File

INTERNAL DOCUMENT 51

I.O.S.

Results of Data Validation procedures carried out at the Institute of Oceanographic Sciences on data from the United Kingdom Offshore Operators Association environmental data collection programme.

[This document should not be cited in a published bibliography, and is supplied for the use of the recipient only].

NATURAL ENVIRONMENT RESEARCH COUNCIL

INSTITUTE OF OCEANOGRAPHIC SCIENCES

INSTITUTE OF OCEANOGRAPHIC SCIENCES

Wormley, Godalming,
Surrey, GU8 5UB.
(042-879-4141)

(Director: Dr. A. S. Laughton)

Bidston Observatory,
Birkenhead,
Merseyside, L43 7RA.
(051-652-2396)
(Assistant Director: Dr. D. E. Cartwright)

Crossway,
Taunton,
Somerset, TA1 2DW.
(0823-86211)
(Assistant Director: M.J. Tucker)

CONFIDENTIAL

Results of Data Validation procedures carried out at the Institute of Oceanographic Sciences on data from the United Kingdom Offshore Operators Association environmental data collection programme.

INTERNAL DOCUMENT NO 51

This document should not be cited in any other paper or report except as 'personal communication', and it is for the use of the recipient only.

Results of Data Validation Procedures carried out at the Institute of Oceanographic Sciences on data from United Kingdom Offshore Operators Association Environmental Data Collection Programme

Current Meter Data

Station: Boyle

Period: April 1976 to the termination of the Boyle Station, May 1977.

No 5 in a series of reports to be submitted to the Department of Industry and the Department of Energy by the Natural Environment Research Council, Institute of Oceanographic Sciences, Taunton. The final year of data collected at the Boyle station completes the validation of the current meter data collected for this programme.

CURRENT METER DATA

CONTENTS

V Report

METHODS OF CURRENT DATA VALIDATION

The current data are received punched on paper tapes in the form of an impeller count and a meter direction (in degrees magnetic) taken every five minutes. The tapes are copied onto magnetic tape and magnetic disk. The latter are edited to remove any obviously false data from the beginning or end of the records. At this stage the data value corresponding to the immersion of the meter is identified as accurately as possible.

A computer program then performs simple checks on the data, removes spurious spikes and applies compass direction corrections. The program includes checking:

(i) the format of the records

(ii) that no compass direction exceeds 360 degrees

(iii) that no impeller count exceeds 1024 (the maximum recorded by the instrument)

(iv) that the differences calculated between consecutive current speeds do not exceed a given value (usually 20 cm sec^{-1}). (Any isolated large or small value of current speed is assumed to be due to instrument malfunction and is replaced by the previous good value. On rare occasions this may cause a 'cascade' effect which results in the replacement of a series of good data values. If this happens on an isolated occasion then the faulty value is manually replaced by an interpolated value and the computer program is rerun. However, if a succession of replacements occurs then the program is rerun on the original data values using a higher difference value (up to 80 cm sec^{-1}) so that the data series can be inspected).

The program produces two files, one containing the five minute checked data and the other containing half-hourly vector averages computed from the five minute checked data.

The five-minute checked data are examined for faults. The individual data values in each record are plotted on a polar diagram and each of the vector points is joined to the next in sequence. This polar plot is viewable on a VDU, whilst a permanent copy is made on a 35mm aperture card using a Microplotter. Both pieces of equipment are peripherals of the computer system. Any gross errors can be confirmed and less obvious errors may be revealed by plotting the individual five minute directions as a time series, and by resolving the five minute vector values into north and east velocity components and plotting these as a time series. When the data are found to be faulty, the records are not processed any further, and are classed as 'unacceptable'.

