Mathematical description of the Standard Wave Analysis Package

Prepared for: MINISTRY OF TRANSPORT, PUBLIC WORKS AND WATER MANAGEMENT

DIRECTORATE-GENERAL OF PUBLIC WORKS AND WATER MANAGEMENT

Date: September 6, 1994

Mathematical description of the Standard Wave Analysis Package

Prepared for: MINISTRY OF TRANSPORT, PUBLIC WORKS AND WATER MANAGEMENT

DIRECTORATE-GENERAL OF PUBLIC WORKS AND WATER MANAGEMENT

By: The Oceanographic Company of the Netherlands b.v.

Author: Drs G. W. Hoekstra

Co-authors: L. Boeré, ir A.J.M. van der Vlugt, ir T. van Rijn

Date: September 6, 1994

CONTENTS

1. Introduction

1.1. Background

Within the framework of the Rijkswaterstaat Monitoring Network Infrastructure the Oceanographic Company of the Netherlands b.v. (OCN) has developed a standard analysis module for wave sensor data called SWAP (Standard Wave Analysis Package). The method for the routine analysis of the raw data and the parameters to be calculated have been defined by Rijkswaterstaat [4]. This definition establishes a standard within the Rijkswaterstaat that facilitates the exchange and interpretation of long-term wave information. Wave information from the North Sea is needed for operational use, research and wave climate monitoring. Usually this data concerns the open sea where waves in the frequency range between 30 and 500 mHz are relevant, but when estuaries are monitored higher frequencies up to 1000 mHz are important.
This report has been written to make a computer implementation of the standard Rijkswaterstaat wave analysis method possible. It gives the exact definitions to a level providing sufficient information for the development of a computer program.

1.2. Overview

Within the Rijkswaterstaat Hydro Meteo Networks several types of sensors are currently being used, being the Wavec, the Directional Waverider, the Waverider and the Stepgauge. The Wavec (measuring vertical displacement and slopes in the east-west and north-south direction) and the Directional Waverider (measuring vertical displacement and horizontal displacement in the east-west and north-south direction) are used for both height and directional measurements, while the Stepgauge and the Waverider can only be used for measuring heights. The output signals of these sensors are pre-processed by a SESAM (SEnsor Signal Adaption Module) that produces raw wave data in a standard format described in [3]. Depending on the sensor type the standard format raw data consists either of only one signal in the vertical direction (for wave height only sensors) or of a signal in the vertical direction plus two more signals in the horizontal plane (for wave height plus wave direction sensors).
The process of analysing the raw data is divided into six different phases, which are depicted schematically in Fig. 1.1 and which are described in the subsequent chapters of this report (Note that the processes described in phase 5, spectral direction analysis, can only be carried out for data from wave height and direction sensors):

1. **Data collection (Chapter 2)**.
   The data delivered by the SESAM is collected by a SCADA (Supervisory Control and Data Acquisition) system that performs data checks and generates a standard error code for invalid datapoints. The time of collection for each datapoint is determined and a timestamp is attached. From the SCADA system the data is sent to SWAP in a standard format message where the cm is used as the unit of distance.

2. **Pre-processing of the raw data (Chapter 3).**
   During this phase the data is prepared for further processing. Parameters describing the quality of the incoming data are calculated (section 3.6). The data samples are grouped together to form an analysis period (section 3.2). For Stepgauge signals tide filtering is performed (section 3.3). Error codes and physically unrealistic datapoints (spikes and signals that remain constant over a long period) are removed from the dataset (section 3.4, which also describes the criteria defining these situations). After error codes are removed the mean is subtracted from the signal (section 3.5).

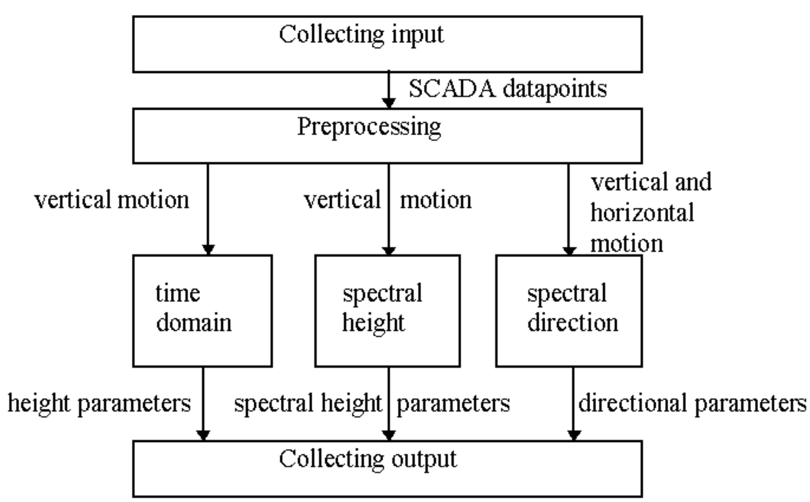
3. **Time domain analysis (Chapter 4).**
   Only the wave height signal is used in the time domain analysis. The number of waves is counted and the waves are sorted according to height and period (section 4.3). The waves are then processed (section 4.4), resulting in wave height and period parameters.

4. **Spectral height analysis (Chapter 5).**
During this phase the basic wave height time series is Fourier transformed and a wave height spectrum is constructed (section 5.2). The spectrum is analysed in the frequency domain, yielding frequency-dependent parameters. A number of frequency bands have been defined over which these parameters are averaged to obtain the characteristics of this band: estimators for wave height (subsection 5.3.3), wave period (subsection 5.3.4) and wave energy (subsection 5.3.2) are derived.

5. **Spectral direction analysis (Chapter 6).**
If there are directional signals available, spectra for both the vertical and horizontal wave motion are constructed. Information on the characteristics of the directional energy distribution is provided. The method used for characterising the directional energy distribution of the waves does not attempt to reconstruct this distribution since only its first four Fourier coefficients can be derived from the spectra, which is not enough to approximate the directional distribution with a truncated Fourier series [5]. Instead, some characteristic parameters such as mean direction, directional width and two parameters characterising the degree of asymmetry and the flatness of the energy distribution are estimated, without assuming a priori a shape of the distribution [2]. This model-free approach does not suffer from the shortcomings of a reconstruction method, where shape assumptions may not be justified, the directional resolution may be misleading and the results hard to interpret. The mean direction and width of the distribution are presented as a function of the frequency. Several frequency bands have been defined over which the mean direction, width, asymmetry, flatness and wave height are averaged.

6. **Collecting output (Chapter 7).**
The results of the analysis are made available at three output levels:
   - The raw wave data.
   - The basic information needed for calculating the characteristic parameters.
   - The estimated parameters.

Figure 1.1. The six phases of the raw data analysis.



2 Data collection

2.1 Introduction

This Chapter describes the data collection from the sensor to the SWAP module.
The raw output signals of the sensors are converted by a SESAM (SEnsor Signal Adaption Module) to the standard format described in [3]. The data from the SESAM is collected by a SCADA (Supervisory Control and Data Acquisition) system that amongst others performs data checks, generates a standard error code for invalid datapoints, determines the time of arrival and attaches a timestamp to each datapoint. The SCADA output is sent to SWAP in a standard message format where the cm is used as the unit of distance. Despite this standard format information on the sensor type is also necessary for the correct processing of the data, as is described in the next section.

2.2 Determining sensor characteristics

Although the incoming data is delivered to SWAP in a standard format, some information on sensor characteristics is required in order to be able to determine how the data should be processed and whether directional analysis can be performed. In SWAP the incoming datastream is therefore linked to a specific sensor type that defines the necessary sensor characteristics.
Depending on the sensor type, only data on the vertical wave motion or a combination of vertical wave motion and wave motion in the horizontal plane will available. The sensor type is therefore a necessary parameter to determine whether to perform directional analysis or not.
It is furthermore required for determining whether tide filtering needs to be done before other processing can be performed (section 3.3). In case of tide filtering a correction filter has to be applied before the spectral processing (subsection 5.2.5).
The type of heave correction filter, i.e. the correction filter needed for compensation of the sensor characteristics, is also determined from the sensor type (subsection 5.2.5).
In Table 2.1 the different sensors currently in use within Rijkswaterstaat are compared with respect to their sample rates, their output type (only height or direction and height) and the correction filters that need to be applied:

Table 2.1 Comparison of the Rijkswaterstaat sensors

Sensor	Sampling rate (Hz)	Output type	Filters
Waverider	2.56	height	heave correction
Stepgauge	2.56	height	tide + tide correction
Wavec	1.28	height/direction	heave correction
Directional Waverider	1.28	height/direction	heave correction

3 Pre-processing

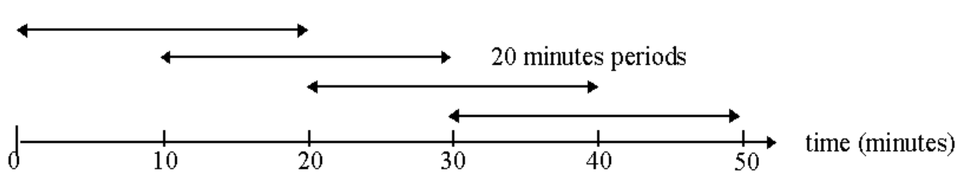
3.1 Introduction

This Chapter describes the pre-processing of the raw sensor data as it is presented to SWAP by the SCADA system. During this phase the data is prepared for further processing. The data samples are first grouped together to form an analysis period (section 3.2). The tide filtering necessary for Stepgauge signals is described in section 3.3. Section 3.4 deals with the removal of error codes and physically unrealistic datapoints such as spikes and constant signals. After the error codes are removed the mean is subtracted from the signal (section 3.5). The calculated parameters, which describe the quality of the incoming data, are presented in section 3.6.

3.2 Construction of the analysis period

Each incoming datasample is provided with a timestamp in MET (Middle European Time) in milliseconds resolution of the time of arrival. The sample rate (depending on the sensor type) and the timestamp are required to be able to put the incoming samples in place in the analysis period.
Independent of sensor type, the analysis of the wave data takes place over a 20 minutes period which is shifted every 10 minutes, creating a 10 minutes overlap with the previous analysis period (see Fig. 3.1).

