Ministry of Environment and Food of Denmark
Environmental Protection Agency

Monitoring of Ballast Water Quality

MUDP report

June 2019

Publisher: The Danish Environmental Protection Agency

Photos:
Frank Struer-Lauridsen
Frey Callesen
Kristina Kern-Nielsen
Camilla Hedberg
Pernille Bohn
Artur Mielczarek

ISBN: 978-87-7038-076-8

The Danish Environmental Protection Agency publishes reports and papers about research and development projects within the environmental sector, financed by the Agency. The contents of this publication do not necessarily represent the official views of the Danish Environmental Protection Agency. By publishing this report, the Danish Environmental Protection Agency expresses that the content represents an important contribution to the related discourse on Danish environmental policy.

Contents

1. Introduction

The IMO's Ballast Water Management Convention entered into force in September 2017 and between 50,000 and 70,000 ships are expected to install a ballast water management system before the end of 2024. Already in 2012 LITEHAUZ decided to develop an automated monitor, which could reliably indicate the performance of these systems directly in the ballast pipe and already during ballasting.

In 2012 LITEHAUZ decided to develop an automated monitor in collaboration with Medico Chemical Lab. The instrument should reliably indicate the performance of a ship's system for treating ballast water under the new Convention and do so directly in the ballast pipe already during ballast water uptake without addition of chemical or interaction from crew.

The bw monitor system[1]
A ballast water treatment system typically includes a unit for removing larger (organic) particles and a disinfection unit for the residual. The separation unit is most often a filter with a mesh size yielding a cut off of 20-50 um and the two major types of disinfection units rely on UV irradiation or producing reactive chlorine. The bw monitor comprises two sensors: one for monitoring particles and hence the performance of a filter and one sensor measuring the fluorescence of chlorophyll, i.e. an indication of viable algae and the efficacy of the disinfection process. The bw monitor operates without chemical addition or extracted sample manipulation in the ballast line (*in-line*) and provides operator or master instant (*on-time*) analysis and results. It can measure and report on pre-treatment performance and the disinfection unit's efficacy in due time for a re-run, repair or retreatment. It gives an instant alarm in case of malfunction and can warn the operator instantly. The bw monitor sensors are located in two positions: before and after pre-treatment and disinfection units.

The bw monitor displays and stores data onboard allowing a vessel master or shipowner to use the data for performance evaluation of the ballast water management system (BWMS). When data becomes abundant this will allow for the development of a predictive maintenance capacity, which also is of profound interest to the BWMS manufacturer. Obviously, the Port State Control that are to inspect the BWMS compliance can use the data to inspect that no gross exceedance of viability criteria takes place. A cloud-based data hub facility is also offered as an option allowing both a backup mechanism and the combined analysis of the BWMS performance over time, of analyses of several identical system's performance or of different systems under similar conditions.

Timeline
The project AitBall was granted under the Ecoinnovation programme in December 2012 to support the development of the monitor for ballast water with a special focus on the data management part. It was quickly realised that the concept prototype monitor required some further development in 2013 before it was ready for the AitBall project and the efforts were postponed only to be picked up again in late 2013 and carried on into early 2015. During these first 18 months the primary activities included methodology and technology determination (Phase 1)

[1] The system was named AitBall in the original application and was later renamed to **bw monitor**.

and specification of the parameters to be used and analysed in the instrument software (Phase 2). This resulted in the application of a patent from the project on the in-line ballast water monitor system with algorithms developed partly in Phase 3. The patent costs were relegated from the consultancy efforts on machine learning, for which insufficient data was available. At this point it was clear that generation and analysis of data would be key to the success of the project. Unfortunately, the hardware, which was financed under another funding mechanism (Markedsmodningsfonden), turned out not sufficiently robust for ships and a complete redesign of the system took place in 2015 and early 2016. This system, now called bw monitor, was market ready and began to generate data in late 2017, and thus the AitBall big data project was rejuvenated. For the major part of 2018 the big data efforts have been carried out: designing the architecture, the data flow and management, and describing the algorithms to handle our own data and Third Party data. The final two Phases 4 and 5 are now complete and the AitBall project is therefore finalised in February 2019.

The measurement technology
The bw-monitor utilizes two key performance parameters: particle density to assess filter retention, and the viability of microalgae as an indicator of disinfection performance. The parameters are measured with optical techniques, i.e. viable algae by fluorescence and particles by laser scattering. The response for each parameter may be logged in real time. i.e. in milliseconds. There are no removal or manipulation of samples nor any addition of chemicals.

Fluorescence
The use of active fluorescence to measure photosynthetic activity has been a generally accepted method for several years. Most of these methods use a flash stimulated fluorescence where it is possible to measure photosynthetic activity using variable fluorescence.
Variable fluorescence Fv is measured as the difference between a low level, steady state fluorescence F0 and the level of a fluorescent transient Fmax. The difference between the two levels is a measure of the functionality of the photosynthetic system of the algae.

The content of living phytoplankton is determined by calculating the variable fluorescence

$$Fv = (Fmax – F0).$$

The minimal fluorescence is determined by using a pulsating light source with preferably 20 ms to 100 ms intervals and a wavelength of 420 nm. To reach the state for determination of the minimal fluorescence faster, a light source with a wavelength longer than 700 nm is used. The maximal fluorescence is determined by using continuous light source with a wavelength of 680 nm.

Laser Scattering
As the determination of the particle size and particle concentration is of great importance in the monitoring of the ballast water compliance, the apparatus is equipped with a light scattering unit. The light scattering unit is a pulsating laser module generating pulsating light with a wavelength of 650 nm. The laser light transmitted, and the diffraction is measured at three angles. The light scattering device is calibrated with priority to a reference particle size standard that allows determination of particle sizes in the range of 10–100 µm. The light scattering measurement is included in the continuously measuring cycle.

2. Monitoring system

The bw-monitor system comprises several tailor-made parts that are designed in order to monitor, analyse and display the monitor status. **Fejl! Henvisningskilde ikke fundet.** shows the arrangement of the following system components:

- Data-HUB – a HMI hub / control place
- Two bw-monitor sensors
- I/O interface – input output interface that gathers data from different sensors installed on board (i.e. temperature, TRO, salinity, geoposition, BWMS operation)

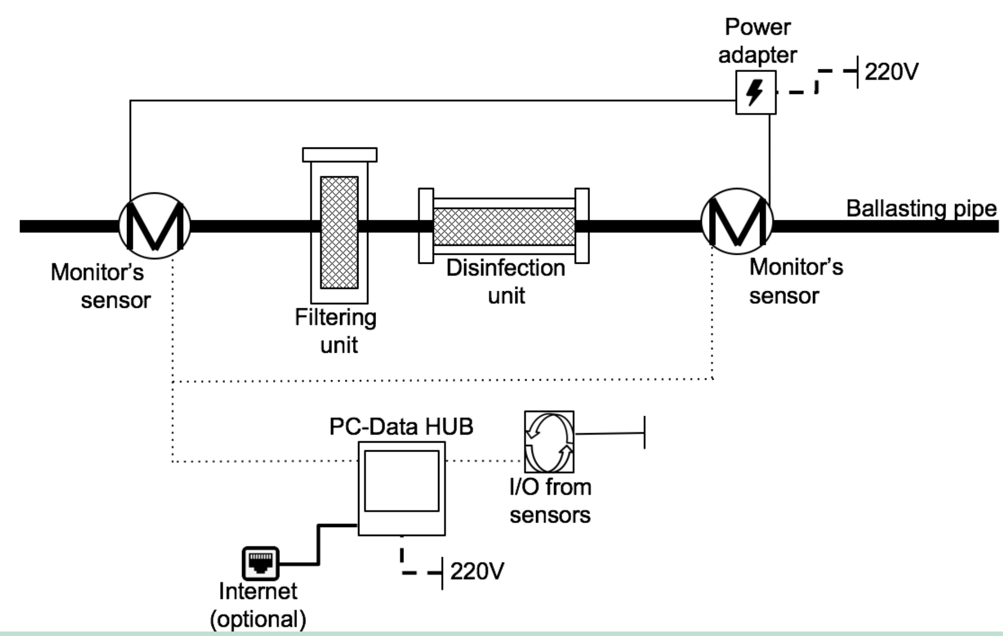


FIGURE 1. Schematic installation of bw-monitor system

2.1 Hardware

The bw–monitor is based on well-known and robust technologies – light scattering and fluorescence. Viable organisms are detected according to a fluorescence signal from chlorophyll, which is detectable directly in the ballast water stream and an indication of the presence of particles, including larger organisms, is provided by a laser diffraction signal related to the density of particles.

The bw-monitor is designed as a two-legged device with the legs extending into the ballast pipe. One leg contains sources of light emission (both LED diodes and laser module) and the other measures the light scattering and holds LED diodes. Perpendicular to the light path, the fluorescence signal is detected in a silicon photodiode provided with a long pass filter. The laser light transmitted and the diffraction is measured at 0, 5 and 25 degree angles referred to as “center”, “inner” and “outer”, respectively.

The ballast water is exposed for three sequences of UV light emissions and one sequence of laser scattering detection of particles. Based on the emission of fluorescence light from the algae in each sequence, the amount of dead and living algae is calculated and based on the result of the laser scattering, the difference in density of particles is determined.

The bw-monitor comes with custom-made software that is installed on a PC-Data HUB. The HUB collects and analyses data incoming from the two connected sensor devices. The HUB is a rugged laptop or (touch-display) rugged all-in-one PC.

FIGURE 2. bw-monitor unit with two sensors (small and large orifice). The two legs are immersed in ballast water and a standard DIN flange is included for installing in ballast pipe.

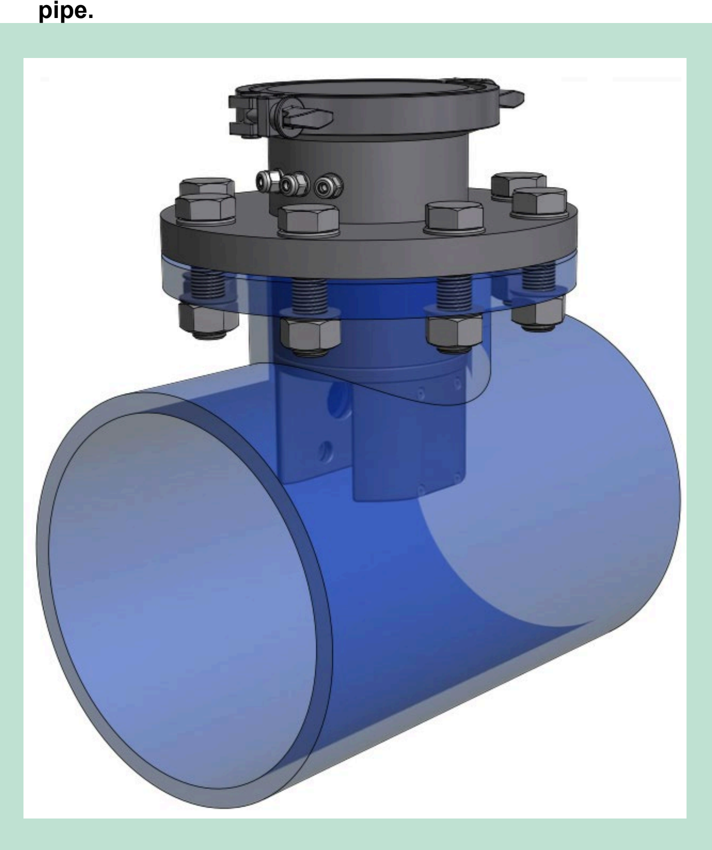


FIGURE 3. The bw-monitor's sensor mounted in the ballast pipe.

2.2 Software

A specially designed .NET-based application, which is installed on the PC connected to bw-monitor, has been created in order to control, gather and analyse incoming data from bw-monitor.

The software has two modes of operation:

- Administration mode – gathering data from the device, executing the algorithm and local storage.
- User mode – user interface displays the information of the quality of the treatment.
    a. It is possible to select a simple display of an indicative fail/ok status for filter and treatment (according to previously assigned error values in the algorithm.
    b. It is also possible to present a real time display of data from the sensors as the data are collected.

The bw-monitor saves all logs locally on the PC-Data-HUB and the user is receiving the status in a form of tamper-free report in pdf/csv file. The software is also pre-programmed (optionally) to transfer an encrypted data log from all past operations to compatible remote cloud storage, when an internet connection becomes available.

The software's inner algorithm has been designed in order to control and manage the bw-monitor and to analyse incoming data in "real time". The algorithm comprises of cascade step-wise check of the incoming values from the sensors. After the evaluation user is instantly informed whether treatment of ballast water is necessary and if yes, whether there is a problem with disinfection or/and filter unit during operation.

