17 July 2012

To whom it may concern:

Within the framework of the Off shore Wind farm Egmond aan Zee project, on the order of Dutch Government and with their financial support, an extensive environmental monitoring program is carried out on the effects of OWEZ on the marine environment. Research area's are birds, marine mammals, fish, benthos, solid substrate and public opinion.

All draft versions of reports made within the framework of the monitoring program were reviewed by Dutch energy agency Agentschap NL and the Waterdienst, a department of the Dutch water authority Rijkswaterstaat. These organisations provided feedback that was discussed with the researchers and taken into account when making the final version of the reports. Before publication all final reports need approval of Agentschap NL and the Waterdienst.

The report on harbour porpoises (OWEZ_R_253_T1_20120202_harbour_porpoises.pdf) is written within the framework of this monitoring program.

Despite approval for publication of the report on harbour porpoises a dispute remains between the researchers of Imares on the one hand and Agentschap NL and the Waterdienst on the other about the validity of the conclusion by Imares that:
"the increase in the PPM number within the wind farm area is due to a true increase in harbour porpoise abundance".

The note at hand contains two documents relating to this dispute:
1: a note, written by TNO Technical Sciences, explaining their point of view and reflecting the position of Agentschap NL and the Waterdienst;
2: a rebuttal by Imares to the note written by TNO.

This note is the final step of the process concerning research on harbour porpoises in OWEZ, the reporting thereof, the review and publication process.

Technical Sciences
Oude Waalsdorperweg 63
2597 AK Den Haag
P.O. Box 96864
2509 JG The Hague
The Netherlands

www.tno.nl

T +31 88 866 10 00
F +31 70 328 09 61
infodesk@tno.nl

TNO report

TNO 2012 R10080

TNO review of IMARES report 'Assessment of the Effects of the Offshore Wind Farm Egmond aan Zee (OWEZ) for Harbour Porpoise (comparison T0 and T1)'

Date	June 2012
Author(s)	G. Blacquière F.P.A. Lam M.A. Ainslie C.A.F. de Jong F.H.A. van den Berg
Number of pages	10
Customer	RWS Waterdienst - Ministerie van EL&I Postbus 17 8200 AA Lelystad
Project name	IMARES review Bruinvis OWEZ
Project number	052.03189

Contents

-

1 Introduction

Two earlier draft versions of IMARES' report "Assessment of the Effects of the Offshore Wind Farm Egmond aan Zee (OWEZ) for Harbour Porpoise (comparison T0 and T1)" [1] have been reviewed by TNO. The results of these reviews were documented in [2] and [3]. Although IMARES agreed with many of the TNO comments and carefully updated the report accordingly, one particular issue remained unsolved. Therefore, Rijkswaterstaat requested TNO to write a brief report that summarizes and clarifies this remaining issue. This report is the result.

-

2 The experimental set-up

Eight T-PODs were deployed: two inside the wind farm area and two-times-three T-PODs outside the wind farm area: three at the North and three at the South side of the wind farm area respectively, see the red dots in Figure 1.

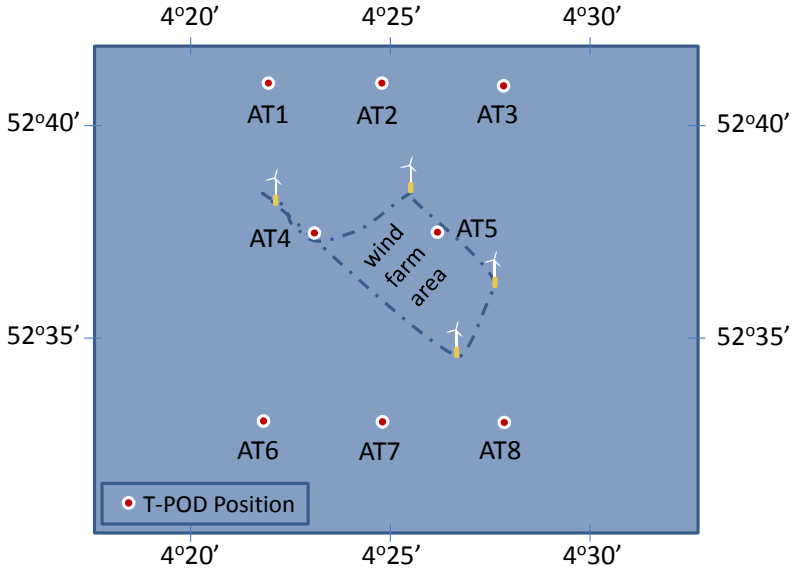


Figure 1 Locations AT1 to AT8 of the T-PODs and the wind farm area (approximate). Note that the actual number of wind turbines in the farm is 36.