Figure 3.1. Construction of the 20 minutes analysis period.



The timestamp attached to the 20 minutes period is defined as the mean of the start time and the end time in MET. The 20 minutes periods are started 6 times per hour: at 0, 10, 20, 30, 40 and 50 minutes past the hour.
The 20 minutes period can be constructed if at least one valid datapoint is present, therefore even if just a single datapoint is collected during a 20 minutes interval an analysis period is constructed and processing is started.

3.3 Tide filtering

All sensors, except for the Stepgauge, are equipped with a high pass filter, the so-called ‘heave filter’, to prevent DC drift of the signal in the frequency range where the sensor response is not defined. These filters also remove the tidal component from the sensor signal. Filtering the Stepgauge signal to remove its tidal component makes it possible to treat the Stepgauge signal in the same way as the other sensor signals. For further processing compensation for these filters is necessary, this is however possible only in the frequency domain (see Chapters 5 and 6), not in the time domain (see remark in section 4.5).

The Stepgauge signal is filtered by subtracting the tidal component, which is defined as all vertical motion with a frequency less than 30 mHz, from the signal. The tidal component is calculated from the raw data before the error correction (section 3.4), but datapoints containing error code or having an ‘extremely large’ deviation from the average are not taken into account. The definition of ‘extremely large’ is made precise by a so-called 4-sigma test: datapoints falling outside the interval [μ-4σ, μ+4σ] are excluded from the tide calculation. Note that rejected points are not set to error code but are simply not taken into account in calculating the tide. The standard deviation σ and the average μ defining this interval are given by:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N-1}} \qquad \mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

where N is the number of datapoints x_i not containing error code. If no valid datapoints are found σ and μ are set to error code and if the number of valid datapoints is one, σ is set to zero and μ is set to the value of the datapoint.
The tide calculation is accomplished using a 30 mHz cut-off low-pass filter. This filter is composed of 4 (identical) moving average (MA) filters. One such MA filter averages each datapoint D_i over a range [-(M-1)/2, (M-1)/2]:

$$\frac{1}{M}\sum_{k=i-(M-1)/2}^{i+(M-1)/2} D_k$$

where the odd integer M is the width of the filter. Applying the MA filter 4 times results in an equivalent 'weighted' MA filter. After filtering the signal with datapoints D_i, the corresponding points T_i of the tide are given by:

$$T_i = \frac{\sum_{k=-(N-1)/2}^{(N-1)/2} E_{i+k} W_k D_{i+k}}{\sum_{k=-(N-1)/2}^{(N-1)/2} E_{i+k} W_k}$$

where W_k is the weight factor and N the total width of the filter which equals 4(M-1)+1 (how W_k and N are determined will be described below). The factor E_{i+k} equals 0 for those datapoints D_{i+k} that contain error code or violate the 4-sigma test described above, and equals 1 for valid datapoints D_{i+k}. Note that in calculating the tide at a certain time, values of the signal in both the past and the future are required. This causes problems at the edges of the analysis period since values outside the analysis period are unknown. Therefore factors E_{i+k} corresponding to $i+k$ outside the analysis period are set to 0.
As mentioned above, datapoints on error code or violating the 4-sigma test are not used in the tide calculation. However, if more than a certain fraction of the weighted datapoints D_{i+k} of the signal is not taken into account, the corresponding tide point T_i is set to error code. The minimum number of weighted datapoints D_{i+k} for tide point T_i to qualify is defined in a way such that the sum of the weight factors W_k used in calculating the tide point T_i should be 90% or more of the total of all weight factors. In formula:

$$\frac{\sum_{k=-(N-1)/2}^{(N-1)/2} E_{i+k} W_k}{\sum_{k=-(N-1)/2}^{(N-1)/2} W_k} \geq 0.9$$

where E_{i+k} and N are defined as above.
This limit may not be reached due to the presence of too many error codes or due to the lack of knowledge of past and future values at the front end and back end of the analysis period. These values are not known and so the fraction of weighted datapoints near the edges of the analysis period drops.
The total filter width N equals 4(M-1)+1, where M depends on the cut-off frequency and the sample rate. The proper value for M can be derived from the following formula [1]:

$$\frac{1}{2} = \left(\frac{\sin(M\pi f\ /\ SR)}{M\ \sin(\pi f\ /\ SR)}\right)^4$$

where SR is the sample rate of the sensor and f the cut-off frequency of the filter (30 mHz).
The width of the filter is determined by solving the above formula for M and choosing the nearest odd value for M. Table 3.1 gives the best values for M and N corresponding with the two sample rates used by the Rijkswaterstaat sensors with around 30 mHz cut-off frequency.

Table 3.1 Recommended values for M and N for the sample rates of the sensors

SR (Sample rate in Hz)	Best value for N	Best value for M	Cut-off frequency (mHz)
1.28	49	13	31.5
2.56	105	27	30.3

The values of the weight coefficients W_k depend on the width M of the MA filter and are symmetric around k=0, i.e. $W_k=W_{-k}$. The weight coefficients are computed in the ranges $k \in [0, M-1\rangle$ and $k \in [M-1, 2M-2]$ (square brackets indicate that boundaries are included) according to the following formulae:

$$W_{k\varepsilon[0,M-1>} = \frac{4}{6}(k-M-1)(k-M)(k-M+1) - \frac{1}{6}(k-2M-1)(k-2M)(k-2M+1)$$

$$W_{k\varepsilon[M-1,2M-2]} = -\frac{1}{6}(k-2M-1)(k-2M)(k-2M+1)$$

After the tide has been computed, it can be subtracted from the signal, but first the gaps (rows of error codes) in the tide have to be repaired, i.e. replaced by valid values.
The gaps at the front and back end of the tide (resulting from error codes in the input data and from the fact that data outside the analysis period is unknown) are repaired by linear extrapolation of a line through the first and last tide point of the analysis period containing no error code. Extrapolation at the front and back end is done only if the gap size is smaller than 75 seconds. For larger front end and back end gaps no extrapolation is performed.
Gaps in the tide surrounded by valid tide points are repaired by linear interpolation if the gap size is smaller than 150 seconds. Linear interpolation is performed between the last valid tide point before the gap and the first valid tide point after the gap.
When subtracting the tide from the signal, only those datapoints are computed for which both the signal and the tide are not on error code, other points are set to error code. Note that the error codes in the signal have not yet been repaired (see section 3.4). Note also that information reduction occurs when valid datapoints are rejected because of error codes in the corresponding tide points (e.g. gaps too large for interpolation or extrapolation).

3.4 Error handling

Once a 20 minutes period has been assembled and optionally tide-filtered, it has to be checked for errors and if possibly repaired before any further processing can take place. The raw data may contain error codes or values which, although valid, is physically unrealistic. Datapoints having the same value as a ‘large’ number of neighbouring datapoints and datapoints having a ‘large’ derivative are considered ‘unrealistic’ and are rejected by setting them to error code in case of the following situations:

- Valid datapoints having a constant value over a time interval exceeding 10 seconds in duration (‘0-sigma test’). Error codes in this 10 seconds interval are not taken into account.
- Valid datapoints outside the interval [-4σ, +4σ] around the average value of 0 (‘4-sigma test’).
- Valid datapoints with a derivative outside the interval [-4δ, +4δ] around the average value of 0 (‘4-delta test’).

The parameter σ is defined as the standard deviation of the valid datapoints:

$$\sigma = \sqrt{\frac{\sum_{j=1}^{N} x_j^2}{N}}$$

where the summation is carried out over all datapoints in the analysis period that do not contain error code and N is the number of these valid datapoints.
The parameter δ is defined as the standard deviation of the derivative. The derivative in datapoint x_i is defined as the difference Δx_i between datapoint x_i and the previous datapoint x_{i-1} if both x_i and x_{i-1} are not on error code. If either x_i or x_{i-1} contains error code the derivative in point x_i is not calculated. The definition of the standard deviation of the derivative is given by the following formula:

$$\delta = \sqrt{\frac{\sum_{j=1}^{M} (\Delta x_j)^2}{M}}$$

where the summation is carried out over all derivatives Δx_j that have been defined above. M is the number of pairs for which the derivative is defined.

Because all signals have been filtered by a high-pass filter the average of the datapoints and the average of the derivatives are assumed to be zero. As the averages are not determined from the data, using the factors N (instead of N-1) and M (instead of M -1) in the denominator in the formulas for σ and δ respectively is justified.
If no valid datapoints are found σ is set to error code, and if no pairs of consecutive valid points are found δ is set to error code.
The 0-sigma, 4-sigma and 4-delta tests described above are carried out in this order. The datapoints that have been rejected by one of the tests are set to error code are not taken into account by the next.
Gaps (one or more consecutive datapoints on error code) can be repaired under certain conditions. If the gap size does not exceed 2 seconds (where the gap size in seconds is defined as the number of error codes divided by the sample rate in Hz) it can be repaired by linear interpolation across the gap. In the same way, gaps at the front end or back end of the analysis period not longer than 1 second are extrapolated by continuation of the value of the first respectively the last valid datapoint. Note that although the *size in seconds* of the largest gap that can be repaired is the same for every sensor, the *number of datapoints* in the maximum gap depends on the sample rate of the sensor type.

3.5 Removing 20 minutes mean

In spite of the high pass filtering already performed by the sensor itself (Wavec, Waverider and Directional Waverider) or by the tide filtering for the Stepgauge within SWAP as described above, a small DC component may still be present in the signal (e.g. caused by removing sigma and delta errors, interpolation or extrapolation). Before further processing a 20 minutes block of data, this DC component is removed by subtracting the 20 minute mean from the signal. This is done mainly for the time domain analysis, for the spectral analysis averages over shorter periods are removed separately (see subsection 5.2.1). Datapoints still on error code after correction by interpolation and extrapolation are not used in the further processing.