The algorithm is designed to self-adjust its critical values (error level) by automatically altering the functional settings (i.e. power of diodes) of the bw-monitor sensor itself. The critical values, which are used in order to determine the BWMS malfunction, are derived from the specifications of the BWMS separation and disinfection units, our empirical laboratory values and on standard curves resembling broad scope of algae and particles concentrations and characteristics.

The conceptual model of an algorithm is depicted in the following graph.

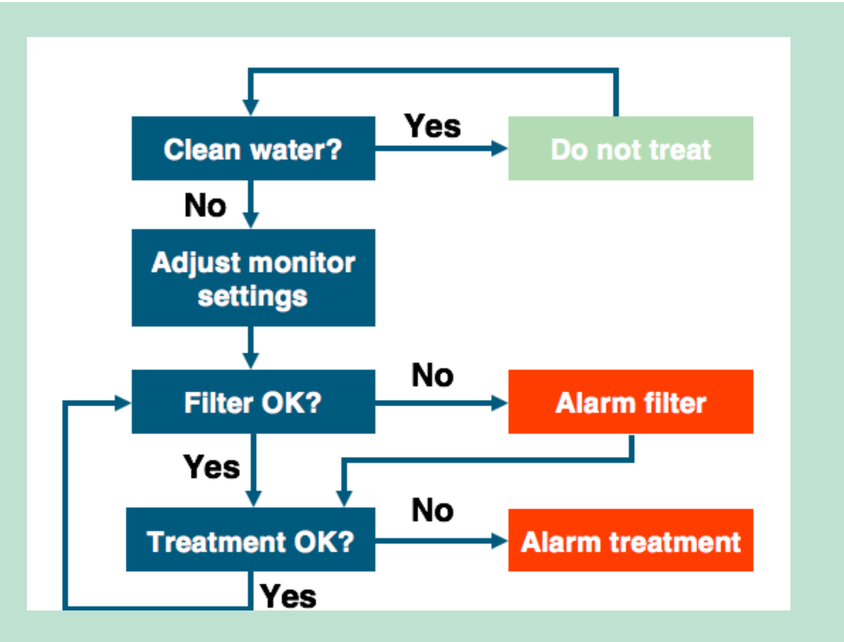


FIGURE 4. Algorithm conceptual design

The bw-monitor is in principle capable of controlling a BWMS as shown in **Fejl! Henvisningskilde ikke fundet.**. In a first step of algorithm's cascade facilitates check of necessity of the treatment and is not used in the current version as it will require bw-monitor to be part of the control mechanism of the type approved BWMS. This step will eventually ensure that ballast water treatment system is switched only when needed, thus leading to energy savings.

In the present version, the second step allows bw-monitor to adjust its sensors response level to properties of water flowing through ballasting system increasing sensitivity in clear water and reducing it in high solids/chlorophyll water. After (first and) second step a repeating loop of analysis of disinfection and filter unit is activated.

In its simplest form the intuitive user interface (UI) of the software allows the BWMS operator to assess in the "real time" performance of the ballast water management system and to take an immediate action in case of malfunction. A sample of UI is depicted below.

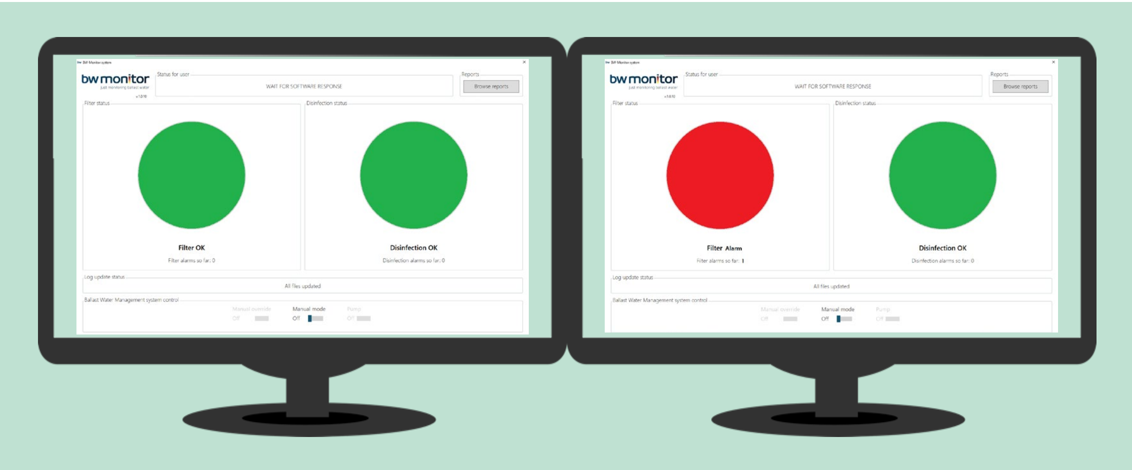


FIGURE 5. bw-monitor's onboard software

There are two different types of output that are produced after completion of the ballasting operation. The bw-monitor software logs, every 1.5 sec., data incoming from the sensors and saves them as raw values in the csv format. These data include: date and time of operation, readout values from fluorescence sensor (Fmax, F0), readout values from laser scattering sensor (Scat inner and outer), power of the diodes. These data are saved on the local hard drive and cannot be altered by the user.

	A	B	C	D	E	F	G	H	I	J	K	L	M
4	02/26/14 10:34:39.842	42	616	139	0	0	0	67	23	325	50	44	0.66
5	02/26/14 10:34:41.839	40	678	136	0	0	0	67	23	325	0	44	0.66
6	02/26/14 10:34:48.63	140	333	183	0	0	0	67	17	325	0	50	0.75
7	02/26/14 10:34:56.331	141	324	200	0	0	0	65	17	325	0	48	0.74
8	02/26/14 10:35:02.556	141	324	200	0	0	0	65	17	325	20	48	0.74
9	02/26/14 10:35:04.631	137	351	234	0	0	0	65	17	325	0	48	0.74
10	02/26/14 10:35:10.855	137	351	234	0	0	0	65	17	325	20	48	0.74
11	02/26/14 10:35:12.930	138	348	269	0	0	0	64	16	325	0	48	0.75
12	02/26/14 10:35:19.154	138	348	269	0	0	0	64	17	325	50	47	0.73
13	02/26/14 10:35:21.213	143	369	312	0	0	0	64	17	325	0	47	0.73
14	02/26/14 10:35:27.438	147	361	287	0	0	0	64	17	325	0	47	0.73
15	02/26/14 10:35:35.737	144	380	350	0	0	0	63	17	325	0	46	0.73
16	02/26/14 10:35:41.962	144	380	350	0	0	0	63	17	325	20	46	0.73
17	02/26/14 10:35:44.36	37	654	104	0	0	0	48	12	325	0	36	0.75
18	02/26/14 10:35:50.261	37	654	104	0	0	0	48	12	325	20	36	0.75
19	02/26/14 10:35:52.336	136	284	81	0	0	0	58	19	325	0	39	0.67
20	02/26/14 10:35:58.560	136	284	81	0	0	0	64	16	325	50	48	0.75
21	02/26/14 10:36:00.650	199	205	70	0	0	0	64	16	325	0	48	0.75
22	02/26/14 10:36:06.875	212	199	72	0	0	0	58	16	325	0	42	0.72
23	02/26/14 10:36:15.174	218	211	98	0	0	0	56	15	325	0	41	0.73
24	02/26/14 10:36:23.458	233	210	88	0	0	0	55	15	325	0	40	0.73
25	02/26/14 10:36:29.698	233	210	88	0	0	0	55	15	325	20	40	0.73
26	02/26/14 10:36:31.757	240	251	148	0	0	0	54	15	325	0	39	0.72
27	02/26/14 10:36:37.966	240	251	148	0	0	0	55	15	325	50	40	0.73
28	02/26/14 10:36:40.40	261	251	105	0	0	0	55	15	325	0	40	0.73
29	02/26/14 10:36:46.265	261	251	105	0	0	0	55	15	325	50	40	0.73
30	02/26/14 10:36:48.355	273	251	143	0	0	0	55	15	325	0	40	0.73
31	02/26/14 10:36:54.580	275	311	168	0	0	0	228	60	325	0	168	0.74
32	02/26/14 10:37:02.879	319	231	94	0	0	0	53	14	325	0	39	0.74
33	02/26/14 10:37:09.103	319	231	94	0	0	0	53	14	325	20	39	0.74
34	02/26/14 10:37:11.178	367	210	76	0	0	0	52	14	325	0	38	0.73
35	02/26/14 10:37:17.387	367	210	76	0	0	0	53	15	325	50	38	0.72
36	02/26/14 10:37:19.462	407	216	93	0	0	0	53	15	325	0	38	0.72
37	02/26/14 10:37:25.686	407	216	93	0	0	0	51	14	325	50	37	0.73
38	02/26/14 10:37:27.761	457	204	110	0	0	0	51	14	325	0	37	0.73
39	02/26/14 10:37:33.970	424	287	180	0	0	0	52	15	325	0	37	0.71
40	02/26/14 10:37:42.269	475	227	110	0	0	0	51	14	325	0	37	0.73
41	02/26/14 10:37:48.493	475	227	110	0	0	0	51	14	325	20	37	0.73
42	02/26/14 10:37:50.553	478	207	85	0	0	0	50	14	325	0	36	0.72
43	02/26/14 10:37:56.777	478	207	85	0	0	0	49	14	325	50	35	0.71
44	02/26/14 10:37:58.852	485	178	89	0	0	0	49	14	325	0	35	0.71
45	02/26/14 10:38:05.76	485	178	89	0	0	0	49	14	325	50	35	0.71
46	02/26/14 10:38:07.151	487	179	88	0	0	0	49	14	325	0	35	0.71
47	02/26/14 10:38:13.344	484	179	87	0	0	0	48	14	325	0	34	0.71
48	02/26/14 10:38:21.643	465	204	126	0	0	0	48	14	325	0	34	0.71
49	02/26/14 10:38:27.868	465	204	126	0	0	0	48	14	325	20	34	0.71
50	02/26/14 10:38:29.927	494	167	62	0	0	0	46	14	325	0	32	0.70
51	02/26/14 10:38:36.151	494	167	62	0	0	0	46	14	325	50	32	0.70
52	02/26/14 10:38:38.226	495	164	68	0	0	0	46	14	325	0	32	0.70
53	02/26/14 10:38:44.466	495	164	68	0	0	0	59	15	325	50	44	0.75
54	02/26/14 10:38:46.541	349	392	386	0	0	0	59	15	325	0	44	0.75

FIGURE 6. Log of raw data in the csv format

For each ballasting procedure a rapport of compliance will be prepared in bw-monitor software. The abovementioned data from the log are compiled and analysed by the algorithm. The captain or operator receives a pdf file in which the operation summary is presented. Each report will include the date of an operation, number of warnings (incompliance with given cut-off values) and a graphical illustration summarizing events from the log. In **Fejl! Henvisningskilde ikke fundet.** to **Fejl! Henvisningskilde ikke fundet.** examples of graphical illustrations of operations are displayed.

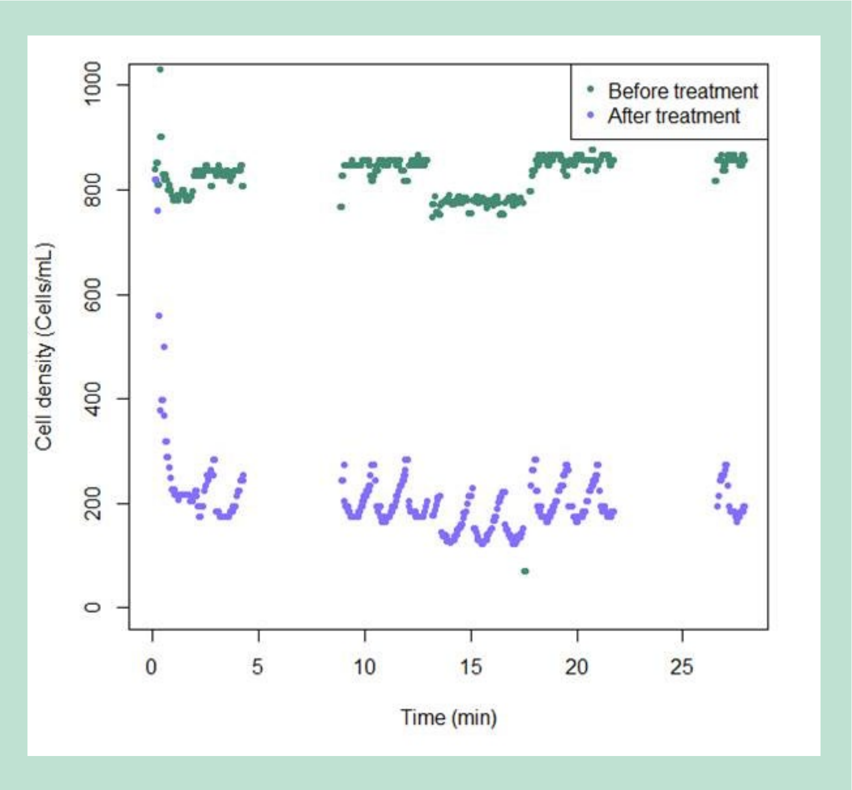


FIGURE 7. Example of cell density graph from a ballasting operation. Shown is the cell density measured over time.