The experimental set-up chosen by IMARES to study the effect of the OWEZ on the Harbour Porpoise is visualized in the middle picture of Figure 2. T-PODs are devices that measure *acoustic activity* (i.e., they do not count harbour porpoises directly). The acoustic activity that is classified as originating from Harbour Porpoises is statistically analysed. By carrying out measurements and analyses both before (T0) and after (T1) the construction of the wind farm, changes in acoustic activity between the T0 and T1 situation can be detected.
IMARES points out that a close correlation between abundance and acoustic activity still remains to be established, but provides various arguments that support such a close correlation ([1], p.48). TNO agrees with these arguments and emphasizes that amongst the currently available methods for the monitoring of harbour porpoises, the acoustic monitoring is indeed the preferred method.

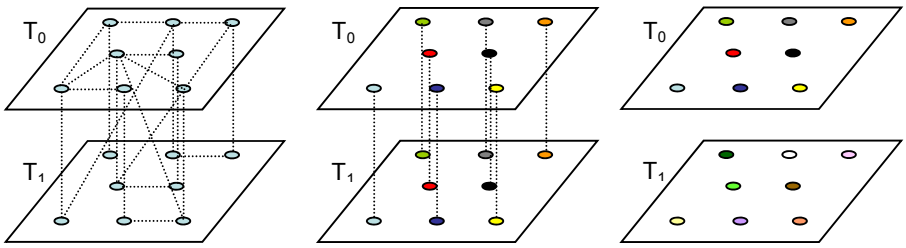
From the literature it is clear that the T-POD sensitivity may differ considerably from one device to the next. E.g., in [4] it is stated that "The results from the intercalibration and sensitivity tests showed that the sensitivity differences between the PODs varied a lot in some PODs, which means that two PODs do not necessarily record exactly the same". Therefore, in the chosen set-up it is very important that the *same device* is deployed at the *same location* during the complete experiment, i.e., both in T0 and in T1, prior to and after the construction of the wind farm respectively. Only then changes can be detected reliably. Note that the results to be obtained from the chosen set-up are not related to the absolute abundance of harbour porpoises. Instead, the results are related to *changes over*

-

time (decrease or increase) of *harbour porpoise acoustic activity* at the particular measurement location.

This possible variation in T-POD sensitivity was taken into account by IMARES when designing the experimental set-up: in [5] they state that “The significant variation between T-PODs was accounted for in the design by keeping the T-PODs at the same position throughout the monitoring period”.

Unfortunately, some of the T-PODs were lost during the experiment. They had to be replaced by other devices. The original T-PODs were of the type V3 whereas the replacement T-PODs were of the type V5. This means that the original experimental set-up (‘same T-POD at same location’) could no longer be maintained. The result is the set-up according to the right picture of Figure 2 (‘various T-PODs at same location’). As a consequence, a careful intercalibration between the various T-POD devices to be used at the same measurement location is required in order to satisfy the original experimental set-up requirements (middle picture of Figure 2).
Note that in the international literature, it is long recognized that T-PODs should indeed be calibrated before setting up a multiple T-POD monitoring programme, see e.g. [4].



Well-calibrated T-PODs. From this set-up acoustic activity can be obtained as a function of both space and time.

T-PODs not calibrated; however: same T-POD at one location; From this set-up temporal changes in acoustic activity can be obtained for each location.

T-PODs not calibrated; different T-PODs used at every location. Not much information on acoustic activity can be obtained from this set-up.

Figure 2 Various options for an experimental set-up for a before (T_0) after (T_1) analysis. The dots represent the eight T-POD locations, see also Figure 1. Dots with the same colour represent ‘the same T-POD’ or ‘well intercalibrated T-PODs’. Dots with different colours represent different, un-calibrated T-PODs.

In summary: because individual T-PODs may differ, it is important to use the same T-POD at each particular location throughout the experiment, or, if this is not possible, to carefully intercalibrate the T-PODs to be deployed at the same location.

