undermine the utility of the tool. It may be for example that entanglement is more likely in some months or areas due to food availability if entanglement is usually a consequence of feeding, independent of the co-occurrence indicator. Some modification of the indicator may then be appropriate. Some modification of the Co-occurrence Indicator might be justified to help minimize undue influence of outlier or extreme values.

In terms of whale sightings per unit effort, it would be appropriate to see whether or not some other model-based distribution of whale density could be derived, and especially if some measure of uncertainty could be introduced so that the final product is not solely dependent upon deterministic predictions of whale density based on an heterogeneous sightings data set spanning several decades. At the least some examination of potential bias, and of likely precision is required if the NARWC data are to be used in their present form.

In terms of the Vertical Line Score calculations, IEc could:

- Consider whether further stratification of the fleet fishing in federal waters might be justified (lobster trap and gillnet at least).
- Consider whether more detailed description of effort distribution in federal waters is possible or needed.
- Consider improvement in vessel stratification in the gillnet fleet in some States – if it is worth it.
- Consider non-reporting vessels – especially in Maine – with low total landings.

Overall, the report would benefit from additional flow diagrams to demonstrate:

- how effort (vertical lines) is allocated by vessel strata
- how the model allocates effort spatially.

Regardless of diagrams, better elaboration of how the model aggregates and disaggregates data at different spatial scales is needed.

Key assumptions should be itemized in a table.

Testing of assumptions should be made more explicit, possibly by finding new ways to test them, for example by telephone surveys of key fishers.

Model results need to be validated wherever possible (within available resources of course), and some suggestions have been made above.

Some sensitivity tests should be made to assess how important the assumptions (perhaps just the most significant ones) are on the end result.

7. Conclusions and recommendations

Overall, I found this a very comprehensive set of documentation of a model that has clearly been thought through in considerable detail. Given the disparate nature of the available fishery data, IEc has done a commendable job in providing a framework for elaborating different conservations scenarios. The model as it stands should represent a useful tool to explore the impacts of potential management measures. To improve transparency and confidence in any resulting predictions in terms of reduced likelihood of whale entanglements, a series of recommended improvements and subjects for further consideration has been made. Specific and detailed recommendations for further work have been listed under Section 6 above while more general recommendations are summarised below:

- Tabulation or clarification of the key assumptions at each stage
- Inclusion of more detailed diagrams to demonstrate processes involved in estimating, inter alia, number of vertical lines for each model vessel and the ways in which effort is allocated geographically.
- Some consideration of interannual variability in fishing effort and in whale sightings rates and the possible consequences to the model predictions.
- Sensitivity analyses to violations in some of the key assumptions
- Ways to validate model predictions, at several levels.

Appendix 1: Bibliography of materials provided for review

Technical Documentation for the Vertical Line Analysis Model Supporting an Amendment to the Atlantic Large Whale Take Reduction Plan:

IEc – Technical Documentation for the Vertical Line Analysis Model. June 2012. 194 pages.

Background material:

(1) Atlantic Large Whale Take Reduction Plan Website
(2) Developing conservation measures intended to reduce the risk of serious injury and mortality of large whales due to entanglement in vertical lines: Atlantic Large Whale Take Reduction Plan Scoping Document
    a. National Environmental Policy Act (NEPA) Informational Sheet
    b. Vertical Line model charts
        i. Northeast Region
        ii. Mid-Atlantic Region
        iii. Southeast Region
    c. Vertical Line Risk Reduction Proposal Criteria
(3) Nov 2010 (Northeast Subgroup) Presentations – Provide good rundown on methods
    a. Vertical Line
    b. SPUE
    c. Co-Occurrence
(4) March 9, 2011 Northeast Working Group presentations
    a. North Atlantic Right Whale Consortium SPUE
    b. Survey effort
    c. Co-Occurrence
(5) April 2011 Mid-Atlantic/Southeast Sub-Group meeting presentations-- Redundant with the above re: methods, but shows results for the Mid-Atlantic and Southeast.
    a. Vertical Line
    b. SPUE
    c. Co-Occurrence
(6) April 2012 webinar presentation—Shows latest use of the model and analysis of proposals
    a. Analysis of Impact on Alternate Management Measures on VL and Co-occurrence Scores

Appendix 2: A copy of the CIE Statement of Work

Attachment A: Statement of Work for Dr. Simon Northridge

External Independent Peer Review by the Center for Independent Experts

Review of Technical Documentation for the Vertical Line Analysis Model Supporting an Amendment to the Atlantic Large Whale Take Reduction Plan

Scope of Work and CIE Process: The National Marine Fisheries Service's (NMFS) Office of Science and Technology coordinates and manages a contract providing external expertise through the Center for Independent Experts (CIE) to conduct independent peer reviews of NMFS scientific projects. The Statement of Work (SoW) described herein was established by the NMFS Project Contact and Contracting Officer's Representative (COR), and reviewed by CIE for compliance with their policy for providing independent expertise that can provide impartial and independent peer review without conflicts of interest. CIE reviewers are selected by the CIE Steering Committee and CIE Coordination Team to conduct the independent peer review of NMFS science in compliance the predetermined Terms of Reference (ToRs) of the peer review. Each CIE reviewer is contracted to deliver an independent peer review report to be approved by the CIE Steering Committee and the report is to be formatted with content requirements as specified in **Annex 1**. This SoW describes the work tasks and deliverables of the CIE reviewer for conducting an independent peer review of the following NMFS project. Further information on the CIE process can be obtained from www.ciereviews.org.