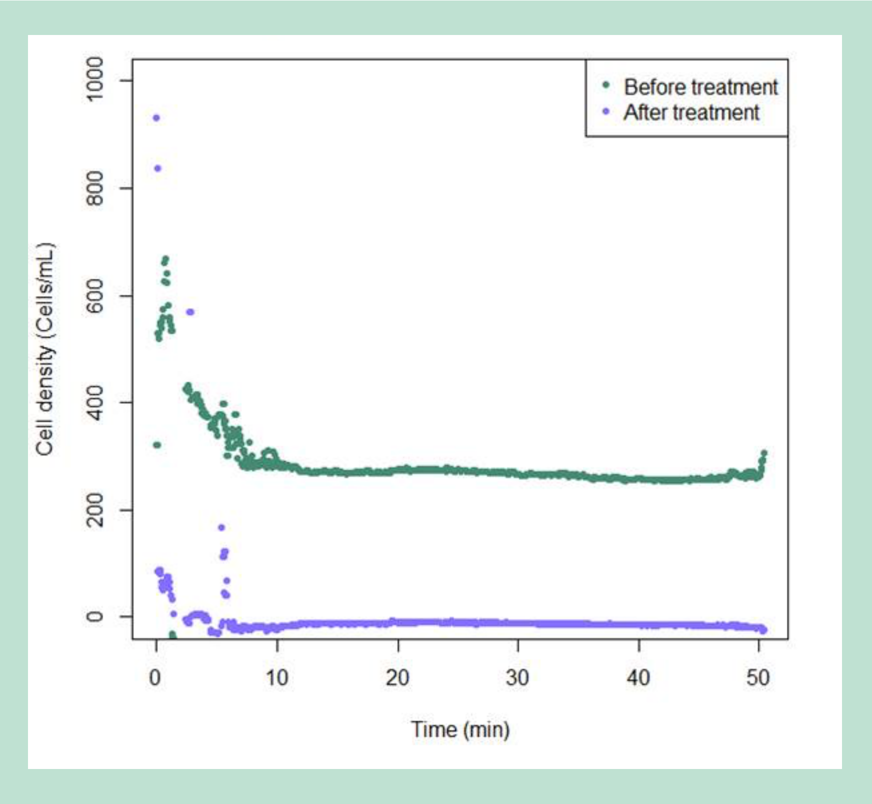


FIGURE 8. Example of cell density graph from a deballasting operation. Shown is the cell density measured over time

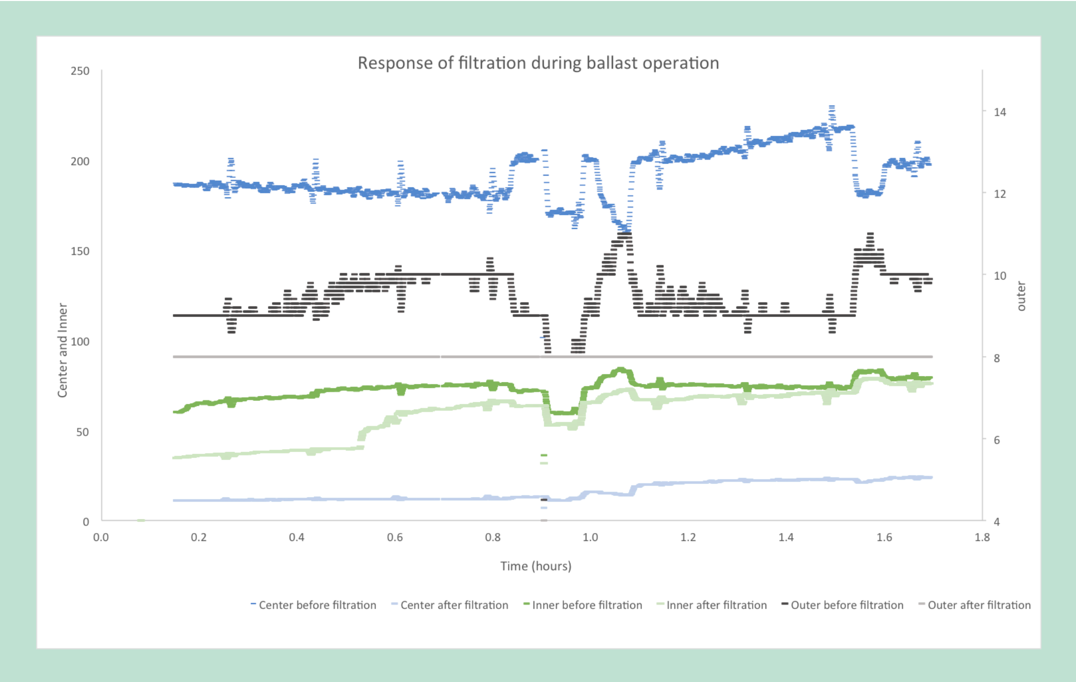


FIGURE 9. Example of particle measurements during ballast operation with a filter present. Shown is the response of particle measurements over time for three particle ranges. The dark colours display the response before filtration and the light colours the response after filtration. A high response indicates a particle rich water while a low response is an expression for a particle poor water. Note the difference in scale.

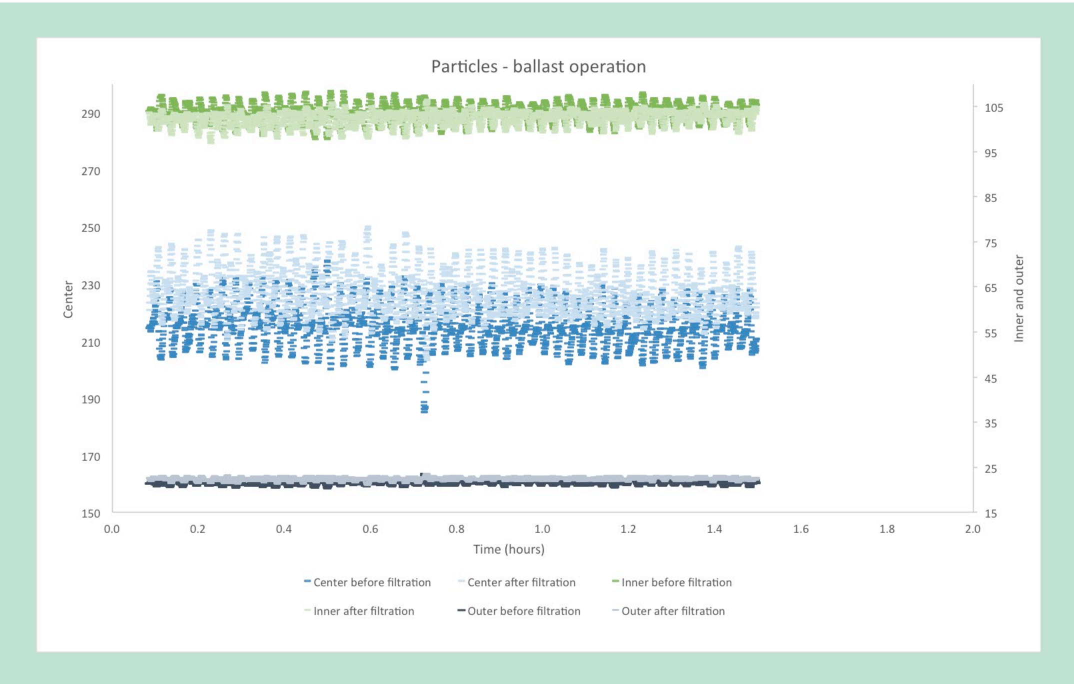


FIGURE 10. Example of particle measurements during ballast operation without a filter. Shown is the response of particle measurements over time. The dark and light colors show the response from before and after measurements. Since no filter is engaged the before and after response is the same for each particle range.

2.3 Autocalibration and service mode

The bw-monitor has been designed to employ optical detection techniques as measurement methods. Within a certain operation time, it is expected that biofouling or coverage of the optical parts (i.e. sensors) with impurities will occur resulting in the decrease of response capabilities of the sensor.

In order to inform the user or ballast water operator about the necessity of service, before each run the series of measurements are taken and compared with predefined default settings. If the signal/read out value that is sent from the sensor is under the preconfigured value user will be instantly notified. During service mode the bw-monitor software will indicate the status as depicted in the picture below.

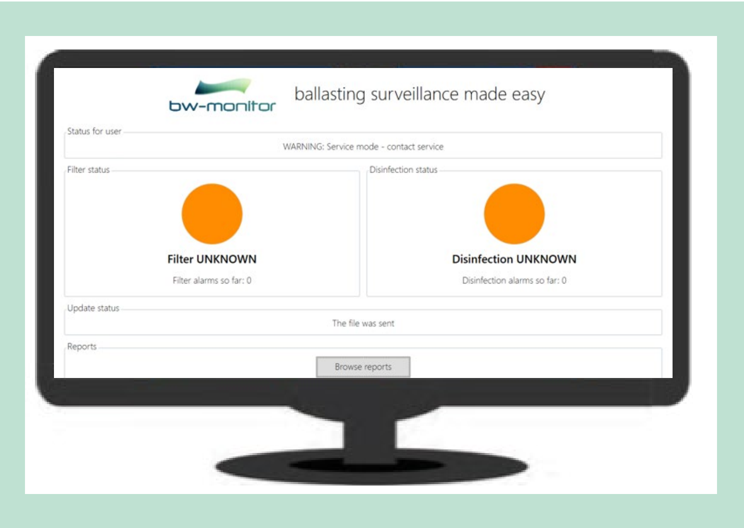


FIGURE 11. Service mode

The maintenance comprises demounting the bw-monitor sensor unit, rinsing it with dilute acid and cleaning it with detergent. A kit containing liquids and tools to clean the sensor windows is supplied. The frequency and effort required depends on the Marine Growth Prevention System (MGPS) and other characteristics of the BWMS and the ship but may range from 3-9 months.

2.4 Auxiliary measuring equipment

The bw-monitor's software handles data incoming from the sensors installed around the BWMS. Additionally, the software is prepared for collecting and logging relevant operational data together with data from other measuring devices related with the ship's operation and ballasting processes.

Depending on availability there are two types of data that may be logged:

1. Ship's activity related data
    a. Vessel description (i.e. load, systems, tonnage)
    b. Fuel consumption
    c. Geoposition and voyage route
    d. Weather, tidal and currents information
    e. Other

2. Ballast Water Management data
    a. Ballast Water System information
    b. Other sensors related to BWMS operation (i.e. total residual oxidant (TRO))
    c. Water temperature
    d. Water salinity
    e. Other

The data that are accessible to the bw-monitor system will be stored encrypted locally on the ships bw-monitor Data-HUB - and when internet connection becomes available - sent to remote cloud storage for further data analysis.

Each BWMS has its particular features and notable differences, thus algorithm custom fitting and update is used to improve the functionality and accuracy of the bw-monitor. In order to implement the future changes in the software and algorithm parameters, the bw-monitor cloud software (further described in section 3) has been designed to remotely access Data-HUBs on the vessels and without user interaction in administration mode rejuvenate the system to reflect changes in the software design.

3. The Data Management

The BW Monitor Azure Cloud system is designed to be fully compatible with bw monitor software. It has two main purposes:

- Database – cloud repository is a secure, encrypted server storage that collects all data from all devices.
- Algorithm update mode – the administrator can remotely update the algorithm with "push-updates". It removes the necessity of user interaction.

The cloud-based system is a central storage place of all data transferred from all users/ devices. The collection of the data will enable pattern recognition and improvements of the algorithm for future software updates. All data that is stored on the BW Monitor Azure Cloud is encrypted. The BW Monitor Azure Cloud is a dedicated server that also can be used to service for remote maintenance of bw monitors installed on the vessels.

3.1 Backup function

Every bw monitor system that is connected to the internet will be offered a remote encrypted storage and access to all of the logged data on bw monitor cloud servers. The feature will provide the shipowner secure storage in case of local Data-HUB malfunction (i.e. hard drive breakdown).