-

3 IMARES approach to the T-POD calibration

IMARES discusses the importance of the calibration issue. However, they mention that "*it was not possible to include a factor describing the specific differences in sensitivity for each replacement of v3 with v5* …" (p. 19 in [1]).

As an alternative, they decided to carry out an *overall* V3-versus-V5 calibration.
At five of the eight locations both a V3 T-POD and a V5 T-POD were deployed at the same location for a limited period of time. The data obtained from these experiments were used to carry out this overall V3-versus-V5 calibration.
As a result of this, in their further statistical analysis IMARES used the following assumptions:
- all T-PODs V3 are identical,
- all T-PODs V5 are identical,
- T-PODs V5 are 8% more sensitive that T-PODs V3, i.e., they produce PPM (porpoise positive minutes) numbers that are 8% higher.

In summary: the various *individual* T-PODs that have been deployed at the same location in the OWEZ study have not been intercalibrated.

-

4 TNO comment

Figure 3 (which is a copy of Figure 17, p.36 in [1]) displays the result of the overall V3-versus-V5 calibration. From this figure, we get the strong impression that at location AT4, the V5 version is much more sensitive than the 8% that has now been adopted for PPM, see the yellow dots. In particular, it seems that if location AT4 would have been excluded from the analysis (as was location AT8), then the general intercalibration of V3-versus-V5 would be close to the line y = x (i.e., mainly based on the intercalibration at location AT7), meaning that V5 and V3 are equally sensitive at AT7 outside the wind farm. However, the individual AT4 V3-versus-V5 calibration (red dashed line, added by TNO) would be close to the line y = 2x (red solid line, added by TNO). It means that V5 seems to be almost twice as sensitive as V3 at AT4 within the wind farm.

IMARES has commented on our question to clarify this issue at location AT4 as follows: 'The short answer to the question is that yes, it appears that the slope of AT4 is higher than 1, but this was not statistically significant and hence not discussed' (email dd. 7 June 2010), see also p.44 of [1].
Note that AT4 is a very important location as it is one of the two locations inside the wind farm. We therefore had expected a more elaborate discussion, e.g., an explanation why the effect was not statistically significant (e.g., not enough data available?).

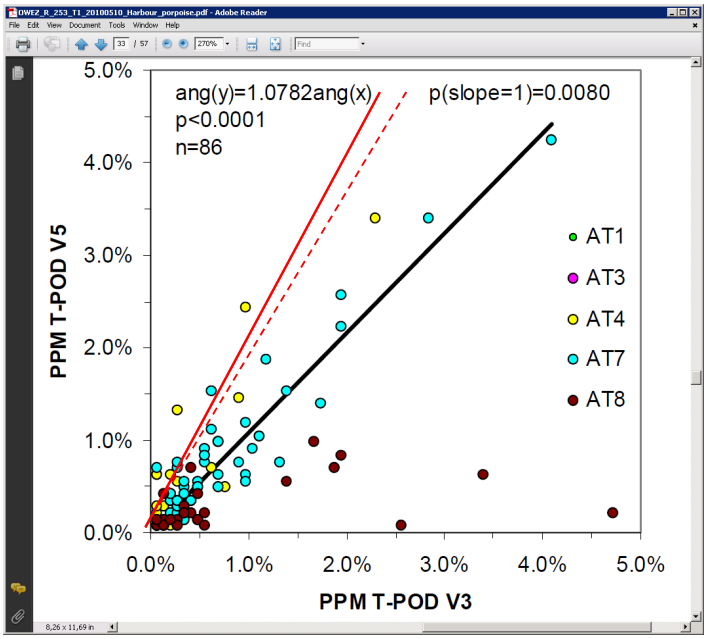


Figure 3 Copy of Figure 17 p. 36 of [1]. The red lines have been added by TNO. The solid red line represents y=2x, the dashed red line represents our estimate of the V3-V5 relation at AT4, based on the locations of the yellow dots. The black line y=1.0782x (or approximately y=1.08x) represents the estimate of the V3-V5 relation at the involved locations combined. The overall factor of 1.08 is used by IMARES.