3.6 Calculated parameters

During the pre-processing phase various parameters indicating the quality of the data are calculated (note that the parameters giving a number of datapoints are sample rate dependent):

- Ngd_zP: percentage of vertical wave motion datapoints that do not contain error code before pre-processing
- Ngd_xP: percentage of horizontal east-west wave motion datapoints that do not contain error code before pre-processing
- Ngd_yP: percentage of horizontal north-south wave motion datapoints that do not contain error code before pre-processing
- Nu_z: number of valid vertical wave motion datapoints that are rejected because of 0-sigma errors
- Nu_x: number of valid horizontal east-west wave motion datapoints that are rejected because of 0-sigma errors
- Nu_y: number of valid horizontal north-south wave motion datapoints that are rejected because of 0-sigma errors
- Nd_z: number of valid vertical wave motion datapoints that are rejected because of 4-delta errors
- Nd_x: number of valid horizontal east-west wave motion datapoints that are rejected because of 4-delta errors
- Nd_y: number of valid horizontal north-south wave motion datapoints that are rejected because of 4-delta errors
- Nv_z: number of valid vertical wave motion datapoints that are rejected because of 4-sigma errors
- Nv_x: number of valid horizontal east-west wave motion datapoints that are rejected because of 4-sigma errors
- Nv_y: number of valid horizontal north-south wave motion datapoints that are rejected because of 4-sigma errors
- Ni_z: number of interpolated or extrapolated vertical wave motion datapoints
- Ni_x: number of interpolated or extrapolated horizontal east-west wave motion datapoints
- Ni_y: number of interpolated or extrapolated horizontal north-south wave motion datapoints

4 Time domain analysis

4.1 Introduction

This Chapter deals with the time domain analysis of the waves, which takes place after the pre-processing described in Chapter 3. Only the wave height signal is used during this phase. The crest height calculation is described in section 4.2, the wave classification process where the waves are counted and sorted according to height and period in section 4..3, the calculation of the wave height and period in section 4.4, and the resulting parameters in section 4.5.

4.2 Crest height

Prior to the wave classification (section 4.3) the crest height is determined. The crest height is defined as the maximum positive value of the vertical wave motion datapoints in the analysis period. Error codes are not taken into account.

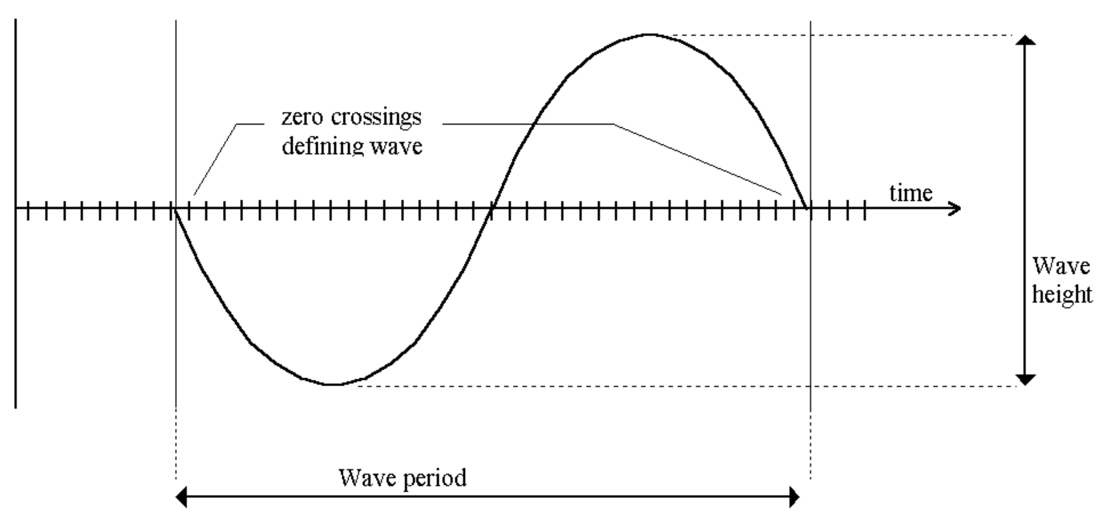
4.3 Wave classification

After the pre-processing and crest height determination, the waves within the analysis period are counted and sorted according to height and period. In order to be able to do this, it is necessary to give the definition of a wave:

A wave is defined as the part between two down-going zero crossings in the wave height signal.

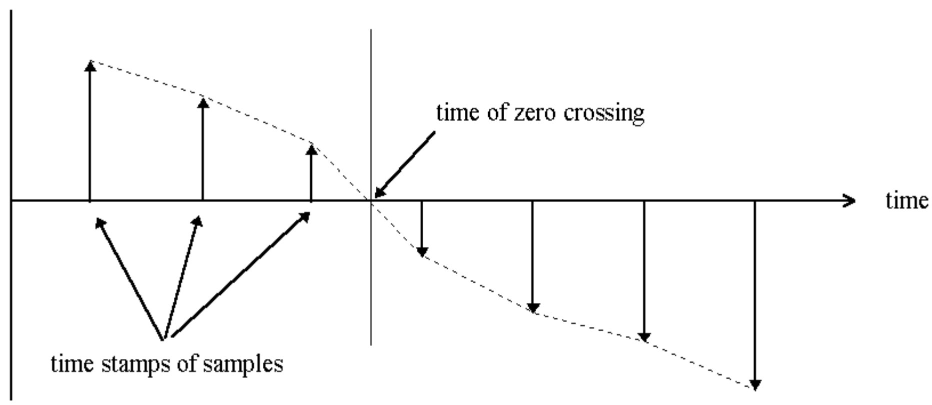
The wave period is defined as the difference between the end time and the start time of the wave. The wave height is defined as the difference between the maximum positive and the maximum negative value in the wave period (see Figure 4.1).

Figure 4.1. A wave.



Linear interpolation between the sampled values is used to determine the zero crossings (see Figure 4.2).

Figure 4.2. Linear interpolation to determine time of zero crossings.



The definition given above is complicated by the additional restriction that two consecutive ‘small’ half-waves have to be eliminated before the remaining waves can be counted and sorted according to height and period. The reason for elimination of small waves is that small waves introduce extra zero crossings thereby cutting the wave periods and wave heights in half.
A ‘small’ half-wave is a half wave period that lasts less than one second. Note that the definition of a small half-wave only depends on its duration and not on its amplitude.
Depending on the sign of the half wave it is designated as either a type ‘A’ or a type ‘B’ small wave (Fig. 4.3 and 4.4). In the figures the small half-waves are printed **bold** and vertical lines surround complete waves.

Figure 4.3. Type ‘A’ one small half-wave. It is part of a normal wave. No elimination is done.

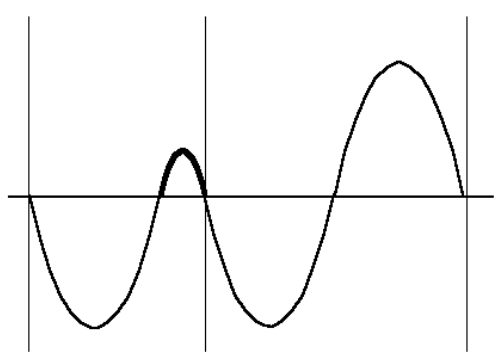
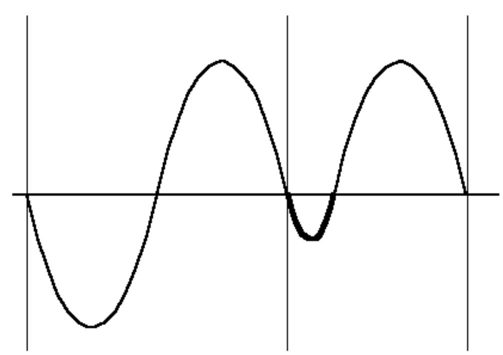


Figure 4.4. Type ‘B’ one small half-wave. It is part of a normal wave. No elimination is done.



If only one small half-wave is encountered it is either the first part (type ‘B’) or the last part (type ‘A’) of a normal wave. However, in case a second small half-wave follows the first one, the two small half-waves are eliminated by incorporating them in the normal waves. Two different situations can be distinguished:

- The first small half-wave is a type ‘A’ small half-wave. In this case the first small half-wave is incorporated in the previous negative period while the second small half-wave is taken with the next positive period, resulting in one wave. The two small half-waves can be said to be located in the middle of a normal wave. See Fig. 4.5.
- The first small half-wave is a type ‘B’ small half-wave. Now the first small half-wave is incorporated in the previous positive period ending the previous wave while the second half-wave is taken with the next negative period starting a new wave, resulting in two waves. The two small half-waves can be said to be located in between two normal waves. See Fig 4.6.

The reconstructed normal waves are indicated by the dotted lines and are surrounded by corresponding vertical lines to mark their begin and end points.

Figure 4.5. Type 'A' two small half-waves. Both are eliminated, incorporating them in one wave.

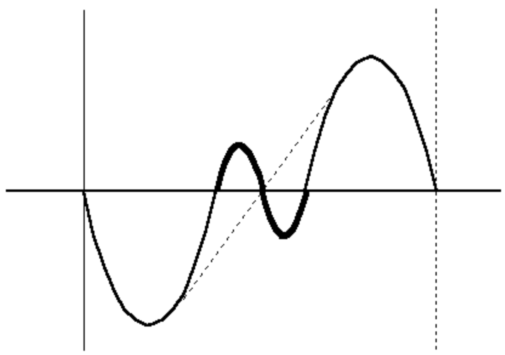
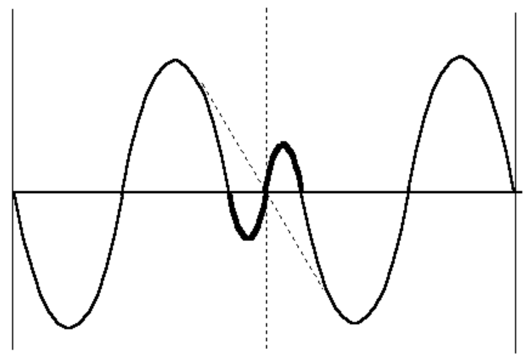


Figure 4.6. Type 'B' two small half-waves. Both are eliminated, incorporating the first one in the previous wave and the second one in the next wave.