At any time, the vessel's captain or a bw monitor operator will have access to all reports created during ballasting operations in the past.

3.2 Data security

The encrypted data stored on the bw-monitor's cloud can be accessed only by certified staff since the logged data collected on the cloud server are considered sensitive information. In order to minimize risk of a security breach data can only be decrypted locally with the decryption program.

Only after agreement with ship owners, data log analysis can be accessible through the dedicated servers.

4. Big data analysis

Data used to evaluate the quality of ballast water comprise monitor measurements, meta-data, and supplemental data as described below. Currently, only monitor measurements and limited meta-data are available. Long-term, other data should be utilized for valuable analytic purposes. The analysis will allow intuitive access of the vessel's performance and pattern identification.

Solutions to the following three requirements were implemented:

1. How can data be efficiently stored and accessed for data exploration purposes?
2. How can third-party databases such as National Oceanic and Atmospheric Administration (NOAA) World Ocean Database and National Aeronautics and Space Administration (NASA) be integrated and incorporated in statistical analysis?
3. How are the data best accessed, analysed and displayed in a way that does not require extensive technical knowledge of the user?

4.1 System Architecture & Data Flow

The data management system includes:

- BW Mate™ which consolidates measurements from BW Monitor™ with meta-data and generated information such as warnings from the PLC/HMI. It saves the data onboard for local analysis and provides facilities to send them BW Master™.
- BW Master™ which consolidates data from all treatment cycles and supplements them with relevant environmental data from 3rd parties. The collective data can then be used for performance analysis, predictive maintenance, etc.

The flow of data from monitor to database is depicted in **Fejl! Henvisningskilde ikke fundet.**

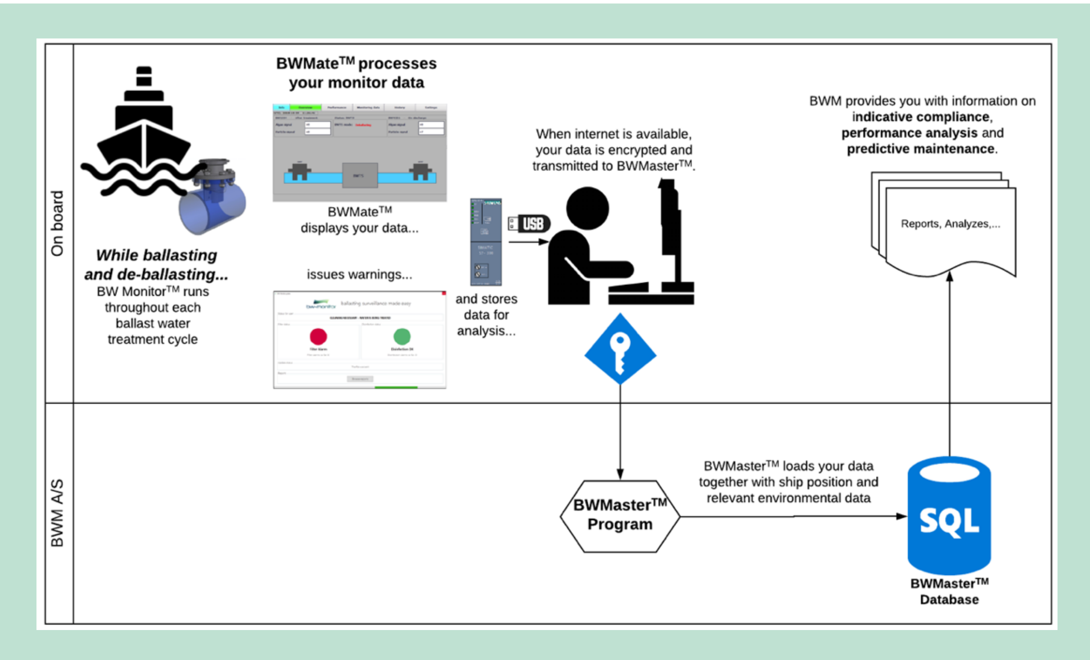


FIGURE 12. BW Monitor™ data flow depiction

Several systems and devices work together to ensure the flow of data from the BW Monitor™ to BW Mate™ and BW Master™. Data flow to and through the various elements is described in Table 1.

Table 1: BW Monitor™ data systems

Device/ System/ Location	**PC version**
BW Mate™ PC on-board	*.bwmonitor csv-file *Each time the monitor is started, the original file (source data) is copied to a uniquely named file, and the original file is overwritten with the current measurements.*
BW Monitor™ Azure Cloud	*.bwmonitor files are encrypted and copied to the cloud. They are accessed by authorized personal only.
Bluebrick Cloud database	A log of all encrypted *.bwmonitor files is maintained. BW Master™ checks daily for new files.
BW Master™	BW Master™ system is comprised of • BW Master™ program: decrypts *.bwmonitor csv-files, checks and loads data into the BW Master™ database, and finds and downloads supplemental environmental data from 3rd party sources. • BW Master™ database: consolidates monitor measurements, meta-data, generated data, and 3rd party environmental data. • (Tableau presentations based on data in BW Master™ database.)
3rd party web servers	Supplemental data are downloaded from various sources via the BW Master™ program.

4.2 On-board PC data / Bluebrick database

Measurements of the monitor system are stored on-site, i.e., on the monitor system's PC onboard ship. They are encrypted and uploaded to the BW Monitor Azure Cloud as *.bwmonitor files, where each file corresponds to exactly one recording (cycle) during ballasting or deballasting. Uploaded *.bwmonitor files are logged in the BlueBrick database. All original *.bwmaster files remain on the monitor's PC in unencrypted csv format. All encrypted *.bwmaster files remain on the BW Monitor Azure Cloud.

4.3 BWMaster™ system

The BWMaster™ program runs automatically at a specified fixed interval – initially once a day at 00:00. The interval and time of BWMaster's execution can be modified anytime.

BWMaster™ checks for new data by comparing file names recorded in the Bluebrick Database with those in the ImportLog table of the BWMaster™ database. When new data is detected, BWMaster™ decrypts the respective *.bwmonitor file, parses the data, and imports them into the BWMaster™ database.

After successful import of monitor data, BWMaster™ imports supplemental data corresponding to the time and location of the imported monitor data. If supplemental data cannot be accessed or loaded, BWMaster™ keeps trying during subsequent run(s) until the supplemental data are successfully loaded. (See Supplemental data for more information.)

Successful imports of monitor data and supplemental data are logged in the ImportLog and Enrichlog, respectively. Unsuccessful attempts to import monitor or supplemental data are logged in the ErrorLog table.

BW Master™ processes are illustrated in **Fejl! Henvisningskilde ikke fundet.** and described thereafter in detail.

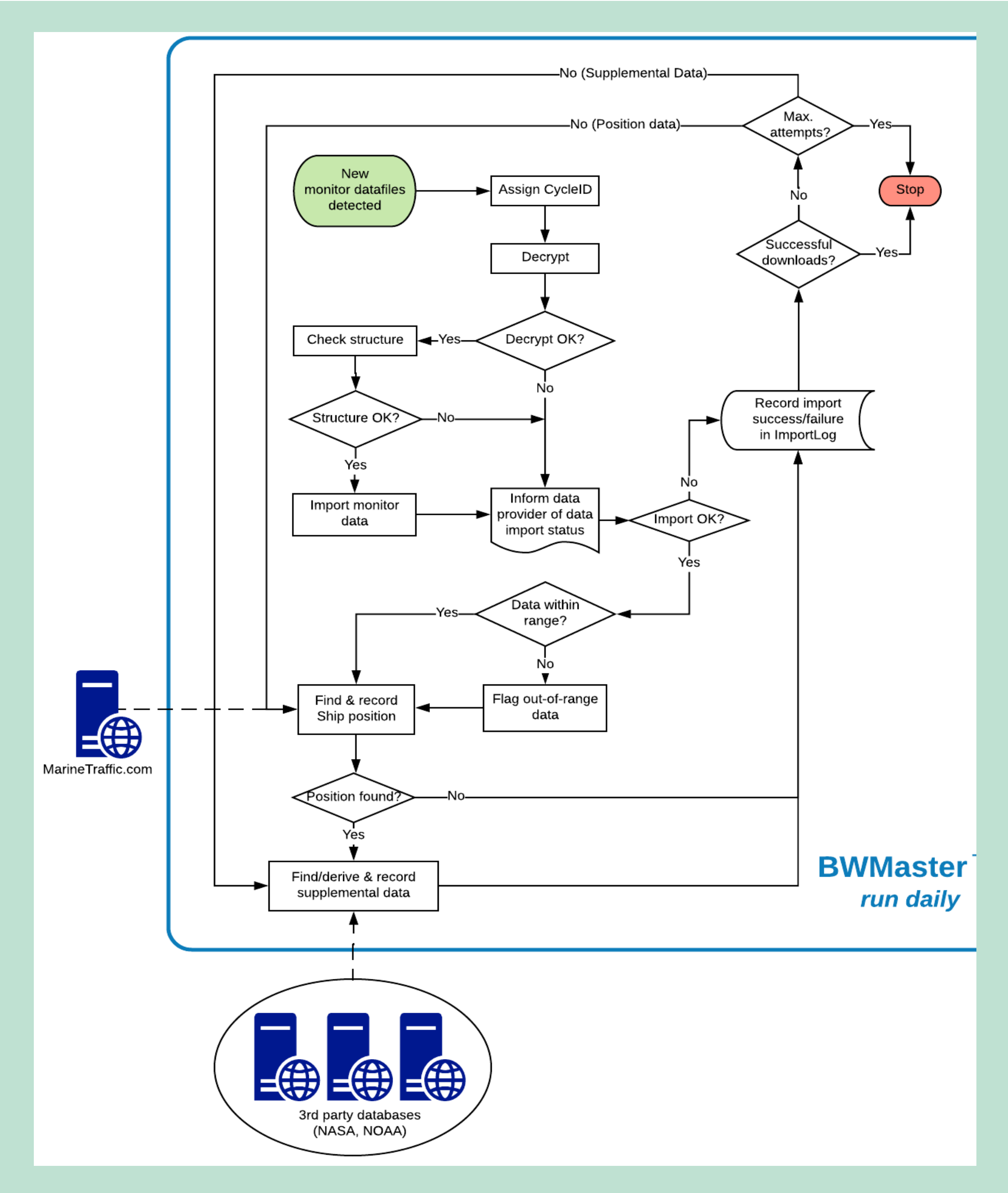


FIGURE 13. BW Master™ processes

4.3.1 Decryption

The encryption algorithm conforms to the AES (Advanced Encryption Standard) specification using the Microsoft C# implementation of the Rijndael block cipher algorithm. In order to decrypt a *.bwmonitor file, one must possess two secrets, namely, the decryption key and the salt used in encryption.

4.3.2 Import

After decryption, the names of the columns comprising the first row of the *.bwmonitor file are checked for conformity to the fixed structure required for import to the BWMaster database. If they do not conform, import is aborted, and the failure is recorded in the ErrorLog table.

If the *.bwmonitor file conforms to import requirements, a unique sequential monitoring cycle number[2] is assigned to it and data are imported row-by-row together with the monitoring cycle number.

All aborted imports are recorded in the ErrorLog table.

4.4 Database size

A conservative (maximum) estimate of data quantity is based on the following:

- The average number of lines generated (rows) in a cycle is estimated from all cycles of the year 2018 to be 565.
- The vast majority of data resides in the Measurement table (~ 52 bytes per row) and the Criteria table (~88 bytes per row). One row of each of these tables is inserted for each row in the *.bwmonitor files.
- On average, a ship produces 4 cycles per month.
- There are approximately 50000 ships worldwide.

The maximum amount of data expected is thus:

  50,000 ships per year
* 4 cycles per month * 12 months per year
* 565 rows per cycle
* (52 + 88) bytes per row
= approximately 190 gigabytes per year.

Even if ships produce 10 or 20 times this load, it is still a very manageable amount of data.

4.5 3rd Party APIs

Organizing, storing and analyzing monitor data is of course in itself valuable for performance monitoring and may help answer questions such as: Is there a correlation between geoposition and warning rate? Are some types of warnings usually followed by other types of warning? etc. However, the overall goal is to develop a predictive maintenance system in order to, e.g., optimize replacement of filters so that they can be ordered ahead of time and thereby minimize downtime for ship owners. To this end, it is hypothesized that external factors which influence seawater content (such as sea surface temperature) have predictive value.

4.5.1 Data Sources and Access

Many such external factors comprise data collected globally from satellites and made available online by governmental organizations such as NASA and NOAA. Access to earth scientific databases have been standardized internationally and can, in most cases, be accessed via DAP (Data Access Protocol) which has been implemented in many programming languages including Python.