-

Note that the intercalibration of V3 and V5 type T-PODs could only be carried out for five of the eight locations, see Figure 3. However, from the figure it is clear that almost all available data refers to three locations, of which location AT8 had to be discarded (too many outliers). As a consequence only two locations remained: one inside the wind farm area, AT4, and one outside the wind farm area, AT7.

In conclusion: Figure 3 suggests that the two TPODs used at location AT4 inside the wind farm had a larger difference in sensitivity (higher sensitivity during T1 than during T0) than the factor of 1.08 as applied by IMARES.
Furthermore, almost all available information on the intercalibration of V3 and V5 type TPODs is from two locations only, AT4 and AT7. Information from AT8 had to be discarded. There is (almost) no information from AT1, AT2, AT3, AT5 and AT6, although such information would have been needed given the experimental set-up. The limited information provided in Figure 3 does not support the assumption that it is sufficient to take into account class differences between V3 and V5 only. Instead it suggests that it is important to take into account differences between individual T-PODs.

Therefore the remaining question is: is the increase in the PPM number within the wind farm area due to a true increase in harbour porpoise abundance (or increased harbour porpoise acoustic activity) or due to a higher sensitivity of the replacement TPODs inside the wind farm or due to a lower sensitivity of the replacement TPODs outside the wind farm or due to a combination of such effects?

-

5 Consequence for the conclusion of the study

The apparent large individual variation between T-PODs – well-known from the literature, very clear at location AT8 which was discarded, and in our view also apparent at location AT4, without (much) information on the other locations – could affect the conclusion of the study. In particular this refers to the text in the *Executive summary* stating that ‘a higher porpoise acoustic activity was recorded inside the wind farm relative to outside, which is most likely linked to an increase in local porpoise occurrence’ (p.3 in [1]) as well as the hypothesized explanation of this phenomenon: ‘The cause behind the increase of porpoises in the farm could not be determined, but may be linked to increased food availability due to the reef effect of the turbine foundations and the exclusion of fishery from the wind farm’ (p.4 in [1]).

TNO is not questioning the presence of harbour porpoises during T0 as well as during T1, inside as well as outside the wind farm area. However, we argue that the question “is the increase of porpoises larger inside the wind farm than outside the wind farm?” cannot be answered with sufficient evidence on the basis of the collected T-POD data.

-

6 References

[1] Scheidat, M., Aarts, G., Bakker, A., Brasseur, S., Carstensen, J., van Leeuwen, P.W., Leopold, M., van Polanen, T., Reijnders, P., Teilmann, J., Tougaard, J., and Verdaat, H., 2012, Assessment of the Effects of the Offshore Wind Farm Egmond aan Zee (OWEZ) for Harbour Porpoise (comparison T_0 and T_1)", coded OWEZ_R_253_T1_20120202; IMARES C012.12.

[2] Blacquière, G., and Verboom, W.C., 2009, Review report IMARES harbour porpoise OWEZ $T_0 – T_1$, memorandum sent to Rijkswaterstaat by email 23 December 2009, file: TNO-memo-RWS 032.30385-01.05.pdf.

[3] Blacquière, G., and Verboom, W.C., 2010, Response to Review report OWEZ_R_253_T1_20100510_Harbour_porpoise, TNO kenmerk DHW1-2010-03794.

[4] Teilmann, J., Henriksen, O.D., Carstensen, J., 2002, Monitoring effects of offshore windfarms on harbour porpoises using PODs (porpoise detectors), Technical report for the Ministry of the Environment, Denmark.

[5] Brasseur, S., Reijnders, P., Damsgaard Henriksen, O., Carstensen, J., Tougaard, J., Teilmann, J., Leopold, M., Camphuysen, K., Gordon, J., 2004, Baseline data on the harbour pospoise, *Phocoena phocoena*, in relation to the intended wind farm site NSW, in the Netherlands, Alterra-report 1043, Wageningen, ISSN 1566-7197.

-

IMARES comment (final) to TNO comment

Point 1: AT4 T-POD

Concerning figure 17 TNO states that "*we get **the strong impression** that at location AT4, the V5 version is much more sensitive than the 8% that has now been adopted for PPM, see the yellow dots*."

It is difficult to interpret if a sensitivity is significantly different or not for a specific station / T-POD based on looking at a figure. IMARES has therefore done a detailed analyses, which can be found in the report. The results show that **AT4 does not have a statistically significant effect**. It did show that AT8 did have a significant effect and therefore it was removed in the later analyses (see below).