If no obvious errors have been seen in the plots of the five-minute checked data, then the corresponding half-hourly vector-averaged data are resolved into tidal and non-tidal (residual) components. The technique used is the response method, in which a computer program is used to correlate the vector-averaged data with a tidal height series at a nearby site, and so calculate the tidal and non-tidal components. (The non-tidal components form the basic data for many of the available techniques for estimating extreme currents). This method works best with data series lengths which are close to the lunar period of about 29 days, and therefore data of only a few days' duration are not processed in this way. The reference tidal series used in this validation work are those for Lerwick and Scilly. The method is fully described in a paper by Munk et al (1966) and more briefly in another by Cartwright et al (1969).

The next stage is an examination of these vector-averaged data and their components. Progressive vector plots are produced of both the vector-averaged currents and the non-tidal components; these show the virtual displacements at the meter locations, and also give an indication of how well the tidal components have been removed. If the plots show a large proportion of tidal energy remaining, this may be due either to incorrectly recorded velocities, in which case reference to the plots of the north and east components as time series may give some confirmation of this; or it may be due to an incorrect sampling interval, which may be detected by spectrally analysing the vector-averaged data and checking whether the peak lies at the tidal frequency as calculated from the stated sampling interval. (A further check on the sampling interval can be made by dividing the total time of deployment by the number of data points obtained; this may be less reliable in those cases when the deployment and retrieval times are not accurately identified).

If no outstanding error has been detected at this stage, the data is classed as 'acceptable'. It should be noted that no check is made on the magnitude alone of the currents before placing them in this category. It is possible to place reliance on the visual inspection of plots only because over a recording period of two or three weeks the current data show certain overall regularities and patterns which change very little from one month to another at each recording station. A typical pattern to emerge in the case of the polar plots is an elliptical distribution of currents, with the ellipse orientated in a specific way for each recording station (for example at Stevenson station good data almost always have the major axis of the ellipse aligned approximately NW-SE). Another prominent characteristic is the sense in which the current vector can be seen, on the VDU, to rotate during plotting, which is almost always clockwise for all stations. When data shows an orientation of ellipse which is a reflection in the north-south axis of that found for the majority of data at a specific station, and when in addition the current vector has been observed to rotate in an anti-clockwise direction, it is usually assumed that the meter was placed on the mooring line upside-down. (The current meter design allows it to be moored upside down and still record data, which will then show both these characteristics). Further confirmation is provided by comparing the progressive vector diagram with those from meters moored at the same site for the same period but at different depths: if the virtual displacement is again a mirror image in the north-south axis of those for other meter positions then there is a high probability that the meter was moored upside-down. In these cases, if the record is otherwise acceptable it is classed as 'reprocessable' since the data that it would have recorded had the meter been properly moored can be salvaged by subtracting the original recorded directions from 360°.

Bottom current meter data does not appear to follow so closely the patterns described above (bottom currents are those recorded only 3 or 4 metres above the seabed). These data only rarely show the regular patterns that have been observed in the other data: the distribution of current vectors does not appear so obviously elliptical, the sense of rotation of the vectors is not always clockwise (at Fitzroy especially it is nearly always anti-clockwise) and the virtual displacement is usually much more complicated than the virtual displacements found in the top and middle waters. These irregularities may accurately describe the water movements at those depths, made complicated by the bottom topography; or they may be indicative of consistently malfunctioning current meters. This was difficult to ascertain in previous IOS validation reports when the same meter was deployed for consecutive months. The problem was highlighted on one occasion when a current meter was deployed for consecutive months producing very consistent, but very obviously faulty, data. However, with the more recent procedure introduced by Marex where a variety of meters is interchanged at various depths it would appear that some of the irregularities recur from month to month and meter to meter, suggesting that they reflect real phenomena at the sites. Many of the irregular bottom current meter records have been placed in the 'acceptable' category because only those records which have obvious and definite faults in them are placed in the 'unacceptable' category. This procedure is performed with more confidence in the light of the recent data returns for bottom meters covered by this report than has previously been the case.