NASA: OPeNDAP (Open-source Project for a Network Data Access Protocol)
NOAA: ERDDAP (Environmental Research Division's Data Access Protocol)

4.5.2 Data Resolution

Normally, ballasting and deballasting are done in port. Occasionally, however, freight is transferred between ships and ballast water must be adjusted at sea. Satellite data, although global, does not always contain data from inside the port area.

[2] Monitoring cycle number is derived from BWMaster's Cycle table = maximum Cycle.CycleID + 1.

Temporal and spatial resolution varies widely but is commonly on the order of:

- 0.05x0.05 to 0.5x0.5 degrees longitude/latitude (0.5 degree is roughly equal to 50 km)
- 1 to 8 days temporal resolution.

Some quantities such as surface temperature may differ significantly between the port area and a few kilometers out at sea, thus requiring a high resolution for accuracy.

Datasets were selected to accommodate as best possible, namely, with highest possible spatial and temporal resolution as required by the respective features while still having adequate coverage in harbour areas. Regardless of the resolution, data points may not always be available for a given ship position at a given time, so the mean of a grid surrounding the respective data point is recorded.

L2 data are near real time (high temporal resolution) measured in the immediate region of a satellite's orbital path. L2 datasets often have low spatial coverage as they take several days to cover the whole globe.

L3-L4 data often have both high spatial resolution (by combining multiple measurements) and high coverage because of interpolation and prediction analysis. However, they typically have lower temporal resolution than L2 data.

L3 and L4 data are thus the best choice for values with little temporal variation, for example, sea surface temperature and salinity. For more highly temporally variable data such as pole tides with their high hourly variation, near real time data (L2) is a better choice.

4.6 Supplemental data

Supplemental data are published at various intervals: some data suppliers provide real time data whereas others may publish data daily, weekly, monthly or even less frequently.

Upon importing monitor data, BW Master™ attempts to download all relevant supplement data. BW Master™ also re-attempts to download any missing supplemental data once per day until downloading is successful or the pre-set number of attempt days is reached. The pre-set number of attempt days is specified separately for each supplemental data type in the DataTypes table. It is based on the data publishing policy of the respective data provider.

The first attempt to import supplemental data – whether successful or not – is logged via a new record in the ImportLog, and thus the ImportLog indicates which data to re-attempt downloading. Upon re-attempting to import supplemental data, the original ImportLog record for the given ship, cycle and supplemental data type is updated with the current import status.

It is not uncommon for supplemental data to be missing for a specific time or position, but available in the vicinity of a position. To ensure the most accurate datasets, actual values are downloaded whenever available. When not available, to ensure a complete dataset, BW Master™ calculates supplemental data values as the mean of the values in a grid surrounding the ship position, or by feeding minimum/maximum position and/or time to supplemental data sources. When such a mean is used, the corresponding data point is flagged in the SuppData.

Below, examples of the supplemental data retrieved are described and images of the typical datasets are detailed.

4.6.1 Ship position

*.bwmonitor files do not contain any position information such as GPS coordinates.
Position coordinates serve as primary keys for all other supplemental data. They are thus needed to retrieve all other supplemental data. The ship's position may be retrieved from the ship's navigational software, but in case this is not available the following solution is applied..

Using the ship's IMO number, and the first and last measurement times within a given cycle, the Vessel Historical Track (PS01) Application Programming Interface (API) of www.marinetraffic.com returns an array of positions for the entire time period in increments of a few seconds. Coordinates of the first non-null position are assigned as ship position for the entire monitoring cycle. That is, the ship's position for a given monitoring cycle is its *earliest* available position.

Ship position data are however only available when the ship is in or close to harbour, as only terrestrial-Automatic Identification System (AIS) data are fetched[3].

[3] Satellite-AIS data is also available but is more expensive.

4.6.2 Sea surface temperature

Sea surface temperature background data are hosted by NOAA and found at https://podaac.jpl.nasa.gov/dataset/UKMO-L4HRfnd-GLOB-OSTIA

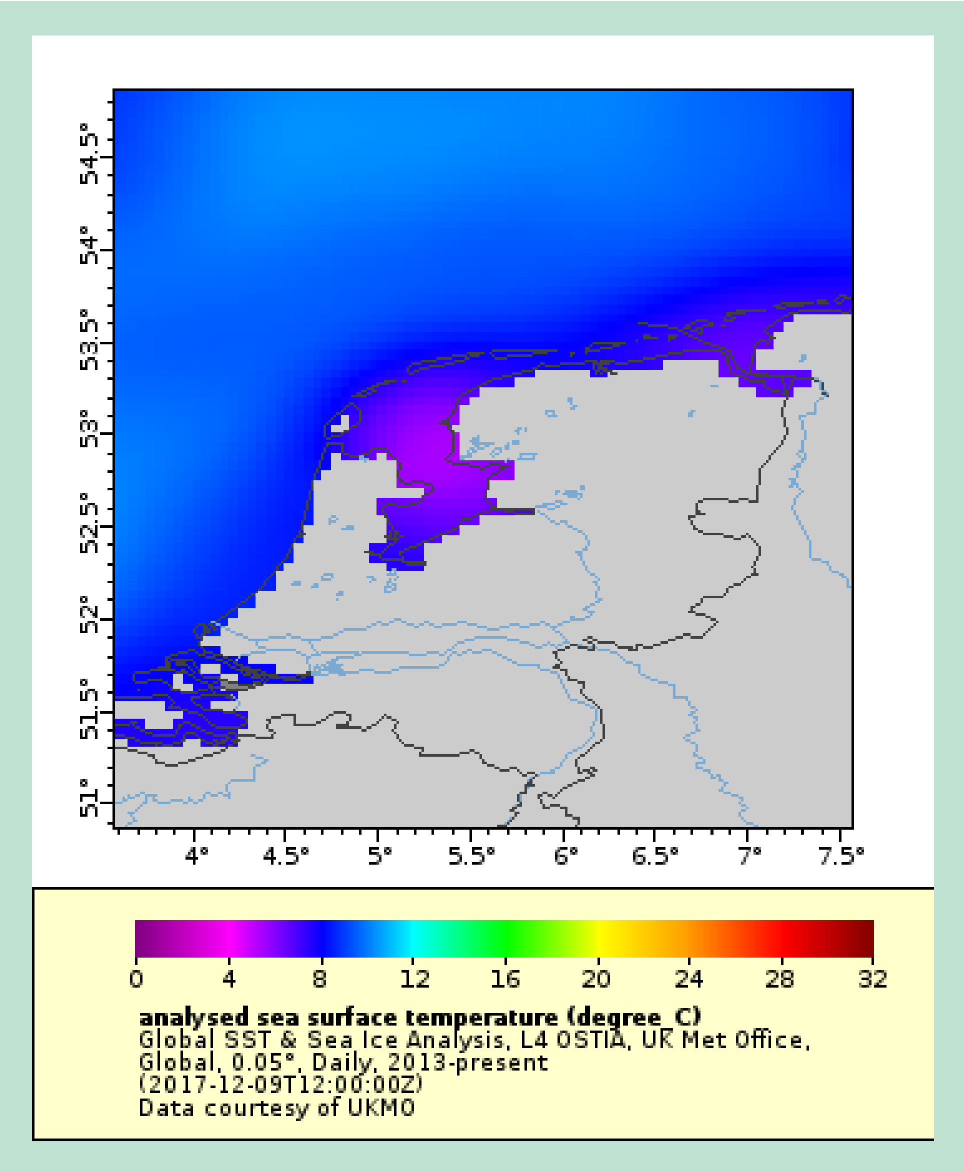


FIGURE 14. Sea Surface temperature supplemental data, example from the Netherlands

- Sea surface temperatures are very high-resolution (0.05 degrees) data with very good coverage as shown in **Fejl! Henvisningskilde ikke fundet.**.
- A 0.2x0.2 degree grid is retrieved and the mean value is inserted.

4.6.3 Sea surface salinity and wind speed

Sea surface salinity and windspeed background data are hosted by NASA and found at https://podaac.jpl.nasa.gov/dataset/SMAP_JPL_L3_SSS_CAP_8DAY-RUNNINGMEAN_V4

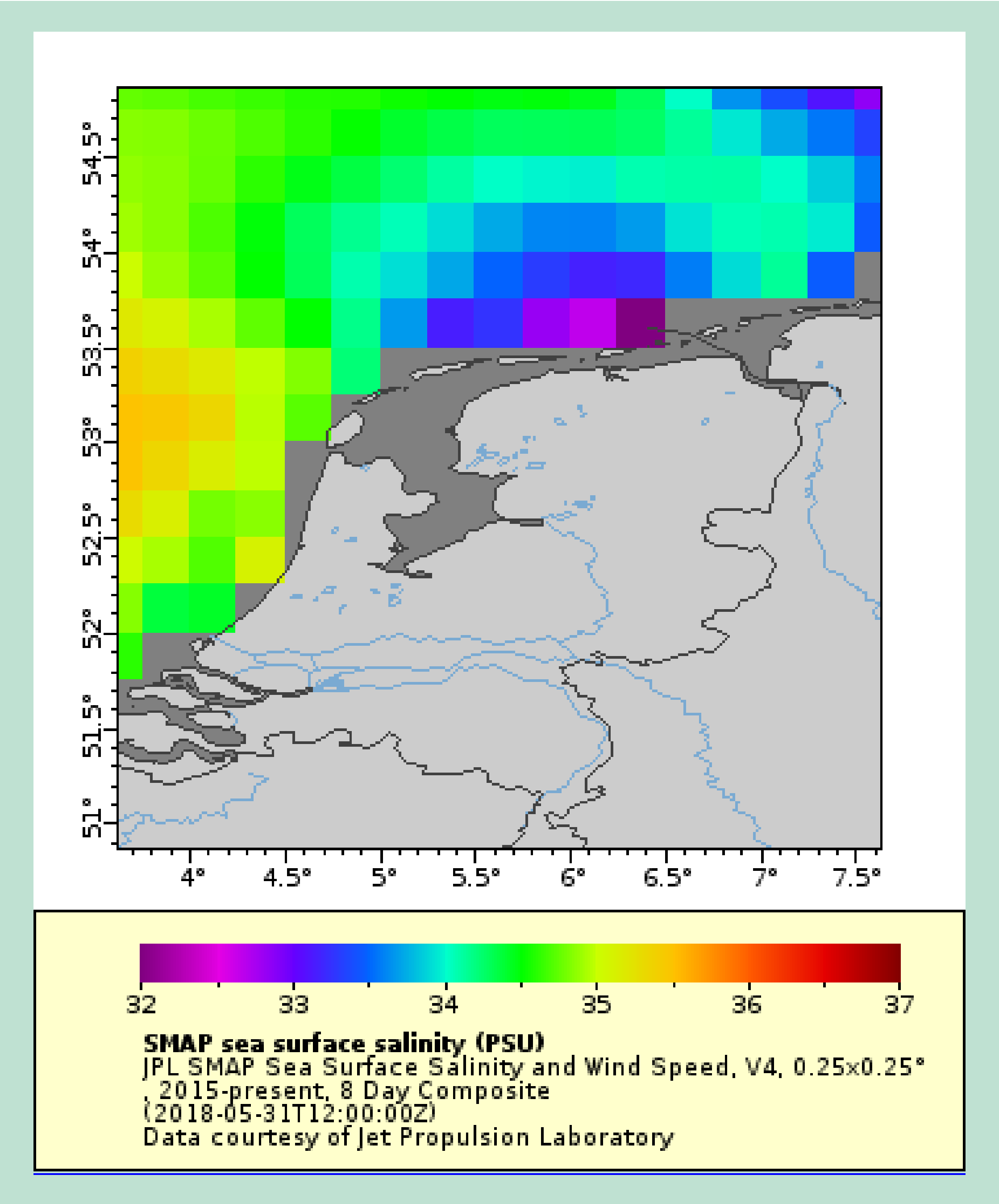


FIGURE 15. Sea Surface Salinity and Wind Speed supplemental data, example from the Netherlands

This dataset has a medium high resolution of 0.25 degrees and medium coverage.

- Some harbours are not covered.
- A 0.75x0.75 degree grid is retrieved and the mean value is inserted.