Project Description: NOAA's National Marine Fisheries Service (NMFS) intends to expand large whale conservation efforts by amending regulations that implement the Atlantic Large Whale Take Reduction Plan (ALWTRP). Since its implementation in 1997, the ALWTRP was modified on several occasions to reduce the risk of injury and mortality of large whales that interact with commercial trap/pot and gillnet fishing gear. The ALWTRP consists of regulatory and non-regulatory programs including: broad-based gear modifications, time-area closures, disentanglement, research and outreach. Despite these efforts, there continues to be injuries and mortalities of large whales from entanglements in vertical lines from commercial trap/pot and gillnet fishing gear. Therefore, additional modifications to the ALWTRP are needed.

At the 2003 Atlantic Large Whale Take Reduction Team (ALWTRT) meeting, the ALWTRT agreed to two overarching principles associated with reducing large whale entanglement risks: reducing entanglement risks associated with groundlines (lines between trap/pots) in commercial trap/pot gear; and reducing entanglement risks associated with vertical lines (endlines or buoy lines) in commercial trap/pot and gillnet gear. NMFS addressed the first principle; reducing entanglement risk from groundlines in October 2007 with the implementation of a sinking groundline requirement for all trap/pot fisheries throughout the entire East coast (72 FR 57104, October 5, 2007). NMFS is addressing the second principle, reducing entanglement risks associated with vertical lines in commercial trap/pot and gillnet gear, in this current process.

In 2009, the ALWTRT agreed on a schedule to develop conservation measures for reducing the risk of serious injury and mortality of large whales that become entangled in vertical lines. NMFS committed to publishing a final rule to address vertical line entanglement by 2014. Unlike the broad-scale management approach taken to address entanglement risks associated with groundlines, the approach for the vertical line rulemaking will focus on reducing the risk of vertical line entanglements in finer-scale high impact areas. Using fishing gear characterization data and whale sightings per unit effort (SPUE) data, NMFS developed a model to determine the co-occurrence of fishing gear density and whale density to serve as a guide in the identification of these high risk areas. The ALWTRT agreed that NMFS should use the model to develop suites of conservation measures that would ultimately serve as options for the ALWTRT to consider when identifying management alternatives. The conservation measures would address vertical line fishery interactions with large whales by reducing the potential for entanglements and minimizing adverse effects if entanglements occur.

Given the significant public interest in this topic, it will be critical for NMFS to obtain a transparent and independent review of the model documentation. It is important that the model contain the best available information on both whale density and fishing gear density and that the associated caveats seem reasonable. Therefore, we seek an independent CIE peer review of the model documentation, and the independent CIE peer review reports formatted as described in **Annex 1** will be made publicly available. The CIE reviewers shall conduct an independent and impartial scientific peer review of this scientific information in accordance with the Terms of Reference (ToRs) for the peer review as specified in **Annex 2**.

Requirements for CIE Reviewers: Three CIE reviewers shall conduct an impartial and independent peer review in accordance with the SoW and ToRs herein. The CIE reviewers shall have combined working knowledge and recent experience in spatial analysis, scenario modeling, marine mammal biology, and fisheries management. Each CIE reviewer's duties shall not exceed a maximum of 10 days to complete all work tasks of the peer review described herein.

Location of Peer Review: Each CIE reviewer shall conduct an independent peer review as a desk review, therefore no travel is required.

Statement of Tasks: Each CIE reviewers shall complete the following tasks in accordance with the SoW and Schedule of Milestones and Deliverables herein.

<u>Prior to the Peer Review</u>: Upon completion of the CIE reviewer selection by the CIE Steering Committee, the CIE shall provide the CIE reviewer information (full name, title, affiliation, country, address, email) to the COR, who forwards this information to the NMFS Project Contact no later than the date specified in the Schedule of Milestones and Deliverables. The CIE Coordinator is responsible for providing the SoW and ToRs to the CIE reviewers. The NMFS Project Contact is responsible for providing the CIE reviewers with the background documents, reports, and other pertinent information. Any changes to the SoW or ToRs must be made through the COR prior to the commencement of the peer review.

Pre-review Background Documents: One week before the peer review, the NMFS Project Contact will send (by electronic mail or make available at an FTP site) to the CIE reviewers the necessary background information and reports for the peer review. The CIE reviewers shall read all documents in preparation for the peer review, and are responsible only for the documents that are delivered to the reviewer in accordance to the SoW scheduled deadlines specified herein. In the case where the documents need to be mailed, the NMFS Project Contact will consult with the CIE Lead Coordinator on where to send documents.