Several consecutive small half-waves can occur. To prevent the waves from growing indefinitely when more small half-waves are encountered, a limit of two eliminations is imposed. The third small half-wave is treated as a normal wave period, after which the process of elimination can start once again. How this is accomplished can best be illustrated with the help of the following figures (see Fig 4.7 - 4.10).

Figure 4.7. Type 'A' three small half-waves. The first two are eliminated, the third is treated as a normal wave period.

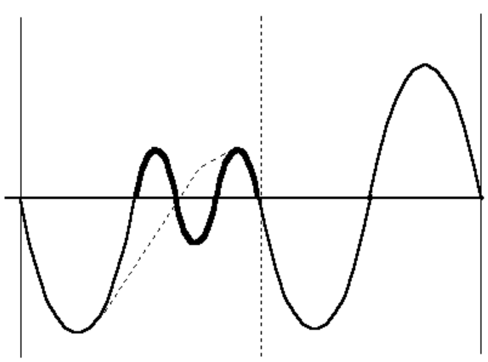


Figure 4.8. Type 'B' three small half-waves. The first two are eliminated, the third is treated as a normal wave period.

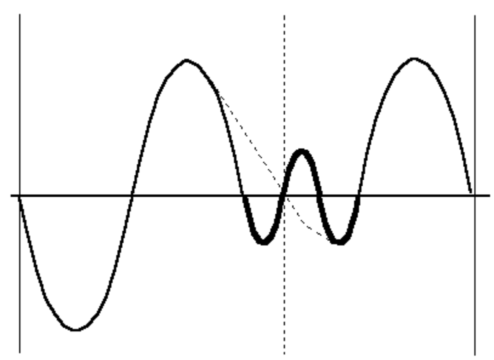


Figure 4.9. Type 'A' four small half-waves. The first two are eliminated (dotted line), the third is treated as a normal wave period and the fourth is not eliminated since it is considered a single small half-wave.

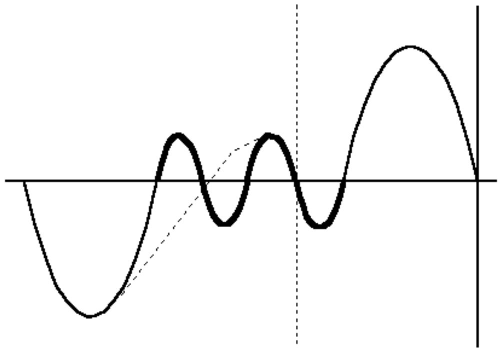
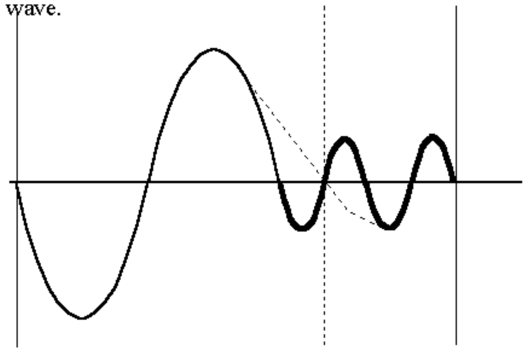


Figure 4.10. Type 'B' four small half-waves. The first two are eliminated (dotted line). The third is treated as a normal wave period and together with the fourth small half-wave this results in a new normal wave.



It is easy to see that, imposing the constraint mentioned above, a train of five consecutive small half-waves is equivalent to two small half-wave eliminations of type 'A' and type 'B' in a row. In the same way even larger trains of small half-waves can be reduced to the situations depicted in the Fig. 4.7 - 4.10.
Although error corrections have taken place (section 3.4) some datapoints may still be on error code. When an error code is encountered, the processing of that wave is terminated and a search for the next wave is started, beginning at the next down-going zero crossing.
For all the processed waves the height and period as defined in Fig. 4.1 are determined and stored in the wave period and wave height tables. The waves in the wave tables are stored in the order of their occurrence in time. Note that only full waves without error codes are stored: waves containing error codes are skipped.
To calculate the partial averages (see next section) the wave tables are sorted twice:

1. The first time they are sorted in order of decreasing wave height (for the calculation of the partial average wave height of a certain part of the highest waves), at the same time sorting the corresponding wave periods (on behalf of the partial average period of 1/3 of the highest waves).
2. The second time they are sorted in order of decreasing period (for the calculation of the partial average period of 1/3 of the longest waves).

4.4 Wave height and period averaging

The individual wave heights and wave periods from the waves determined above are averaged to obtain height and period estimators. The standard deviations of the wave height and wave period are calculated by:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N - 1}}$$

where μ is the average wave height (or wave period), x_i the individual wave height (or period) and N the number of waves taken into account.
For some parameters the averaging takes place over part of the waves from the in decreasing order sorted wave table, for instance over 1/3 or 1/10 of the highest (longest) waves. This number is not necessarily an integer: one wave may only partially fall within the required range and is therefore weighted correspondingly in the calculations. For example, the total number of waves found is N_T, the part required for averaging of the waves is 1/P. Therefore the number of waves needed for averaging is N_T/P. Only an integer number N_R waves falls completely within this range. The weighting factor for the N_R+1-th wave (i.e. the next highest (longest) wave after the N_R -th wave) is $N_T/P - N_R$. In formula:

$$\bar{x} = \frac{\left(\frac{N_T}{P} - N_R\right)x_{N_R+1} + \sum_{n=1}^{N_R} x_n}{\frac{N_T}{P}}$$

For calculating the partial average of the waves a minimum number of waves is required. This number is set to half the value of P (rounded upwards if necessary). For example: 0.5*50 => 25 in case of the 1/50 partial average and 0.5*3 => 2 in case of the 1/3 partial average. If the required minimum number of waves is not reached the average is set to error code.

4.5 Calculated parameters

Note that the heave filters in the Waverider, Directional Waverider and Wavec and the tide filter applied to the Stepgauge signal are not compensated for in the time domain wave analysis. This may lead to an underestimation of the wave height at low frequencies.
In the time domain analysis phase the wave periods and heights, their maxima and their averages and standard deviations are calculated. One parameter concerning the quality of the measurements is determined: the percentage of the total analysis time taken up by complete waves in the above sense.
For each analysis period the following 17 parameters concerning wave height and wave period are calculated as described above (unit of wave heights is cm, unit of wave periods is s):

- WTBH(i), i ∈ [1,AG] Table of wave heights
- WTBT(i), i ∈ [1,AG] Table of wave periods
- AG Total number of waves
- Hmax Height of highest wave
- H1/50 Average of the height of the highest 1/50 of the waves
- H1/10 Average of the height of the highest 1/10 of the waves
- H1/3 Average of the height of the highest 1/3 of the waves
- GGH Average of the height of all waves
- SPGH Standard deviation of the wave height
- Tmax Period of longest wave
- THmax Period of highest wave
- T1/3 Average of the period of the longest 1/3 of the periods
- TH1/3 Average of the period of the highest 1/3 of the waves
- GGT Average of the period of all waves
- SPGT Standard deviation of the wave period
- HCM Crest height, maximum positive value of all data within one analysis period
- Nwt_zP Sum of periods of waves divided by analysis period

5 Spectral height analysis

5.1 Introduction

This Chapter explains the spectral height analysis. During this phase the basic wave height timeseries is Fourier transformed to obtain an energy density spectrum (section 5.2). This spectrum is analysed in the frequency domain, yielding several frequency-dependent parameters. These parameters are averaged over a number of frequency bands to obtain the characteristics of these bands: the wave energy, the wave height and wave period per band are derived.

5.2 Construction of the energy density spectrum

The vertical motion data contains information on the frequency dependence of the energy density function, the wave height and wave period. To extract this information the wave height frequency spectra of the signal are constructed by Fourier transforms.
This process requires 6 steps, which are elaborated on in the subsequent subsections:

1. The sampled 20-minute period is split up into 6 subseries of 200 seconds each.
2. A 10% cosine tapering is applied to each subseries to minimise leakage in the frequency domain [6].
3. Fast Fourier Transformation is carried out on each of the tapered subseries.
4. The energy density spectrum is formed by addition of the resulting wave height frequency spectra.
5. Correction filters are applied to compensate for the previously applied sensor and tide filters.
6. From the corrected 5 mHz resolution energy density spectrum a smoothed 10 mHz resolution energy density spectrum is derived.

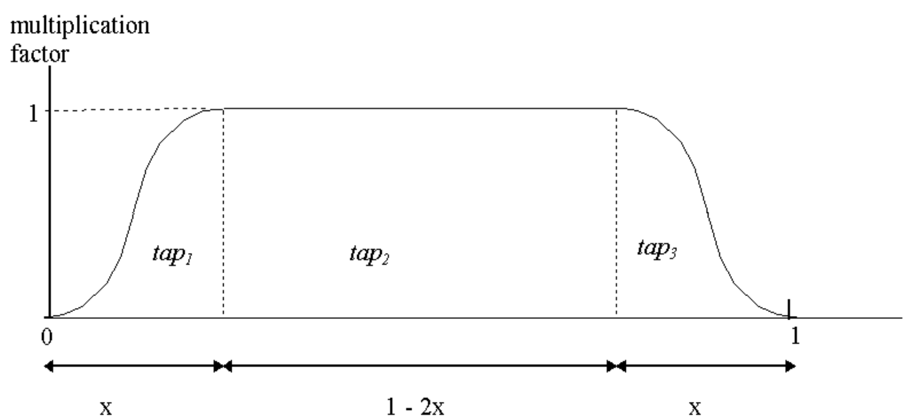
5.2.1 Construction of subseries

After the pre-processing outlined in Chapter 3 the 20 minutes analysis period is split up into six subseries, each subseries lasting 200 seconds. The number of valid subseries (i.e. subseries not containing error codes) is counted and only the valid subseries are retained, subseries containing error codes are ignored. From each of the subseries the 200 seconds mean is removed. (Note that although the mean of the entire 20 minutes period has already been removed (section 3.5) an offset per 200 seconds subseries may still be present.)