Due to the complexities discussed above concerning current systems in confined waters, visual inspections of data plotted in a number of different ways have necessarily formed a large part of the quality control. The limitation of this approach is that only those errors which are evident from the plots are detected, although, as previously mentioned, spectral analysis is sometimes used to uncover less obvious errors.

The results of the quality control are presented in three sections, one for each location. At the end of each section the total durations are given in hours for each of the following:

(i) the data it was attempted to collect

(ii) the data found to be 'acceptable'

(iii) the data found to be 'unacceptable'

These figures have been calculated by multiplying the number of data points on each record (as notified by Marex) by the sampling interval in hours. In cases where a meter has been lost, the record length which might have been obtained has been taken as equal to the record lengths obtained by other meters deployed at the same time. In cases where all meters have been lost an average record length for that particular station has been substituted.

REFERENCES

1. Munk, W.H. and Cartwright, D.E. 1966. Tidal spectroscopy and prediction. Phil. Trans., A259, 533-581.

2. Cartwright, D.E., Munk, W. and Zetler, B. 1969. Pelagic tidal measurement. Transactions of the American Geophysical Union, 50, 472-477.

DISCUSSION OF THE RESULTS

This report completes the validation work carried out on the current meter data collected by Marex Ltd. The data for the Boyle station considered in this report cover the period 10 April 1976 to 31 May 1977 and two sets collected in July 1974, which had not been sent with the early data (refer to validation report No 2, p 9). The results from these two deployments are included in the summary figures on page 8 and tables 1 and 2. During the main period under consideration there were 45 deployments of current meters which attempted to collect data. Seven of these meters failed to collect data and these have been acknowledged by Marex in their reports. As they usually have no record numbers the deployment periods, meter numbers and nature of the fault as notified by Marex are listed:

7. 5.76	to	1. 6.76	meter no 429	:	fault in tape
23.10.76	to	16.11.76	" " ")		Successive deployments of a faulty meter
17.12.76	to	11. 1.77	" " ")		
11. 2.77	to	8. 3.77	" " ")		
11. 3.77	to	4. 4.77	" " 414	:	Subsurface buoy waterlogged. Sank meter to the sea bed.
8. 4.77	to	3. 5.77	" " 428	:	Faulty meter
6. 5.77	to	31. 5.77	" " 416	:	Meter flooded

Of the remaining 40 records all were examined by IOS. One record with acceptable data, AS 78, is a short record and has not been included in any Marex report. The four faulty records placed in the unacceptable category are:

AS 80 : 3.7.76 to 27.7.76 : meter no 414.

Marex reported that this tape was faulty and the IOS analysis confirmed this. The cause appears to have been a malfunction in the compass electronics which caused some direction readings to be restricted to 90 and 180 degrees.

AS 84 : 27.8.76 to 21.9.76 : meter no 429.

Marex reported this record faulty, which was confirmed by IOS analysis. The velocity series contained excessively large readings, mostly in the second half of the data. This was the first of four successive but unsuccessful deployments of this faulty meter at 15m and the third of sevel total deployments in the period under review. Apart from one short record (AS 78) six of the seven failed to collect acceptable data.

AS 92 : 23.10.76 to 16.11.76 : meter no 428.

The validation procedures show that this data was collected with an average sampling interval of 4m 50sec (Marex reported this as 4.8m). A DNC-2 current meter uses a set of miniature plugs and sockets to set its sampling period from 15 sec to 31 min 45 sec in 15 second increments. Clearly it is not possible to set the meter to record at 4 m 50 sec and indeed this is only an average sample period derived from the number of data points and the time the meter was deployed. It is not possible to establish from the analysis what the variability of sampling interval was, if any, and therefore this record has been classed as unacceptable. However it has been included in the Marex data reports (see table 1).

AS 103 : 11.2.77 to 8.3.77 : meter no 428.

The direction readings for this data show no particular pattern for the second half of the record. This could be interpreted as the meter fouling the sea bed or the mooring as the velocities are also small for this period. As a result the data has been classified as unacceptable by both IOS and Marex.