4.6.4 Ocean currents

Ocean currents background are hosted by NASA and found at https://podaac.jpl.nasa.gov/dataset/OSCAR_L4_OC_third-deg

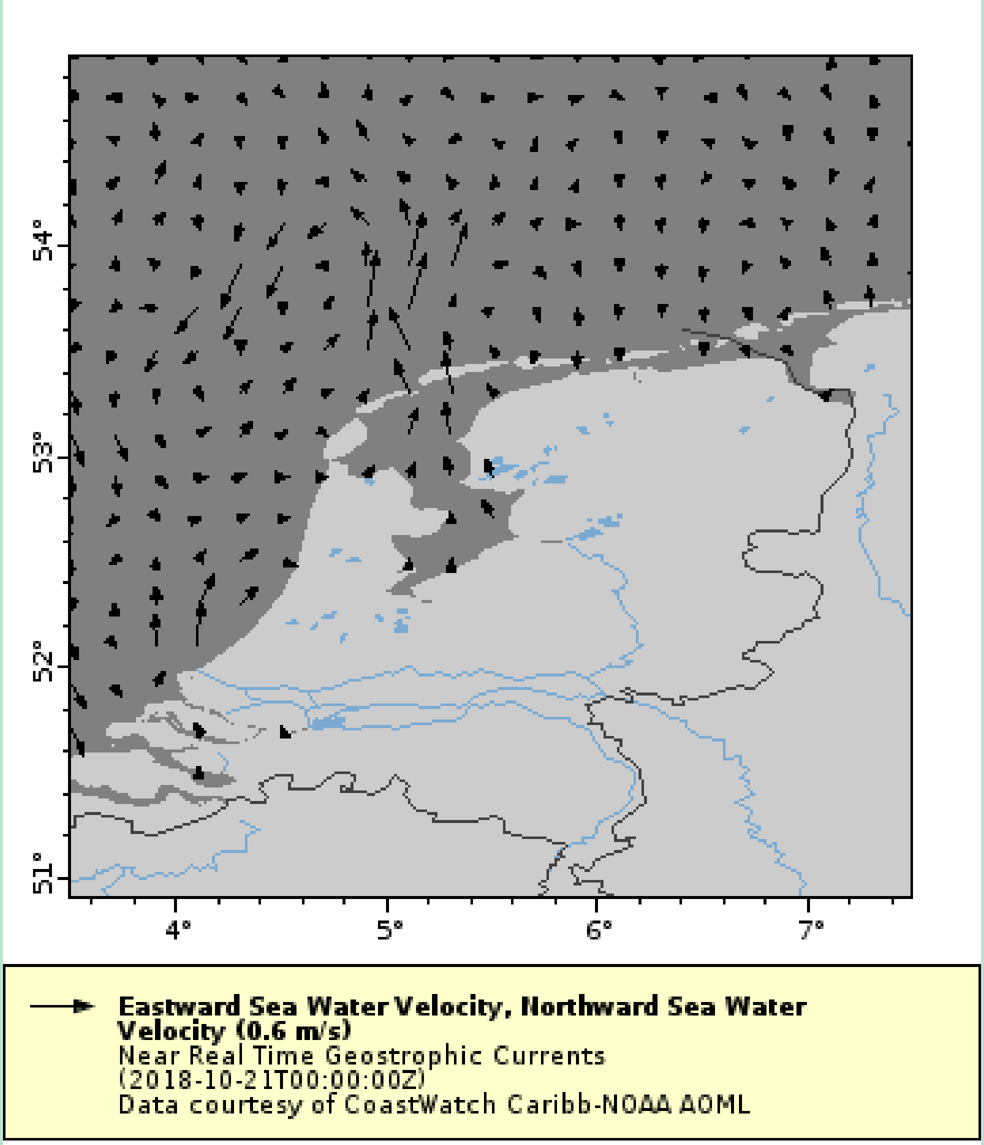


FIGURE 16. Ocean currents supplemental data, example from the Netherlands

- Currents are represented with a u_current and v_current value
- The spatial resolution is 0.2 degrees
- Coverage is good: most harbours are covered
- A 0.4x0.4 degree grid is retrieved and the mean value is inserted.

4.6.5 Chlorophyll concentration

Chlorophyll background data are hosted by NOAA and found at
http://www.aoml.noaa.gov/phod/dataphod/work/trinanes/INTERFACE/index.html

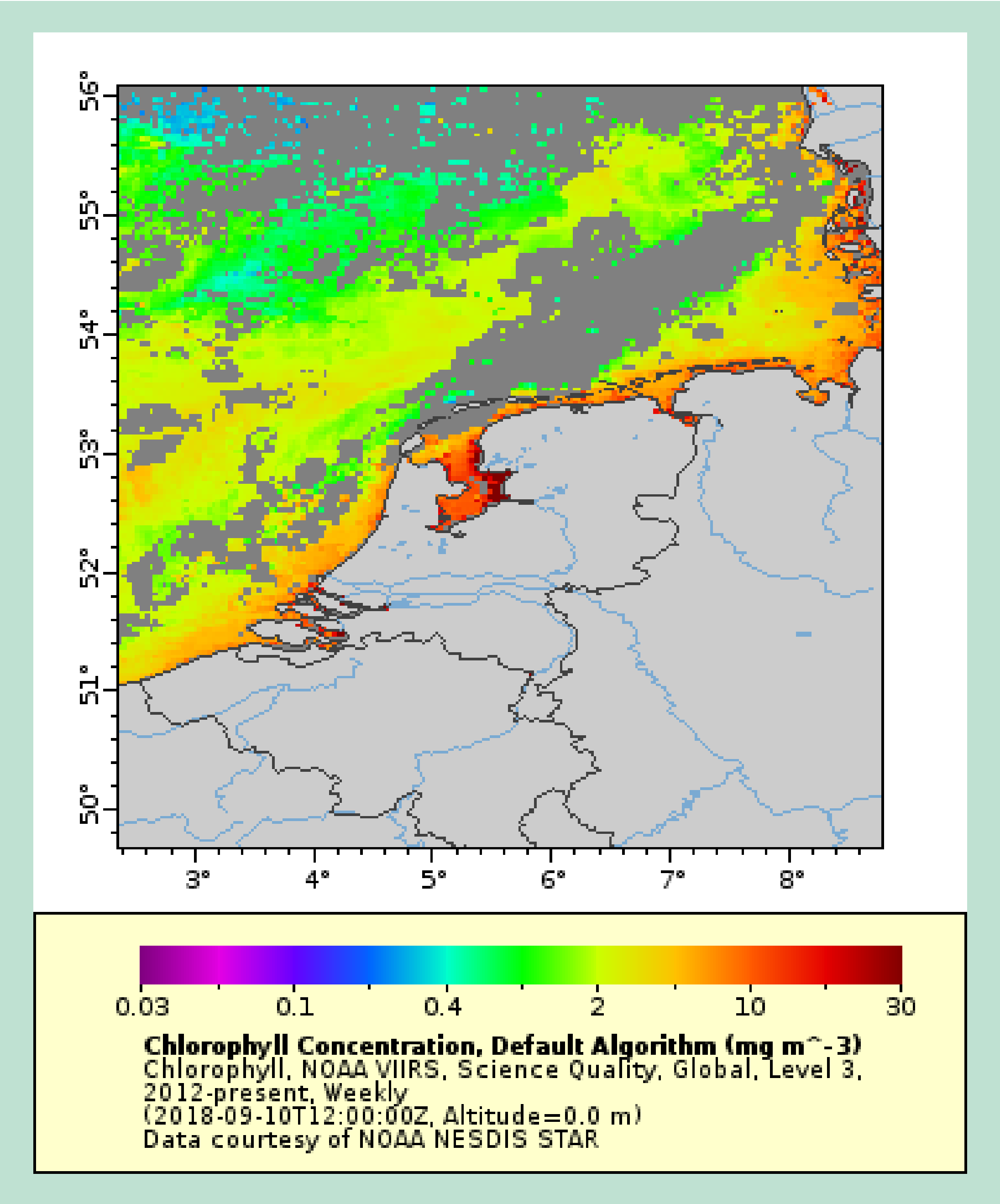


FIGURE 17. hlorophyll Concentration supplemental data, example from the Netherlands

- The spatial resolution is 4.17 km or approximately 0.037 degrees.
- The weekly aggregated dataset is chosen for its better coverage.
  A daily aggregated dataset is also available but with lower coverage.
- A 0.222x0.222 degree grid is retrieved and the mean value is inserted.

4.6.6 Pole tide

Pole tide background data are hosted by NASA and found at https://podaac.jpl.nasa.gov/dataset/ALTIKA_SARAL_L2_OST_XOGDR

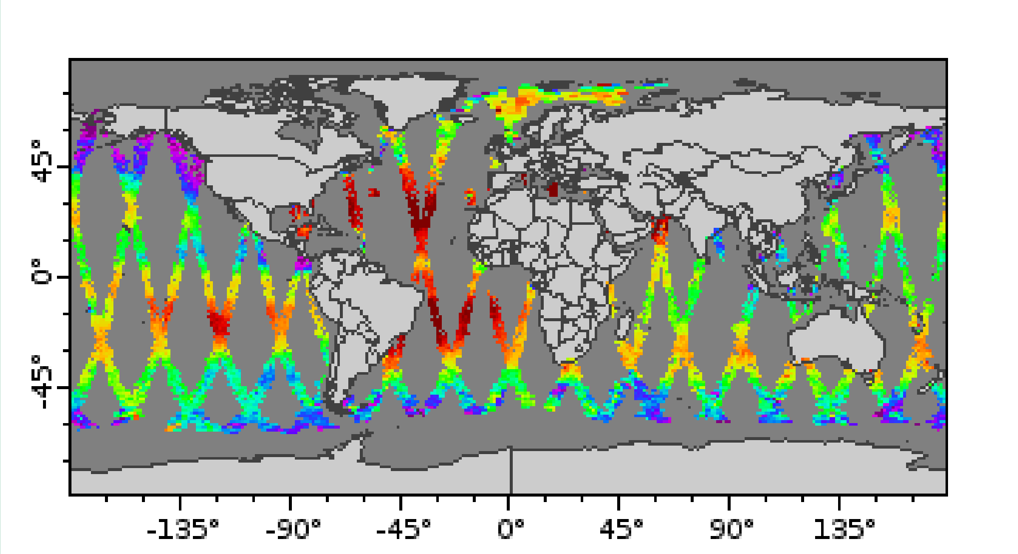


FIGURE 18. Pole Tide supplemental data

Fejl! Henvisningskilde ikke fundet. provides an example of this swathed data on a given day (image is not from the real dataset).

- Temporal repeat is 35 days, so on average 1/35 of the ships will be enriched with this dataset.
- Data is near real time which is required by pole tide which changes hourly.

4.6.7 On-demand vs. scheduled access to supplemental data

Initially, on-demand, temporary downloading of weather data was proposed. i.e., not stored but rather temporarily downloaded for display and analysis purposes. However, scheduled, permanent downloading has been implemented for the following reasons:

1. Downloading can take up to several minutes which is an unacceptably long waiting period for users.
2. Some 3rd party web servers have downtime issues, again imposing long waiting times for users.
3. The size of this data is minimal as there will only be one row for each dataset per monitor cycle.

4.6.8 Adding other supplemental data

All datasets from https://coastwatch.pfeg.noaa.gov/erddap/griddap are easily implementable Datasets from https://podaac.jpl.nasa.gov/datasetlist with OPeNDAP direct access are also easily implementable despite their variable structures.

4.7 Data visualization tool

Incorporation of the data visualization tool for the bw-monitor system was recommended in order to improve the software's algorithm to be more predictive and more accurate in defining the potential malfunction of ballasting systems. The scope is to perform the pattern recognition and statistical clustering of the result logged in the bw-monitor cloud and imported to the BWMaster database, which is a part of the big data analysis, in pursuance of most suitable and fitted monitoring system for each particular treatment type. However, it is now not intended to be an automated exercise.

The data visualization is realised by using the Tableau software. The multiple options to represent data in different views, applying filters /drill downs /formatting, creating sets, groups, generating trend lines and performing forecasting is a good match with the envisioned needs. Data from the BWMaster^{TM} database are easily connected to Tableau and certified users are able to, in a graphical way, explore the analysis of the pattern.

The analysis will allow instant intuitive access of the vessel’s performance and pattern identification. Ship owners, manufacturer and Port State Control (PSC) potentially have the same level of interest in recognition of patterns.

1. Ship owner:
   - Proof of treatment effectiveness
   - Which ship has repeated issues with the BWMS
   - Pattern for where/when is an alarm monitored?
   - Is there a correlation between the type of BWMS and alarms?
   - Which BWMS is better suited for which ship type
   - Is there a correlation between when and where the BWMS is installed (e.g. same crew)
   - Avoiding port fees by anticipating position port of discharge
   - Pattern for weather, salinity etc. conditions in correlation with alarms

2. Manufactures
   - It there a trend that certain types of BWMS have issues with certain conditions
   - Prove of system effectiveness
   - BWMS product development

3. Port State Control
   - Better use of resources through targeted sampling if there is a pattern that ships with a certain port of origin have a higher risk of being noncompliant

By connecting the BWMaster database with Tableau the collected data can be manipulated in various ways. Before starting the analysis, a data source has to be chosen. Data tables originating from the BWMaster database can be linked in four ways namely, inner join, left join, right join or full outer join, as shown in **Fejl! Henvisningskilde ikke fundet.**. Depending on the selected join, a preselection is made as to which data is available for the analysis. For example, will a left join produce a complete set of records from Table A, with the matching records (where available) in Table B. If there is no match, the right side will contain null. In **Fejl! Henvisningskilde ikke fundet.** and **Fejl! Henvisningskilde ikke fundet.** are examples given for possible tables links. Furthermore, are the available tables from the BWMaster data base displayed in **Fejl! Henvisningskilde ikke fundet.**.