5.2.2 Cosine tapering

To prevent leakage between frequencies during the Fourier transform, a cosine taper is applied to the valid subseries, i.e. the subseries are multiplied by an invertable function that gently falls down to zero near the edges. The taper function consists of an up-going part tap_1, a unity part tap_2 and a down-going part tap_3 that is the opposite of the first part tap_1 (see Fig. 5.1).

Figure 5.1 Continuous tapering function in the range <0,1>.



The first part tap_1 is described by:

$$tap_1(t) = \frac{1}{2}\left[1 - \cos\left(\pi \frac{t}{x}\right)\right]$$

where t runs from 0 to x. When tapering a discrete, sampled subseries, the tapering function is sampled with the same number of points as the subseries. The datapoints of the subseries are multiplied by the corresponding points of the tapering function, yielding the tapered subseries. The components of the discrete taper function $tap_1(i)$ are calculated by:

$$tap_1(i) = \frac{1}{2}\left[1 - \cos\left(\pi \frac{i - 1/2}{Nx}\right)\right]$$

where $i \in$ [1, int(Nx)] with int(Nx) the integer part of Nx and N the number of datapoints in the subseries. The values of $tap_3(i)$ can be expressed as a function of $tap_1(i)$: $tap_3(i) = tap_1(N+1-i)$.
As a consequence of the tapering, the total signal is scaled by a factor equal to the area of the tapering function, which depends on the tapering fraction x (in SWAP a value of x=0.1 (10%) is used). The scaling factor r is equal to the integral of the square of the tapering function:

$$r = \int tap_1^2 + \int tap_2^2 + \int tap_3^2$$

$$r = 2\int_0^x tap_1^2(t)dt + (1 - 2x) = 1 - \frac{5x}{4}$$

where x is the tapering fraction (between 0 and 0.5, see Fig 5.1). The integral of the first and third part of the tapering function are equal and have been grouped in the first term.
The loss caused by the tapering is compensated for by multiplying all datapoints by the square root of the inverse of the scaling factor (√1/r) before the Fourier transformation and subsequent squaring of the absolute values. Note that the scaling factor is determined using a continuous tapering function and not by calculating the integral of the sampled tapering function.

5.2.3 Fourier transformation

With the sample rates used for the sensors within Rijkswaterstaat (1.28 and 2.56 Hz) the number of datapoints in each subseries is a power of two, making it possible to use a radix-2 Fast Fourier Transform routine. For each valid subseries, the Fourier transform yields a complex spectrum with components FZ_f between 0 Hz and the Nyquist frequency (which equals half the sample rate) and a resolution of Δf =1/(length of subseries) Hz. The definition of the Fourier transform used within SWAP is given by:

$$FZ_m = \sum_{n=1}^{N} Z_n e^{-2\pi i(m-1)(n-1)/N}$$

where FZ_m is the m-th spectral component of the Fourier Transform and m has integer values in the range [1, N], Z_n are the samples of the time subseries and N is the number of datapoints within one subseries.
Negative frequencies are ignored, and m is therefore limited to the range [1, (N/2)+1]. The frequency f and the index m are related according to the formula $f = (m-1)\Delta f$ where Δf equals the spectral resolution (=0.005 Hz). The lowest frequency is 0 Hz, the maximum frequency depends on the sample rate SR and is equal to the Nyquist frequency $SR/2$. This yields the following spectral components FZ_f (where the frequency f in Hz is related to m as described above): FZ_0, $FZ_{0.005}$, $FZ_{0.010}$, ...,$FZ_{N/2 * 0.005}$.

5.2.4 The 5 mHz energy density spectrum

After the Fourier transformation the spectral components FZ_f of each subseries are multiplied by their complex conjugate FZ_f^* and added to obtain a real valued spectrum. This spectrum is converted to an energy density spectrum. Normalisation is done over the number of valid subseries. The components of the resulting one sided energy density spectrum equal (expanding the complex spectral components into real and imaginary parts):

$${}_5C_{zz}(f) = \frac{1}{nN}\frac{2}{SR}\sum_{i=1}^{n} Re({}^iFZ_f)Re({}^iFZ_f) + Im({}^iFZ_f)Im({}^iFZ_f)$$

where *SR* is the sampling rate in Hz, *N* is the number of datapoints within one subseries, *n* is the number of valid subseries, and the factor 2 results from the conversion from a two sided density function to a one sided function. The functions Re and Im take the real respectively the imaginary part of their complex arguments. The subscript $_5$ denotes that the resolution of the energy density spectrum is 5 mHz.
The result is an uncorrected one sided energy density spectrum consisting of the spectral components $_5C_{zz}(f)$.

5.2.5 Correction filters

In the Wavec, the Directional Waverider and the Waverider high-pass filters which influence both amplitude and phase characteristics of the signal are employed to prevent the signal from drifting in the frequency range where the sensor response is not defined. A heave correction filter is applied to the spectra of the Waverider, Directional Waverider and Wavec data to compensate for the filter characteristics of the heave filter in the sensors. The spectral components $_5C_{zz}(f)$ of the energy density spectrum as determined in subsection 5.2.4 are multiplied by a frequency and sensor dependent correction factor H_f (f is the frequency in Hz, which is a multiple of the resolution as in subsection 5.2.3: 0.005 Hz) to give the corrected components $_5C^c_{zz}(f)$:

Correction for Wavec and Waverider [7]:

$$_5C^c_{zz}(f) = {_5C_{zz}}(f)\,H_f$$

$$H_f = R_f^2 + I_f^2$$

$$R_f = 1 - a_f^2 - a_f\,b_f\,\sqrt{2}$$

$$I_f = a_f^2\,b_f - a_f\,\sqrt{2} - b_f$$

$$a_f = \frac{1}{30.8 * f}$$

$$b_f = \frac{1}{170 * f}$$

Correction for Directional Waverider [8]:

$$_5C^c_{zz}(f) = {_5C_{zz}}(f)\,H_f$$

$$H_f = 1 + a_f^6$$

$$a_f = \frac{1}{30.8 * f}$$

The tide filter applied to the Stepgauge data as described in section 3.3 is a high pass filter with a 30 mHz cut-off frequency which also alters the signal. Some influence of the filter is still noticeable above 30 mHz and this is corrected by multiplying the spectral components of the energy density spectrum by a frequency dependent correction factor G_f [1]:

Correction for Stepgauge:

$$ {}_5C^c_{zz}(f) = {}_5C_{zz}(f)G_f $$

$$ G_f = \left[1 - \left(\frac{\sin(M\pi f\ /\ SR)}{M\sin(\pi f\ /\ SR)} \right)^4 \right]^{-2} $$

where *SR* represents the sample rate and *M* (as determined in section 3.3) is the width of one filter section such that the cut-off frequency approximates 30 mHz.
The superscript ^{c} indicating the corrected spectra will be dropped from now on.

5.2.6 Smoothed 10 mHz energy density spectrum

The band parameters that are calculated in the next section are based on a 10 mHz resolution spectrum, which components are obtained by smoothing the 5 mHz resolution spectrum components (which effectively comes to halving the number of frequency points):

$$ {}_{10}C_{zz}(f) = \frac{{}_5C_{zz}(f - \Delta f)}{4} + \frac{{}_5C_{zz}(f)}{2} + \frac{{}_5C_{zz}(f + \Delta f)}{4} $$

where Δf is the resolution of the 5 mHz spectrum in Hz (i.e. 0.005 Hz). The components of the 10 mHz spectrum are defined at multiples of the resolution (0.010 Hz), starting at 0 Hz.
In case one of the 5 mHz spectral components is not available, the corresponding 10 mHz component is set to error code.

5.3 Energy density spectrum parameters

In [4] several frequency bands have been defined over which some characteristic parameters are determined: the wave energy (subsection 5.3.2), the wave height (subsection 5.3.3) and the wave period (subsection 5.3.4). For these calculations different moments of the energy density spectrum are necessary, which are described in subsection 5.3.1.

5.3.1 Moments

The band parameters are computed with the help of the moments of the 5 mHz or the 10 mHz resolution energy density spectrum. The *n*-th order moment of a spectrum band is defined as:

$$ M_n = \Delta f \left[\frac{1}{2} l^n C_{zz}(l) + \sum_{f=l+\Delta f}^{h-\Delta f} f^n C_{zz}(f) + \frac{1}{2} h^n C_{zz}(h) \right] $$

where *l* is the lower frequency of the band, *h* the upper frequency and Δf is the resolution of the spectrum in Hz.

5.3.2 Wave energy

The wave energy per frequency band equals the sum of the components of the 10 mHz resolution energy density spectrum multiplied by the spectral resolution. Therefore the band energy equals M_0 as defined in the previous subsection.

5.3.3 Wave height

Since the wave energy is proportional to the square of the wave height, an estimate for the wave height over a frequency band can be derived from the band energy. The following formula from [9] is used:

$$H = 4\sqrt{M_0}$$

where M_0 is defined over the desired frequency band using the 10 mHz energy density spectrum.

5.3.4 Wave period

The wave period of a specific frequency band is defined as the reciprocal of a weighted average of the frequency over the band. Two different wave periods are provided:

$$average\ period = \sqrt{\frac{M_o}{M_2}}$$

$$minus\ \ first\ moment\ \ period = \frac{M_{-1}}{M_0}$$

where the moments are taken from the 10 mHz energy density spectrum. In case the denominator in the one of the above expressions is zero, the value for the wave period is set to error code.

5.4 Calculated parameters

The following parameters are calculated during the spectral height analysis phase. Note that the boundaries are included in the ranges given below. The unit for wave height is cm, energy density is given in cm^2s, energy in cm^2 , period in s and frequency in Hz. The parameters in the range up to 1000 mHz can only be calculated for sensors with a sample rate SR ≥ 2 Hz, which excludes the Wavec and the Directional Waverider since their sample rate is 1.28 Hz.