Data is only classified as unacceptable if it can be clearly demonstrated that the data is faulty. Data which are borderline cases are given the benefit of the doubt, and are placed in the acceptable category. Several such records occurred in these series of data and should be treated with caution. They are:

AS 67 : 10.4.76 to 4.5.76
AS 74 : 4.6.76 to 29.6.76
AS 81 : 30.7.76 to 24.8.76
AS 87 : 25.9.76 to 18.10.76
AS 91 : 19.11.76 to 14.12.76

These successive deployments of meter 383 all show brief series of zero velocities at the point of current reversal in the velocity time series plots. It is possible that these indicate that the rotor was sticking. Alternatively as these deployments are all at the bottom (3 m above the sea bed) they may represent a real phenomenon, an explanation which seems to be confirmed by two other factors. Other meters deployed at this depth also exhibit this behaviour and meter 383 when deployed at 76 m for its last three deployments, no longer displays this symptom.

AS 106 : 11.3.77 to 4.4.77 : meter no 331

AS 108 : 8.4.77 to 3.5.77 : meter no 363

AS 111 : 6.5.77 to 31.5.77 : meter no 363

These are the three bottom meter deployments which follow those referred to in the previous section. All of them exhibit the short periods of zero velocities and in addition they also contain intermittent spikes in the direction time series plots when the compass appears to briefly stick or the electronics malfunction and the direction is recorded as 0, 90, 180 or 270 degrees. The data should be treated with caution.

AS 99 : 16.1.77 to 8.2.77 : meter no 383

AS 100 : 16.1.77 to 8.2.77 : meter no 331

These two meters, deployed on the same mooring at the same time, appear to have been deployed in the reverse way to that notified by Marex. The progressive vector plots and the velocity time series plots both show that the magnitude of the currents is generally larger for the data from meter 383, also Marex reported the maximum velocities to be 33.58 cm sec^{-1} for AS99, and 46.96 for AS100. This all tends to suggest that AS99 is the bottom meter data and that AS100 is the lower meter data. The two sets of data should be used with caution.

The total durations of the various categories are:

Attempted collection	-	24614 hours
Acceptable data	-	19482 hours
Unacceptable data	-	2354 hours

Expressing the durations of the data in the validation categories as a percentage of the time covered by the meter deployments (attempted collection):

Analysed by IOS	-	88.7%
Acceptable data	-	79.2%

For a comprehensive summary of the results see Tables 1 and 2.

R Gleason
Institute of Oceanographic Sciences
Crossway, Taunton.

10 May 1979

TABLE 1

BOYLE

Summary of the Results

			Record Sub-total	Record Totals
Acceptable		AS: 65,66,67*,72,73,74*,75, 76,77,78*,79,81*,82,83, 85,86,87*,88,89,90,91*, 93,96,97,98,99*,100*,102, 105,106*,107,108*,110,111*, 4, 6.	36	36
Unacceptable	Faults in compass direction	AS: 80, 103	2	4
	Sampling interval errors	AS: 92	1	
	Velocity faults	AS: 84	1	
Deployed, but no data return	Meter malfunction, mechanical failure		7	7
			Total	47

The following record, classed as unacceptable by IOS, appears in a Marex report.

Report	Figure No	Record No	Dates
Boyle Autumn 1976	19	AS92*	23.10.76 to 16.11.76

*See comments on these records in the main text

TABLE 2

SUMMARY OF RETURNS BY DEPTH

Measuring Site	Acceptable data (in hours and as a percentage of the time covered by this report)				Total time covered by this report (ii)
	Top	Upper	Lower	Bottom	
Boyle (i)	3524	-	7000	8047	10008
(iii)	35%	-	70%	80%	100%

This table shows:

(i) the number of hours of acceptable data returned for each meter depth;

(ii) the total time covered by this report, ie the maximum time over which data could have been continuously collected;

(iii) the acceptable data from (i) expressed as a percentage of the total time covered by (ii).