Report (page 29/30)

⋯ since the T-POD sensitivity is specific to T-POD unit rather than the T-POD version, differences between v3 and v5 across stations was investigated. For clicks per PPM an additive difference was expected on the log-scale, and **differences between v3 and v5 was analysed by means of a paired t-test for each station individually. There were no differences between the two T-POD versions at all stations except AT8 where v5 recorded 26% less clicks per PPM ($t_{29}=2.24$; $P=0.0329$).** There were no station-specific differences in the intercalibration slope ($F_{3,80}=0.79$; $P=0.5015$) for clicks per PPM, when data from AT8 were not included, and **none of the individual slopes or common slope (Figure 17) were significantly different from 1** suggesting that the difference for clicks per PPM between v3 and v5 was generally small and that one of the T-PODs deployed at AT8 could have been malfunctioning.

For PPM there were significant differences among the station-specific slopes ($F_{4,108}=11.61$; $P<0.0001$), but this significant difference was entirely due to the observations from AT8. **There were no significant differences among the station-specific slopes when excluding observations from AT8** ($F_{3,82}=1.74$; $P=0.1654$). **PPM was significantly higher (~8%) for v5 than v3 when observations from AT8 were not included in the intercalibration regression** (Figure 17).

For the days with simultaneous deployments of v3 and v5 T-PODs 1008 encounters and 1001 waiting times were recorded at the 5 positions. Differences between T-POD versions, positions and their interaction were investigated on the log-transform of encounters and waiting times by means of analysis of variance. Again, AT8 behaved completely different from the other positions and if observations from this position were excluded the interaction between position and T-POD version was not significant for neither encounters ($F_{3,772}=0.60$; $P=0.6119$) nor waiting time ($F_{3,766}=0.90$; $P=0.4393$). **Without AT8 observations there was no difference in encounter duration between v3 and v5 ($F_{1,775}=0.07$; $P=0.7980$), whereas waiting times were longer for v3 (22%) although not significant at a 95% confidence level**

($F_{1,769}$=2.76; P=0.0970). Thus, there is potentially a general bias towards v5 T-PODs being more sensitive that v3 T-PODs, except for AT8, where the opposite was the case.

Point 2: taking into account individual differences of T-PODs

Due to the fact that only some of the stations had the two types of T-PODs available for calibration, TNO states "it suggests that it is important to take into account differences between individual TPODs".

IMARES agrees with this and therefore in the BACI analyses we included the factors "podtype" as well as "podid". Additionally the results from the intercalibration were applied to the further analyses. Details can be found in the report.

Methods section (page 16):

…, the indicators were analysed according to a modified Before-After Control-Impact (BACI)-design (Green 1979) that included station-specific and seasonal variation as well. Variation in all four indicators reflecting different features of the same porpoise echolocation activity were assumed to be potentially affected by the following factors (5 fixed and 3 random) and combinations thereof:

- *Area* (fixed factor with 2 levels: *impact* and *control*) describes the spatial variation between control areas and impact area (wind farm).
- *Subarea(area)* (fixed factor with 3 levels: *control N, control S* and *impact*) describes the spatial variation between the three areas. As this factor is nested within *area*, it describes differences between the two control areas *control N* and *control S.*
- *Station (area subarea)* (random factor with 8 levels: *AT1-AT8*) describes the station-specific variation (variation among stations) within each of the three areas.
- *Period* (fixed factor having 2 levels: T_0 and T_1) describes the difference between baseline and operation period.
- *Year(period)* (random factor with 5 levels: *2003, 2004, 2007, 2008, 2009*) describes the variation between years within the two periods T_0 and T_1.
- *Month* (fixed factor with 12 levels: *Jan-Dec*) describes the seasonal variation by means of monthly values.
- ***Podtype*** **(fixed factor with 2 levels: v3 and v5) describes the difference between v3 and v5 T-PODs.**
- ***Podid*** **(random factor with 20 levels: serial number of T-POD) describes the random variation between different T-PODs for v3 and v5 separately.**

Four of the fixed factors (main factors *area*, *period*, *month* as well as nested factor *subarea(area)*), and their 7 interactions, describe the spatial-temporal variation in the

echolocation activity, whereas *podtype* describes a potential monitoring bias from replacing v3 with v5 T-PODs.