• Czz5(i), i ∈ [0,100]	5 mHz energy density spectrum ${}_5C_{zz}(f)$ Czz5(i)= ${}_5C_{zz}(f)$ where f=i*5 mHz, $f \in$ [0,500] mHz
• Czz5_M(i), i ∈ [0,200]	5 mHz energy density spectrum ${}_5C_{zz}(f)$ Czz5_M(i)= ${}_5C_{zz}(f)$ where f=i*5 mHz, $f \in$ [0,1000] mHz
• Czz10(i), i ∈ [0,50]	10 mHz energy density spectrum ${}_{10}C_{zz}(f)$ Czz10(i)= ${}_{10}C_{zz}(f)$ where f=i*10 mHz, $f \in$ [0,500] mHz
• Czz10_M(i), i ∈ [0,100]	10 mHz energy density spectrum ${}_{10}C_{zz}(f)$ Czz10_M(i)=${}_{10}C_{zz}(f)$ where f=i*10 mHz, $f \in$ [0,1000] mHz
• M0	Band energy from ${}_{10}C_{zz}(f)$ in the range $f \in$ [30, 500] mHz
• Hm0	Significant wave height from M0
• M0_M	Band energy from ${}_{10}C_{zz}(f)$ in the range $f \in$ [30, 1000] mHz
• Hm0_M	Significant wave height from M0_M
• Tm02	Average period from M_0 and M_2 in the range $f \in$ [30, 500] mHz
• Tm02_M	Average period from M_0 and M_2 in the range $f \in$ [30, 1000] mHz
• TE0	Band energy from ${}_{10}C_{zz}(f)$ in the range $f \in$ [500, 1000] mHz
• TE1	Band energy from ${}_{10}C_{zz}(f)$ in the range $f \in$ [200, 500] mHz
• TE1_M	Band energy from ${}_{10}C_{zz}(f)$ in the range $f \in$ [200, 1000] mHz
• TE2	Band energy from ${}_{10}C_{zz}(f)$ in the range $f \in$ [100, 200] mHz
• TE3	Band energy from ${}_{10}C_{zz}(f)$ in the range $f \in$ [30, 100] mHz
• HTE3	Wave height from TE3
• Fp	Frequency f where ${}_{10}C_{zz}(f)$ has its maximum in the range $f \in$ [30, 500] mHz
• Fp_M	Frequency f where ${}_{10}C_{zz}(f)$ has its maximum in the range $f \in$ [30, 1000] mHz

• HS7	Wave height from band energy from ${}_5C_{zz}(f)$ in the range $f \in [30, 142.5]$ mHz[1]
• Tm-10	Minus first moment period from M_{-1} and M_0 in the range $f \in [30, 500]$ mHz
• Tm-10_M	Minus first moment period from M_{-1} and M_0 in the range $f \in [30, 1000]$ mHz
• Ndlr_H	Quality parameter: number of valid subseries
• AV10_H	Theoretical number of degrees of freedom of the energy density spectrum (= 4 * Ndlr_H)

[1]This specific wave height parameter is calculated over 1/7-th of the 1000 mHz spectrum. The lower bound of the frequency band defined this way is fixed at 30 mHz since the sensors are not defined below 30 mHz, and the upper bound is fixed at 1000/7 mHz which is approximated by 142.5 mHz. The wave height parameter is calculated with the following formula (see [10]):

$$HS7 = 4\sqrt{\frac{1}{2}C_{zz}(l)\Delta f + \Delta f \sum_{f=l+\Delta f}^{h} f^n C_{zz}(f)}$$

where $C_{zz}(h)$ is the energy density at 140 mHz (which is the interval 137.5-142.5 mHz) and $C_{zz}(l)$ the energy density at 30 mHz. Note that in this specific case the 5 mHz resolution energy density spectrum is used.

6 Spectral directional analysis

6.1 Introduction

The data on both vertical and horizontal motion is analysed in the frequency domain to give the directional spectrum parameters. The spectral analysis of the vertical motion has already been described in detail in the previous Chapter, this Chapter focuses on the spectral directional analysis. From the auto, co and quad spectra (see section 6.2) four types of parameters are calculated:

1. Fourier coefficients (section 6.3),
2. Directional spectra (section 6.5),
3. Directional spectrum parameters (section 6.6),
4. Band parameters (section 6.7).

6.2 Construction of auto, co and quad spectra

In total 9 spectra are constructed: 3 auto spectra, 3 co spectra and 3 quad spectra. These spectra are determined analogously to the energy density spectrum in section 5.2 and will be represented by C_{zz}, C_{xx}, C_{yy}, C_{zx}, C_{zy}, C_{xy}, Q_{zx}, Q_{zy} and Q_{xy} with x and y indicating the 2 orthogonal components of the horizontal motion and z indicating the vertical motion. Only 8 of the spectra are used for the calculation of the directional wave parameters, Q_{xy} is not used.

6.2.1 Subseries construction, cosine tapering and Fourier transformation

To obtain the auto, co and quad spectra, each of the three signals (x, y, and z) is processed exactly as described in subsection 5.2.1. The resulting valid subseries are then cosine tapered as described in subsection 5.2.2 and Fourier transformed as in subsection 5.2.3, resulting in the three spectra: *FZ*, *FX* and *FY*.

6.2.2 The 5 mHz auto, co and quad spectra

Data in the vertical direction and the two horizontal directions is required in order to be able to do spectral directional analysis. Even if data in just one direction is missing, no directional processing can be performed. Consequently, only groups of valid subseries can be used, where a group is defined as a combination of three subseries (one in each direction and corresponding in time).
The auto spectra are constructed similar to subsection 5.2.4 (resulting in the spectra ${}_5C_{zz}$, ${}_5C_{xx}$ and ${}_5C_{yy}$ with spectral components ${}_5C_{zz}(f)$, ${}_5C_{xx}(f)$ and ${}_5C_{yy}(f)$), but instead of normalising over the number of valid subseries, normalisation is done over the number of groups n of valid subseries. SR is the sampling rate and N the number of datapoints within one subseries:

$${}_5C_{zz}(f) = \frac{1}{nN}\frac{2}{SR}\sum_{i=1}^{n} Re({}^iFZ_f)Re({}^iFZ_f) + Im({}^iFZ_f)Im({}^iFZ_f)$$

$${}_5C_{xx}(f) = \frac{1}{nN}\frac{2}{SR}\sum_{i=1}^{n} Re({}^iFX_f)Re({}^iFX_f) + Im({}^iFX_f)Im({}^iFX_f)$$

$${}_5C_{yy}(f) = \frac{1}{nN}\frac{2}{SR}\sum_{i=1}^{n} Re({}^iFY_f)Re({}^iFY_f) + Im({}^iFY_f)Im({}^iFY_f)$$

The three co spectra ${}_5C_{xy}$ are constructed analogously (normalising over the number of valid groups n):

$${}_5C_{zx}(f)=\frac{1}{nN}\frac{2}{SR}\sum_{i=1}^{n}Re({}^iFZ_f)Re({}^iFX_f)+Im({}^iFX_f)Im({}^iFZ_f)$$

$${}_5C_{zy}(f)=\frac{1}{nN}\frac{2}{SR}\sum_{i=1}^{n}Re({}^iFZ_f)Re({}^iFY_f)+Im({}^iFY_f)Im({}^iFZ_f)$$

$${}_5C_{xy}(f)=\frac{1}{nN}\frac{2}{SR}\sum_{i=1}^{n}Re({}^iFX_f)Re({}^iFY_f)+Im({}^iFY_f)Im({}^iFX_f)$$

The two quad spectra ${}_5Q_{zx}$ and ${}_5Q_{zy}$ are described by the following formulae (normalising over the number of valid groups n):

$${}_5Q_{zx}(f)=\frac{1}{nN}\frac{2}{SR}\sum_{i=1}^{n}Im({}^iFZ_f)Re({}^iFX_f)-Im({}^iFX_f)Re({}^iFZ_f)$$

$${}_5Q_{zy}(f)=\frac{1}{nN}\frac{2}{SR}\sum_{i=1}^{n}Im({}^iFZ_f)Re({}^iFY_f)-Im({}^iFY_f)Re({}^iFZ_f)$$

Note that the quad spectrum ${}_5Q_{xy}$ is not used.

6.2.3 Correction filters

The necessary correction filters for the spectral height analysis have been described in subsection 5.2.5, but for the spectral direction analysis all or most (depending on the sensor type: Wavec or Directional Waverider) of the auto, co and quad spectra have to be corrected.
The corrections on the various auto, co and quad spectra of the Wavec are described below (Note that ${}_5C_{xx}$, ${}_5C_{yy}$ and ${}_5C_{xy}$ remain unchanged and ${}_5Q_{xy}$ is not used), with the factors H_f, R_f and I_f as described in subsection 5.2.5:

$${}_5C^c_{zz}(f)={}_5C_{zz}(f)H_f$$

$${}_5C^c_{zx}(f)={}_5C_{zx}(f)R_f-{}_5Q_{zx}(f)I_f$$

$${}_5C^c_{zy}(f)={}_5C_{zy}(f)R_f-{}_5Q_{zy}(f)I_f$$

$${}_5Q^c_{zx}(f)={}_5Q_{zx}(f)R_f+{}_5C_{zx}(f)I_f$$

$${}_5Q^c_{zy}(f)={}_5Q_{zy}(f)R_f+{}_5C_{zy}(f)I_f$$

Heave filter correction for the Directional Waverider is accomplished by multiplying all the elements of the 3 auto, 3 co and 2 quad spectra by H_f as described in subsection 5.2.5.
From now on the superscript ^{c} indicating corrected spectra will be dropped.

6.2.4 Smoothed 10 mHz spectra

From the 5 mHz resolution spectra the 10 mHz resolution spectra are obtained analogously to subsection 5.2.6. These 10 mHz spectra will be indicated by the subscript $_{10}$.