The analysis in Tableau is initiated by opening a new sheet in the Tableau desktop application and dragging and dropping the desired data into the dedicated fields. The diversity in possibilities is displayed in **Fejl! Henvisningskilde ikke fundet.** i.e. which data, plot, filter, labels etc. to choose from and how to arrange them.

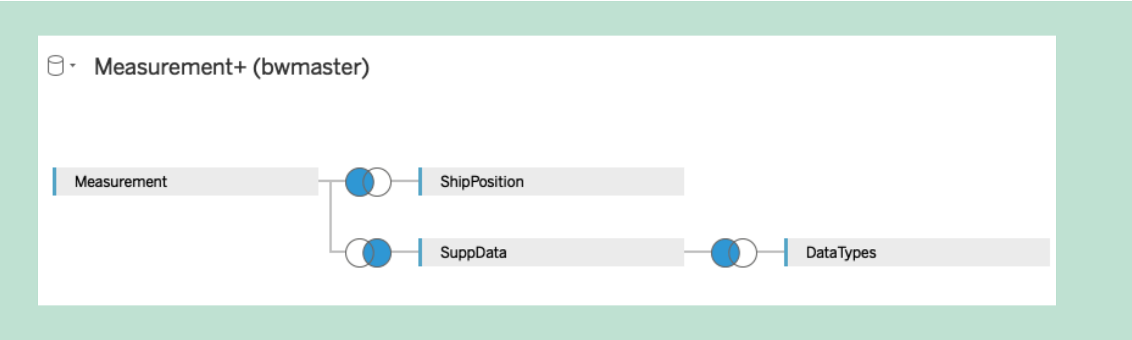


FIGURE 19. Example of table connection tree

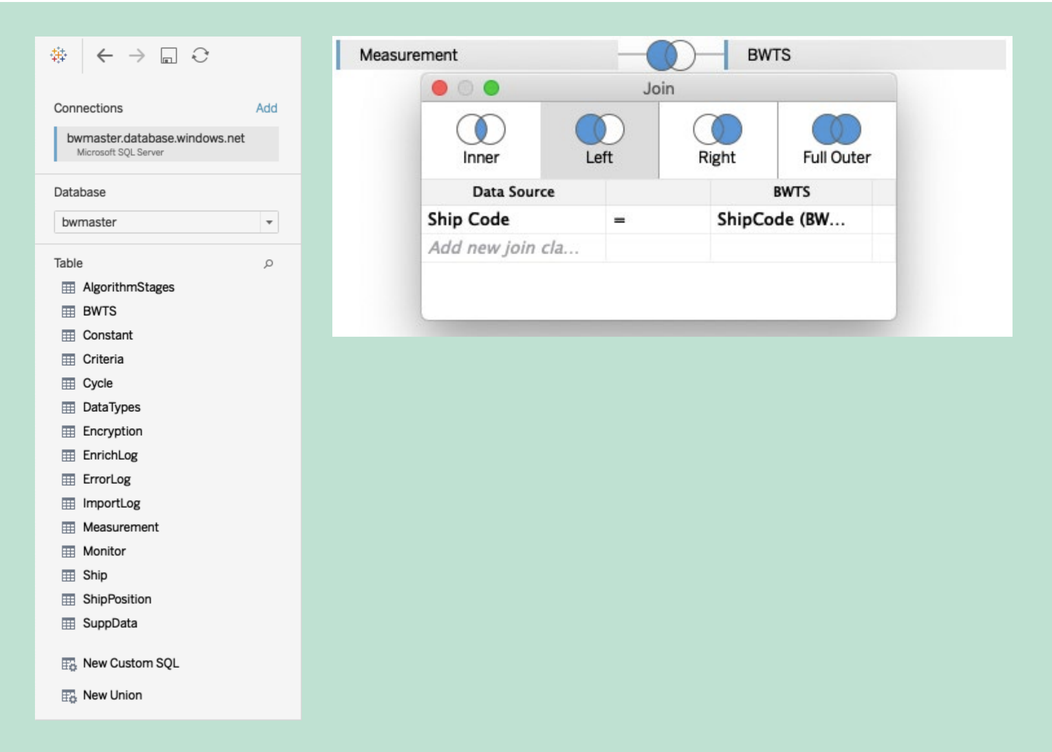


FIGURE 20. Tableau Desktop screenshot. To the left are tables with various data to the right are the four join possibilities of table connections.

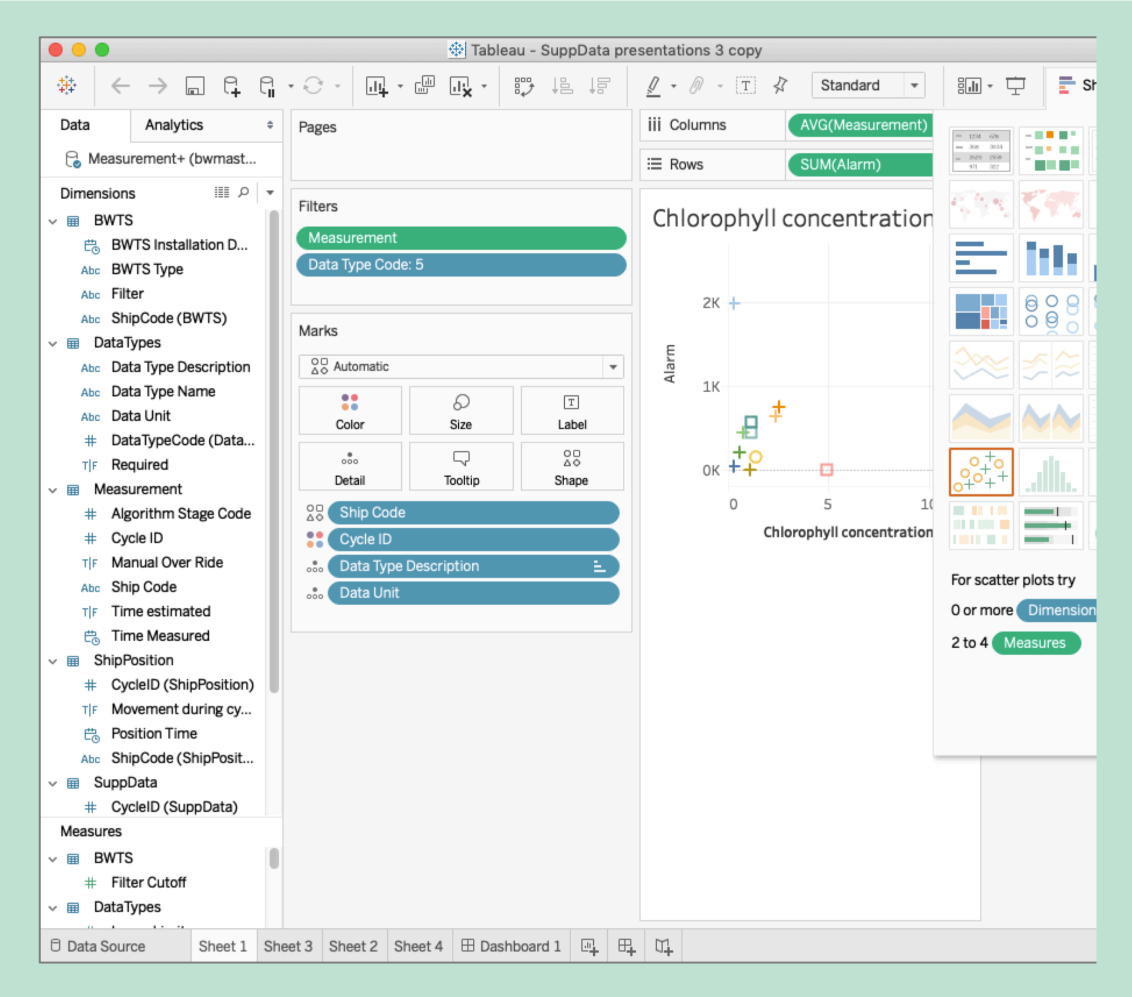
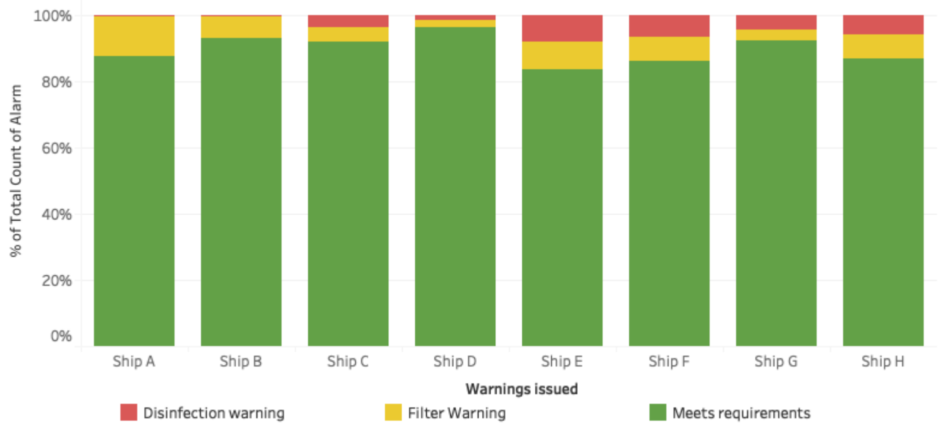


FIGURE 21. Screenshot of Tableau Desktop's user interface. On the left is the available data that can be dragged and dropped to different areas in order to illustrate the selected analysis. The window on the right displays the different plot choices.

In the following section examples of the variety of analyzing and visualizing big data with Tableau are shown. For the purpose of demonstrating the functionality of the illustration tool real date was extended by dummy data. The examples shown are intended to give an idea of the tool options but are not exhaustive.

Here is an overview of the examples:

- **Fejl! Henvisningskilde ikke fundet.** and **Fejl! Henvisningskilde ikke fundet.** are illustrations of warnings per ship in various formats.
- **Fejl! Henvisningskilde ikke fundet.**, **Fejl! Henvisningskilde ikke fundet.** and **Fejl! Henvisningskilde ikke fundet.** displaying warnings per ship and ship cycle on a map with different focus.
- **Fejl! Henvisningskilde ikke fundet.** shows the warnings per Ballast Water Treatment System (BWTS) type, i.e. UV, ozonation, chlorelectrolysis etc.
- **Fejl! Henvisningskilde ikke fundet.** is an example of identifying the supplemental data available for a ship and ship cycle in table format.
- **Fejl! Henvisningskilde ikke fundet.** shows a scatter plot of the number of alarms in correlation with the associated Chlorophyll concentration.
- **Fejl! Henvisningskilde ikke fundet.** is an example for including the time of ballast operation, the number of alarms and the associated Chlorophyll concentration on a map.



Warnings per ship

Warnings issued	Ship A	Ship B	Ship C	Ship D	Ship E	Ship F	Ship G	Ship H
Disinfection warning	0.52%	0.31%	3.90%	1.66%	8.22%	6.65%	4.50%	5.84%
Filter Warning	11.89%	6.70%	4.33%	1.89%	8.23%	7.24%	3.07%	7.50%
Meets requirements	87.59%	92.99%	91.77%	96.45%	83.55%	86.11%	92.44%	86.66%

FIGURE 22. Example of showing two types of alarms, disinfections and filter, for different ships. On top a histogram is shown and, in the bottom, the same data is shown in table format.

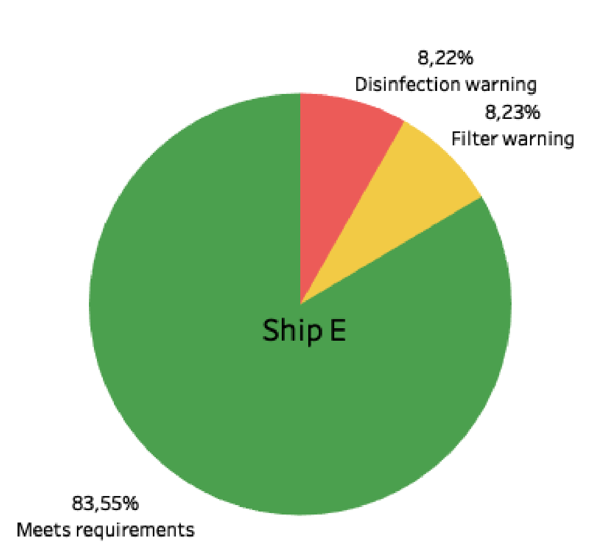


FIGURE 23. Pie diagram of warnings per ship. Shown are disinfection alarms, filter alarm and no alarm.