Results section (page 30):
The model for spatial-temporal variation as well as T-POD specific variation (Eq. 1) and an ARMA(1,1) correlation structure was computed for the 4 indicators. Only 6 out of the 12 fixed effects in Eq. (1) could significantly explain variation in the echolocation indicators (Table 6). **For none of the four indicators the T-POD specific variation was found significant, neither as a systematic bias between v3 and v5 nor as a difference in the variation between T-PODs for the two versions.** Although v5 yielded slightly higher echolocation activity than v3 in the models, the bias was not significant relative to the large overall residual variation, when the T-PODs were deployed in a natural environment. These results correspond to the general results (without station-specific intercalibration) obtained from the intercalibration of the two T-POD types on a reduced data set (Section 3.2.3). However, in the intercalibration analysis it was also realised that the T-PODs at position AT8 behaved significantly different from any other pair of T-PODs deployed simultaneously, and that T-POD recordings from this position may impair the overall conclusion that the change from v3 to v5 T-PODs did not affect conclusions (see also further discussion below). This deviating pattern with a decrease in click monitoring from v3 to v5 could be due to an extraordinary sensitive transducer in POD323 (v3) or an equally insensitive transducer in POD702 (v5).

The BACI analyses were consequently carried out on two data sets: 1) assuming no effect of switching from v3 to v5 T-PODs and 2) **using the results from the intercalibration analysis to recalculate indicators from v5 to v3. In the intercalibration data set data from the v5 T-pod at AT8 were discarded, because the v3 (POD323) was deployed at AT8 in both T_0 and T_1 and thus even if this T-POD had a deviating sensitivity this difference would be the same for both T_0 and T_1 and thus not affect the BACI results.** Moreover, **PPM values calculated from v5 data was divided by the estimated intercalibration factor of 1.078 and waiting times calculated from v5 data was multiplied by the estimated intercalibration factor of 1.22.** These two data sets will be referred to as the non-calibrated and intercalibrated data sets, respectively. The models obtained with both data sets, after eliminating non-significant effects, were similar in structure and allowed for a direct comparison of the intercalibration exercise.
….

Differences between the BACI analyses carried out on non-calibrated data (Table 6) vs. intercalibrated data (Table 7) were generally small. The significance of the different factors was generally reduced with the intercalibrated data set, and the variation between Control N and Control S (*subarea(area)*) turned insignificant and was removed. The most important difference between the two analyses was that the BACI factor (*area×period*) became non-significant for waiting time with the intercalibrated data set. This change was mainly caused by excluding data from POD702 (v5) that increased the mean waiting time in the control area during T_1. **Due to the suspect data from this T-POD and the expected improved sensitivity switching from v3 to v5, the analysis based on the intercalibrated data set is believed to more correct than the analysis based on non-calibrated data, and in the following results from the BACI analysis using the intercalibrated data set will be shown only.**

….

All four indicators also showed a significant increase in echolocation activity from T_0 to T_1: clicks per PPM increased from 33.9 clicks/min to 47.2 clicks/min, PPM more than tripled from 0.22% to 0.67%, encounter duration increased from 3.5 minutes to 4.6 minutes, and waiting times decreased from 14.0 hours to 7.7 hours. However, **the significance of *area×period* for clicks per PPM and PPM as well as a tendency for relatively longer encounters and shorter waiting times in the impact area during T_1 suggested that echolocation activity in the impact area increased more than in the reference area (Figure 18). Echolocation activity was similar in the two areas during the baseline, but increased significantly more during the operation period in the impact area. The increase in the impact area relative to the reference areas was 24% for clicks per PPM, 109% for PPM, 15% for encounter duration and a 20% decrease in waiting times.**

Point 3: final conclusion

TNO asks the following question: “is the increase in the PPM number within the wind farm area due to a true increase in harbour porpoise abundance (or increased harbour porpoise acoustic activity) or due to a higher sensitivity of the replacement TPODs inside the wind farm or due to a lower sensitivity of the replacement TPODs outside the wind farm or due to a combination of such” ?

IMARES trusts that based on the analyses done and presented in detail in the report, the answer to this question is: **yes, the increase in the PPM number within the wind farm area is due to a true increase in harbour porpoise abundance.**