6.3 Fourier coefficients

It is impossible to reconstruct the complete directional energy distribution of the waves from the buoy data. However, it is possible to calculate the first four Fourier coefficients of the directional energy distribution, which are used to determine the mean direction, width, asymmetry and flatness of the directional energy distribution [2]. These first 4 Fourier coefficients of the directional energy density spectrum can be estimated directly from the 8 auto, co and quad spectra [5]:

$$A_1(f) = \frac{Q_{zx}(f)}{W(f)C_{zz}(f)}$$

$$A_2(f) = \frac{C_{xx}(f) - C_{yy}(f)}{C_{xx}(f) + C_{yy}(f)}$$

$$B_1(f) = \frac{Q_{zy}(f)}{W(f)C_{zz}(f)}$$

$$B_2(f) = \frac{2C_{xy}(f)}{C_{xx}(f) + C_{yy}(f)}$$

$$W(f) = \sqrt{\frac{C_{xx}(f) + C_{yy}(f)}{C_{zz}(f)}}$$

where $W(f)$ is the wavenumber. Note that the spectra of Fourier coefficients obtained this way have the same resolution as the auto, co and quad spectra from which they are derived (5 or 10 mHz).

6.4 Centred Fourier coefficients

Centred Fourier coefficients [2] are used to calculate the width, skewness and kurtosis of the directional distribution and are defined by:

$$m_1(f) = \sqrt{(A_1(f))^2 + (B_1(f))^2}$$

$$m_2(f) = A_2(f)\cos 2\theta_m(f) + B_2(f)\sin 2\theta_m(f)$$

$$n_1 = 0$$

$$n_2(f) = B_2(f)\cos 2\theta_m(f) - A_2(f)\sin 2\theta_m(f)$$

The coefficients are centred around the direction $\theta_m(f)$, which is *not* the nautical mean direction $\theta_0(f)$ (see subsection 6.5.1). Note that the centred Fourier coefficients exist only if the normal Fourier coefficients (see previous subsection) can be determined.

6.5 Directional spectra

There are two directional spectra constructed during the spectral directional analysis:

1. The mean wave direction spectrum,
2. The directional width spectrum.

Note the difference between these parameters (which are functions of the frequency) and the average mean direction and average width as directional spectrum parameters and as band parameters. Both spectra are calculated from the 10 mHz spectra.

6.5.1 Mean wave direction spectrum

According to [2] the lowest Fourier coefficients define an angle which can be regarded as the mean wave direction. This mean direction spectrum is defined such that a 0° angle indicates east and 90° indicates north, with the waves propagating away from the sensor:

$$\tan\theta_m(f) = \frac{B_1(f)}{A_1(f)}$$

However, according to nautical convention the wave direction is the direction from which the waves propagate towards the sensor, with a 0° angle indicating waves coming from the north and a 90° angle indicating waves from the east.
The nautical mean direction spectrum $\theta_0(f) \in [0°;360°\rangle$ is given by:

$$\tan\theta_0(f) = \frac{-A_1(f)}{-B_1(f)}$$

To get rid of the singularity at $B_1(f) = 0$, the expression can be rewritten using the sine function giving:

$$\sin\theta_0(f) = \frac{-A_1(f)}{\sqrt{(A_1(f))^2 + (B_1(f))^2}}$$

The signs of the Fourier coefficients $A_1(f)$ and $B_1(f)$ determine the interval in which $\theta_0(f)$ can be found (boundaries indicated by square brackets are included, those indicated by angular brackets are excluded). See Table 6.1 and 6.2:

Table 6.1 θ_0(f)-range as determined by the sign of A_1(f)

sign of A_1(f)	range of θ_0(f)	θ_0(f) if B_1(f) is zero
positive	[180°, 360°〉	270°
negative	[0°, 180°〉	90°

Table 6.2 θ_0(f)-range as determined by the sign of B_1(f)

sign of B_1(f)	range of θ_0(f)	θ_0(f) if A_1(f) is zero
positive	[90°, 270°〉	180°
negative	[0°, 90°〉 ∪ [270°, 360°〉	0°

In case both $A_1(f)$ and $B_1(f)$ are zero, $\theta_0(f)$ is set to error code.
The angle thus obtained is defined such that waves coming from an angle of 0º propagate from the magnetic north. In order to compensate for the slight deviation of the magnetic poles from the real poles, a θ_0-correction angle $\delta\theta$ is added to $\theta_0(f)$. The resulting $\theta_0(f)$ is converted to the range of [0°, 360°〉. The magnitude of the correction $\delta\theta$ and its annual variation depends on the location of the sensor and can be found on maps or in tables.

6.5.2 Directional width spectrum

The directional width spectrum $\sigma(f)$ is calculated according to:

$$\sigma(f) = \sqrt{2(1 - m_1(f))}$$

where $m_1(f)$ is the centred Fourier coefficient around the mean wave direction $\theta_m(f)$ as defined in section 6.4. If $m_1(f) \geq 1$ or if $A_1(f)$ or $B_1(f)$ could not be determined $\sigma(f)$ is set to error code. The directional width is obtained in radians and has to be multiplied by a factor $180/\pi$ for conversion to degrees.

6.6 Directional spectrum parameters

There are two different types of directional spectrum parameters:

1. Average mean directions,
2. Average width.

The directional spectrum parameters are calculated from the 10 mHz spectra.

6.6.1 Average mean direction

The directional spectrum parameters include two average mean directions, one in the band [30, 500] mHz and the other in the band [30, 100] mHz. The calculation of the average mean direction is discussed in subsection 6.7.4.

6.6.2 Average width

The average width from the directional spectrum parameters is calculated in the band [30, 500] mHz. The precise definition is given in subsection 6.7.5.

6.7 Band parameters

Over several frequency bands the wave height, the number of degrees of freedom of the energy density, the average mean direction, the average width, the average asymmetry, the average flatness and the average frequency are calculated. Table 6.3 lists those frequency bands, while the calculation of the various band parameters is discussed in the subsections below. The definitions used in the formulae of the band parameters are given in subsection 6.7.1.

Table 6.3 Frequency bands

Band name	Lower bound (mHz)	Upper bound (mHz)	Resolution (mHz)
B0	30	500	5
B1	200	500	5
B2	100	200	5
B3	30	100	5
B4	$(f_{max}-\Delta f)$, $f_{max} \in [30,500]$	$(f_{max}+\Delta f)$, $f_{max} \in [30,500]$	5
G1	30	45	5
G2	45	60	5
G3	60	85	5
G4	85	100	5
G5	100	125	5
G6	125	165	5
G7	165	200	5
G8	200	250	5
G9	250	335	5
G10	335	500	5
-	30	500	10
-	30	100	10

Note that in determining the boundaries of the band B4 f_{max} is defined as the frequency with the largest energy density in the range [30, 500] mHz and Δf is the resolution of the spectra.

6.7.1 Definitions

To calculate the band parameters for the frequency bands defined above the following definitions are used, where Δf is the resolution of the spectra in Hz, f the frequency in Hz, and l and h designate the lower and upper bound of the frequency band respectively. The Fourier coefficients A_1, A_2, B_1 and B_2 are indicated by X_n . A horizontal bar over a parameter denotes that this is an average value:

The n-th moment

$$M_n = \Delta f \left[\frac{1}{2} l^n C_{zz}(l) + \sum_{f=l+\Delta f}^{h-\Delta f} f^n C_{zz}(f) + \frac{1}{2} h^n C_{zz}(h) \right]$$

Total of square of energy:

$$S = (\Delta f)^2 \left[\frac{1}{2} C_{zz}^2(l) + \sum_{f=l+\Delta f}^{h-\Delta f} C_{zz}^2(f) + \frac{1}{2} C_{zz}^2(h) \right]$$

Average Fourier coefficients:

$$\overline{X_n} = \Delta f \, \frac{\frac{1}{2} X_n(l) C_{zz}(l) + \sum_{f=l+\Delta f}^{h-\Delta f} X_n(f) C_{zz}(f) + \frac{1}{2} X_n(h) C_{zz}(h)}{M_0}$$

Average centred Fourier coefficients:

$$\overline{m_1} = \sqrt{\overline{A_1}^2 + \overline{B_1}^2}$$

$$\overline{m_2} = \overline{A_2} \cos 2\theta_m + \overline{B_2} \sin 2\theta_m$$

$$\overline{n_2} = \overline{B_2} \cos 2\theta_m - \overline{A_2} \sin 2\theta_m$$

with θ_m as defined in section 6.5.1.

6.7.2 Wave height

As mentioned in subsection 5.3.3, the wave height per frequency-band is determined by:

$$H = 4\sqrt{M_0}$$

6.7.3 Degrees of freedom of energy density

The degrees of freedom of the energy density per frequency band is defined by:

$$N = \frac{M_0^2}{S} 2n(1 - x)$$

where *n* is the number of valid groups and *x* the tapering fraction as defined in section 3.3.

6.7.4 Average mean direction

The average mean direction of the energy density distribution per frequency band is defined as the angle given by the averaged Fourier coefficients:

$$\tan \overline{\theta_0} = \frac{-\overline{A_1}}{-\overline{B_1}}$$

The interval in which the angle can be found is determined as in subsection 6.5.1. Compensation for the deviation of the magnetic poles from the real poles is done as described in subsection 6.5.1. The resulting angle is converted to the range of [0°, 360°〉.

Note that the angle can not be calculated if both $\overline{A_1}$ and $\overline{B_1}$ equal zero, in which case the angle is set to error code.

6.7.5 Average width

The average width of the energy density distribution is calculated per frequency-band using the average centred Fourier coefficients:

$$\overline{\sigma} = \sqrt{2(1 - \overline{m_1}}$$

If $\overline{A_1}$ or $\overline{B_1}$ cannot be determined or if $\overline{m_1} \geq 1$, the average width is set to error code. Note that the width is given in radians and has to be converted to degrees by multiplication by 180/π.