Desk Review: Each CIE reviewer shall conduct the independent peer review in accordance with the SoW and ToRs, and shall not serve in any other role unless specified herein. **Modifications to the SoW and ToRs can not be made during the peer review, and any SoW or ToRs modifications prior to the peer review shall be approved by the COR and CIE Lead Coordinator.** The CIE Lead Coordinator can contact the Project Contact to confirm any peer review arrangements.

Contract Deliverables - Independent CIE Peer Review Reports: Each CIE reviewer shall complete an independent peer review report in accordance with the SoW. Each CIE reviewer shall complete the independent peer review according to required format and content as described in Annex 1. Each CIE reviewer shall complete the independent peer review addressing each ToR as described in Annex 2.

Specific Tasks for CIE Reviewers: The following chronological list of tasks shall be completed by each CIE reviewer in a timely manner as specified in the **Schedule of Milestones and Deliverables**.

1) Conduct necessary pre-review preparations, including the review of background material and reports provided by the NMFS Project Contact in advance of the peer review.

2) Conduct an independent peer review in accordance with the ToRs (**Annex 2**).

3) No later than **28 September 2012**, each CIE reviewer shall submit an independent peer review report addressed to the "Center for Independent Experts," and sent to Mr. Manoj Shivlani, CIE Lead Coordinator, via email to shivlanim@bellsouth.net, and CIE Regional Coordinator, via email to Dr. David Sampson david.sampson@oregonstate.edu. Each CIE report shall be written using the format and content requirements specified in Annex 1, and address each ToR in **Annex 2**.

Schedule of Milestones and Deliverables: CIE shall complete the tasks and deliverables described in this SoW in accordance with the following schedule.

24 August 2012	CIE sends reviewer contact information to the COR, who then sends this to the NMFS Project Contact.
31 August 2012	NMFS Project Contact sends the stock assessment report and background documents to the CIE reviewers. Background documents may be sent to the CIE reviewers one week earlier.
7-21 September 2012	Each reviewer conducts an independent peer review as a desk review.
28 September 2012	CIE reviewers submit draft CIE independent peer review reports to the CIE Lead Coordinator and CIE Regional Coordinator.
12 October 2012	CIE submits the CIE independent peer review reports to the COR.
19 October 2012	The COR distributes the final CIE reports to the NMFS Project Contact and regional Center Director.

Modifications to the Statement of Work: This 'Time and Materials' task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery Management Council, and Council's SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COR within 10 working days after receipt of all required information of the decision on changes. The COR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

Acceptance of Deliverables: Upon review and acceptance of the CIE independent peer review reports by the CIE Lead Coordinator, Regional Coordinator, and Steering Committee, these reports shall be sent to the COR for final approval as contract deliverables based on compliance with the SoW and ToRs. As specified in the Schedule of Milestones and Deliverables, the CIE shall send via e-mail the contract deliverables (CIE independent peer review reports) to the COR (William Michaels, via William.Michaels@noaa.gov).

Modifications to the Statement of Work: This 'Time and Materials' task order may require an update or modification due to possible changes to the terms of reference or schedule of milestones resulting from the fishery management decision process of the NOAA Leadership, Fishery Management Council, and Council's SSC advisory committee. A request to modify this SoW must be approved by the Contracting Officer at least 15 working days prior to making any permanent changes. The Contracting Officer will notify the COR within 10 working days after receipt of all required information of the decision on changes. The COR can approve changes to the milestone dates, list of pre-review documents, and ToRs within the SoW as long as the role and ability of the CIE reviewers to complete the deliverable in accordance with the SoW is not adversely impacted. The SoW and ToRs shall not be changed once the peer review has begun.

Support Personnel:

William Michaels, Program Manager, COR
NMFS Office of Science and Technology
1315 East West Hwy, SSMC3, F/ST4, Silver Spring, MD 20910
William.Michaels@noaa.gov Phone: 301-427-8155

Manoj Shivlani, CIE Lead Coordinator
Northern Taiga Ventures, Inc.
10600 SW 131st Court, Miami, FL 33186
shivlanim@bellsouth.net Phone: 305-383-4229

Roger W. Peretti, Executive Vice President
Northern Taiga Ventures, Inc. (NTVI)
22375 Broderick Drive, Suite 215, Sterling, VA 20166
RPerretti@ntvifederal.com Phone: 571-223-7717

Key Personnel:

NMFS Project Contact:

Kate Swails
NOAA Fisheries, Northeast Regional Office
55 Great Republic Drive, Gloucester, MA 01930
Email: Kate.Swails@noaa.gov Phone: (978) 282-8481

Annex 1: Format and Contents of CIE Independent Peer Review Report

1. The CIE independent report shall be prefaced with an Executive Summary providing a concise summary of the findings and recommendations, and specify whether the science reviewed is the best scientific information available.

2. The main body of the reviewer report shall consist of a Background, Description of the Individual Reviewer's Role in the Review Activities, Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs.

3. The reviewer report shall include the following appendices:

Appendix 1: Bibliography of materials provided for review

Appendix 2: A copy of the CIE Statement of Work