6.7.6 Average asymmetry

The average skewness or asymmetry per frequency band is defined as:

$$\overline{\gamma} = \frac{-\overline{n}_2}{\left[\frac{1}{2}(1-\overline{m}_2)\right]^{3/2}}$$

If $\overline{A}_1$ or $\overline{B}_1$ cannot be determined or are both equal to zero, or in case $\overline{m}_2$ equals 1, the average skewness is set to error code.
Note that the sign of the skewness indicates the direction in which the asymmetric tail of the distribution extends. The nautical definition of the mean direction implies an inversion of the direction and therefore in the nautical coordinate system the skewness indicates the direction opposite to the direction to which the asymmetric tail points.

6.7.7 Average flatness

The average flatness or kurtosis of the distribution per frequency band is given by:

$$\overline{\delta} = \frac{6 - 8\overline{m}_1 + 2\overline{m}_2}{(2(1-\overline{m}_1))^2}$$

In case $\overline{A}_1$ or $\overline{B}_1$ cannot be determined or are both equal to zero, or in case $\overline{m}_2$ equals 1, the average kurtosis is set to error code.

6.7.8 Average frequency

The average frequency per band is defined as the weighted average over the band:

$$F = \frac{M_1}{M_0}$$

6.8 Calculated parameters

The following parameters are calculated during the spectral direction analysis phase. The units of the calculated parameters are: energy is expressed in cm^2, wave height in cm, frequency in Hz, period in s, direction in ° (degrees)

Directional spectra

- Th010(i), i ∈ [3,50] Spectrum of mean direction $\theta_0(f)$ from 10 mHz spectra
  Th010(i) = $\theta_0(f)$ where f =i*10 mHz, $f \in$ [30,500] mHz
- S0bh10(i), i ∈ [3,50] Spectrum of directional width $\sigma(f)$ from 10 mHz spectra
  S0bh10(i) = $\sigma(f)$ where f = i*10 mHz, $f \in$ [30, 500] mHz

Fourier coefficients

- A1_5(i), i ∈ [0,100] Spectrum of Fourier coefficients $A_1(f)$ from 5 mHz spectra
  A1_5(i)= $A_1(f)$ where f=i*5 mHz, $f \in$ [0, 500]
- A2_5(i), i ∈ [0,100] Spectrum of Fourier coefficients $A_2(f)$ from 5 mHz spectra
  A2_5(i)= $A_2(f)$ where f=i*5 mHz, $f \in$ [0, 500]
- B1_5(i), i ∈ [0,100] Spectrum of Fourier coefficients $B_1(f)$ from 5 mHz spectra
  B1_5(i)= $B_1(f)$ where f=i*5 mHz, $f \in$ [0, 500]
- B2_5(i), i ∈ [0,100] Spectrum of Fourier coefficients $B_2(f)$ from 5 mHz spectra
  B2_5(i)= $B_2(f)$ where f=i*5 mHz, $f \in$ [0, 500]
- W_5(i), i ∈ [0,100] Wavenumber W(f) from 5 mHz spectra
  W_5(i)= $W(f)$ where f=i*5 mHz, $f \in$ [0, 500]

Directional spectrum parameters

• Th0	Average mean direction $\overline{\theta}_0$ from 10 mHz spectra in the range $f \in$ [30, 500] mHz
• S0bh	Average width $\overline{\sigma}$ from 10 mHz spectra in the range $f \in$ [30, 500] mHz
• Th3	Average mean direction $\overline{\theta}_0$ from 10 mHz spectra in the range $f \in$ [30, 100] mHz
• DL_index	Wave number DL_index = $W(f)$ at f=100 mHz from 5 mH spectra

Band parameters (over all bands B0 - B4 and G1 - G10)[2]

• Hm0_..	Wave height H from 5 mHz spectra
• Ndfe_..	Number of degrees of freedom of the energy density N from 5mHz spectra
• Th0_..	Average mean direction $\overline{\theta}_0$ from 5 mHz spectra
• S0bh_..	Average width $\overline{\sigma}$ from 5 mHz spectra
• G1_..	Average asymmetry $\overline{\gamma}$ from 5 mHz spectra
• G2_..	Average flatness $\overline{\delta}$ from 5 mHz spectra
• Fm01_..	Average frequency F from 5 mHz spectra

Quality parameters

• Ndlr_H	Number of valid subseries of the signal in the vertical direction
• Ndlr_R	Number of valid groups of subseries
• AV10_R	Number of degrees of freedom of the directional spectra (= 4 * Ndlr_R)

[2] The two dots in the parameter names stand for the name of the band (see Table 6.3). Taking the Hm0 as an example: 15 Hm0 band parameters exist with names Hm0_B0, Hm0_B1, Hm0_B2, Hm0_B3, Hm0_B4, Hm0_G1, Hm0_G2, Hm0_G3, Hm0_G4, Hm0_G5, Hm0_G6, Hm0_G7, Hm0_G8, Hm0_G9 and Hm0_G10.

7 Output message definition

7.1 Introduction

This Chapter describes the output of the SWAP module. The wave information becomes available at three levels: at level 0 the raw data used mostly for inspection of the sensors and for research is provided, the basic information needed for calculation of the characteristic wave parameters is made available at level 1, while at level 2 the estimated parameters for operational use and long time archiving are given.
The information provided at each level is given in the subsequent sections.

7.2 Level 0

Level 0 provides the raw wave data in 20 minutes blocks, which consist of:

For wave height sensors:

- Samples of vertical signal

For wave height and direction sensors:

- Samples of vertical signal
- Samples of east-west horizontal signal
- Samples of north-south horizontal signal

7.3 Level 1

Level 1 provides the basic information needed for calculating the characteristic parameters, which consists of:

For wave height sensors:

- Czz5(i), $i \in [0,100]$
- Czz5_M(i), $i \in [0,200]$
- WTBH(i), $i \in [1,AG]$
- WTBT(i), $i \in [1,AG]$
- AG

For wave height and direction sensors:

- Czz5(i), $i \in [0,100]$
- WTBH(i), $i \in [1,AG]$
- WTBT(i), $i \in [1,AG]$
- AG
- A1_5(i), $i \in [0,100]$
- A2_5(i), $i \in [0,100]$
- B1_5(i), $i \in [0,100]$
- B2_5(i), $i \in [0,100]$
- W_5(i), $i \in [0,100]$

7.4 Level 2

Level 2 provides the estimated parameters, which are:

For wave height sensors:

- Czz10(i), $i \in [0,50]$
- Czz10_M(i), $i \in [0,100]$
- M0
- Hm0
- M0_M
- Hm0_M
- Tm02
- Tm02_M
- TE0

- TE1
- TE1_M
- TE2
- TE3
- HTE3
- Fp
- Fp_M
- HS7
- Tm-10
- Tm-10_M
- AV10_H
- AG
- Hmax
- H1/50
- H1/10
- H1/3
- GGH
- SPGH
- Tmax
- THmax
- T1/3
- TH1/3
- GGT
- SPGT
- HCM
- Nwt_zP
- Ndlr_H
- Ngd_zP
- Nu_z
- Nv_z
- Nd_z
- Ni_z

For wave height and direction sensors:

- Czz10(i), $i \in [0,50]$
- M0
- Hm0
- Tm02
- TE1
- TE2
- TE3
- HTE3
- Fp
- HS7
- Tm-10
- AV10_H
- AG
- Hmax
- H1/50
- H1/10
- H1/3
- GGH
- SPGH
- Tmax
- THmax
- T1/3
- TH1/3

- GGT
- SPGT
- HCM
- Nwt_zP
- Th010(i), $i \in [3,50]$
- S0bh10(i), $i \in [3,50]$
- Th0
- S0bh
- Th3
- AV10_R
- DL_index
- Ndlr_H
- Ngd_zP
- Nu_z
- Nv_z
- Nd_z
- Ni_z
- Ndlr_R
- Ngd_xP
- Ngd_yP
- Nu_x
- Nv_x
- Nd_x
- Ni_x
- Nu_y
- Nv_y
- Nd_y
- Ni_y

The following parameters are provided for the bands B0-B4 and G1-G10[3]:

- Hm0_..
- Ndfe_..
- Th0_..
- S0bh_..
- G1_..
- G2_..
- Fm01_..

[3] The two dots in the parameter names stand for the name of the band (see Table 6.3). Taking the Hm0 as an example: 15 Hm0 band parameters exist with names Hm0_B0, Hm0_B1, Hm0_B2, Hm0_B3, Hm0_B4, Hm0_G1, Hm0_G2, Hm0_G3, Hm0_G4, Hm0_G5, Hm0_G6, Hm0_G7, Hm0_G8, Hm0_G9 and Hm0_G10.

8 Literature

1. Zwanenburg, Tj., "Onderzoek naar getijfilters ten behoeve van golfonderzoek", div-nr 672, 1988.
2. Kuik, A. J., Van Vledder, G. Ph., and Holthuijsen, L. H., "A method for the routine analysis of pitch-and-roll buoy wave data", Journal of Physical Oceanography, Vol. 18, No. 7, 1988.
3. Benschop, A. W., Hakkesteegt, H. C., and Vroon, C. A., "RMI aanpassings module, SESAM verwerkingsmethodiek en sensoruitgangsspecificaties", part A, TPD-HAI-RPT-93-0001, 1993.
4. Heinen, P. F., "Standaard voor bewerking en opslag van golfgegevens", Rapport DGW-92.008, 1992.
5. Longuet-Higgins, M. S., Cartwright, D. E., Smith, N. D., "Observations of the directional spectrum of sea waves using the motions of a floating buoy", Ocean wave spectra, Prentice-Hall, 1963.
6. Kuik, A. J., "Beschrijving van de methode van verwerking voor Wavec", TPD-rapport, 1981.
7. Datawell B.V., "Manual for the Wavec buoy from serial no. 22039", 1986.
8. Datawell B.V., Hippy sensor documentation for Wavec and Waverider and for Directional Waverider.
9. Bishop, J.M., "Applied Oceanography", John Wiley & Sons, 1984.
10. Fax from KNMI to Rijkswaterstaat 5-Oct-1993.