The next three illustration below show how to add data on a map. It is possible to include several ships and ship cycles and the associated alarms together on one map as shown in **Fejl! Henvisningskilde ikke fundet.**. It is possible to enlarge a certain region on this map and thereby get more details as displayed in **Fejl! Henvisningskilde ikke fundet.**. Another option is to select only one ship and put the alarms per ship cycle on a map, as shown in **Fejl! Henvisningskilde ikke fundet.**. These are only few examples of how data is displayed on a map.

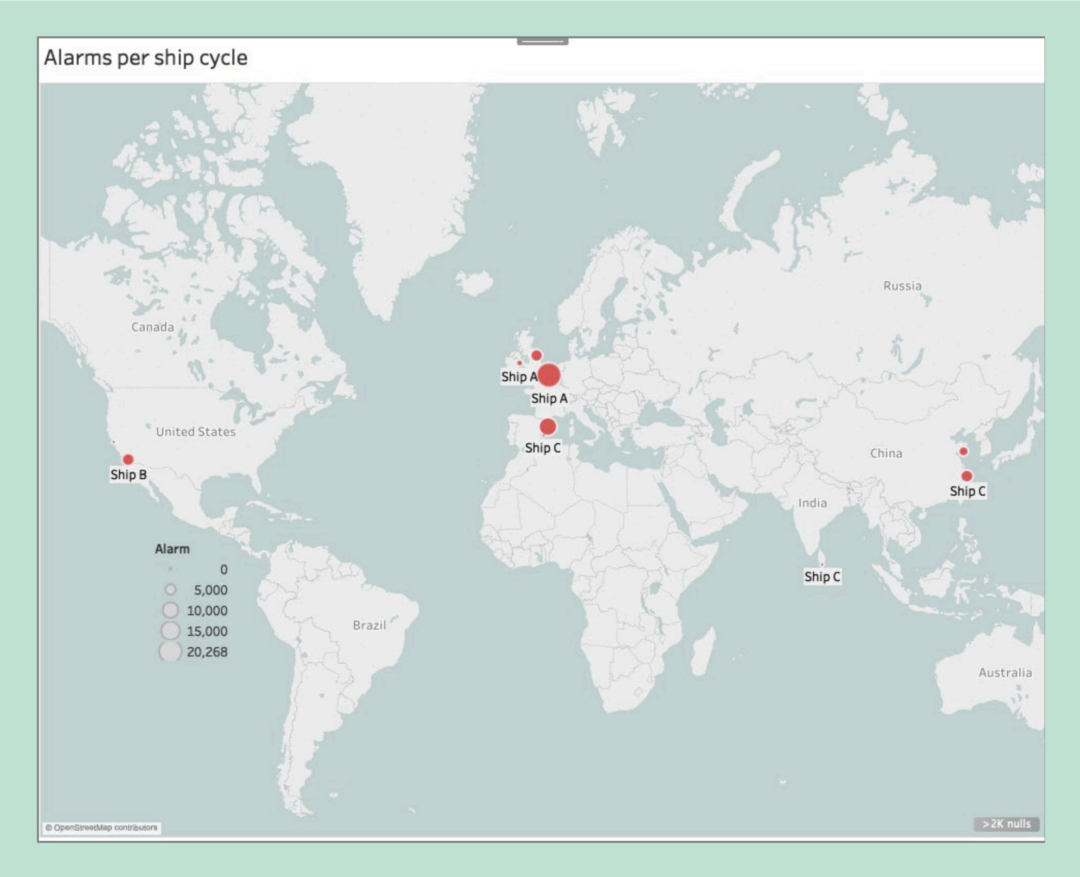


FIGURE 24. Example of three ships mapped with the location and the number of alarms per ship cycle. Each dot is representing a ballast operation while the dot size corresponds to the numbers of alarms.

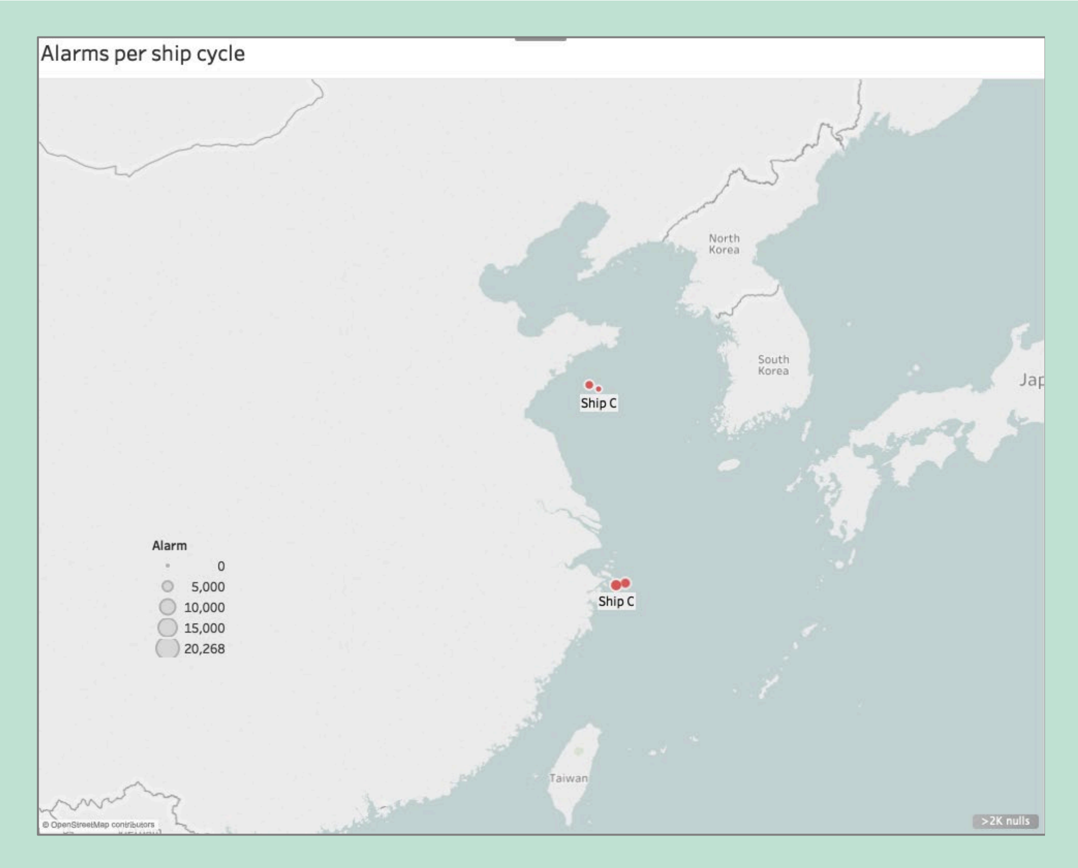


FIGURE 25. Same illustration as Fejl! Henvisningskilde ikke fundet. **except here it was zoomed in on a selected location.**

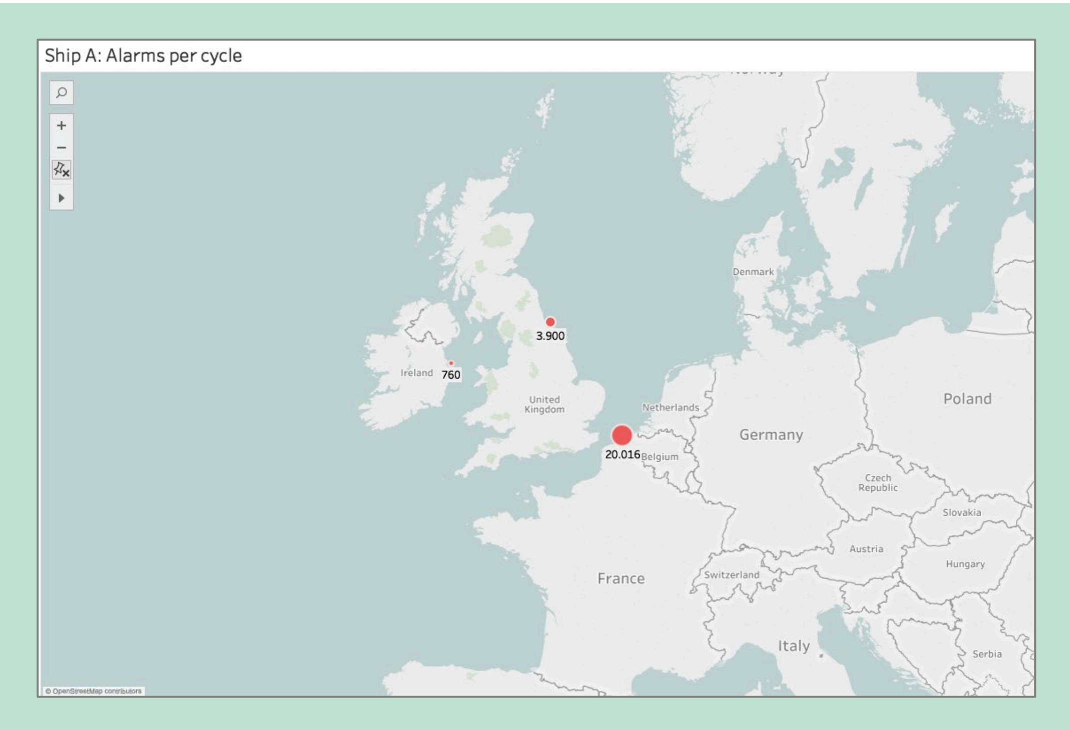


FIGURE 26. Example of alarms per cycle for on specific ship. The labels show the sum of alarms.

A correlation between a BWTS type, such as UV, ozonation or chlorelectrolysis and the frequency of warnings could be investigated as illustrated below.

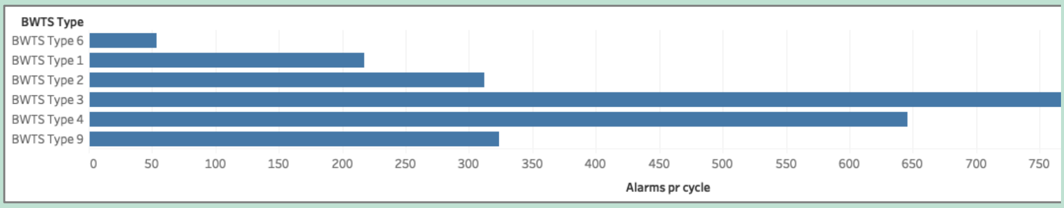


FIGURE 27. Example of warning frequency by BWMS type.

The next illustration gives an example of displaying the available supplemental data per ship cycle sorted by ships. This can be useful to make a selection of what to investigate further. Here it was chosen to illustrate Chlorophyll data in relationship with alarms as shown in **Fejl! Henvisningskilde ikke fundet.** and **Fejl! Henvisningskilde ikke fundet.**.

		Data Type Description / Data Unit							
		analysed sea surface temperatu..	Chlorophyll Concentration, Def..	geocentric pole tide height	Northward Sea Water Velocity	Eastward Sea Water Velocity	SMAP 10 m wind speed	SMAP sea surface salinity	tidal h
Ship Code	Cycle ID	degree_C	mg m^-3	m	m/s	m/s	m s-1	1e-3	
Ship A	3601	15.08	1.36		-0.12	0.05			
	5299	8.28	11.19	0.00	-0.03	0.10	4.47	36.30	
	5308	6.84			-0.06	0.05	10.63	35.18	
Ship B	5289	16.74	1.16		-0.02	0.01	2.46	33.29	
	5290	16.74	1.16		-0.02	0.01	2.46	33.29	
	5328	13.35	4.93		0.00	0.05	5.32	33.27	
Ship C	9	23.55	0.30		-0.01	-0.06			
	61	27.58	0.27		-0.01	0.01			
	72	25.65	2.56		0.15	0.07			
	74	26.44	2.38		0.18	-0.04			
	95	22.72	0.60		0.03	0.00			
	96	22.10	0.73		-0.05	-0.02			
	1277	27.58	1.09		-0.09	0.00			

FIGURE 28. Example of a table showing the supplemental data from third parties, e.g., NOAA per ship and ship cycle.

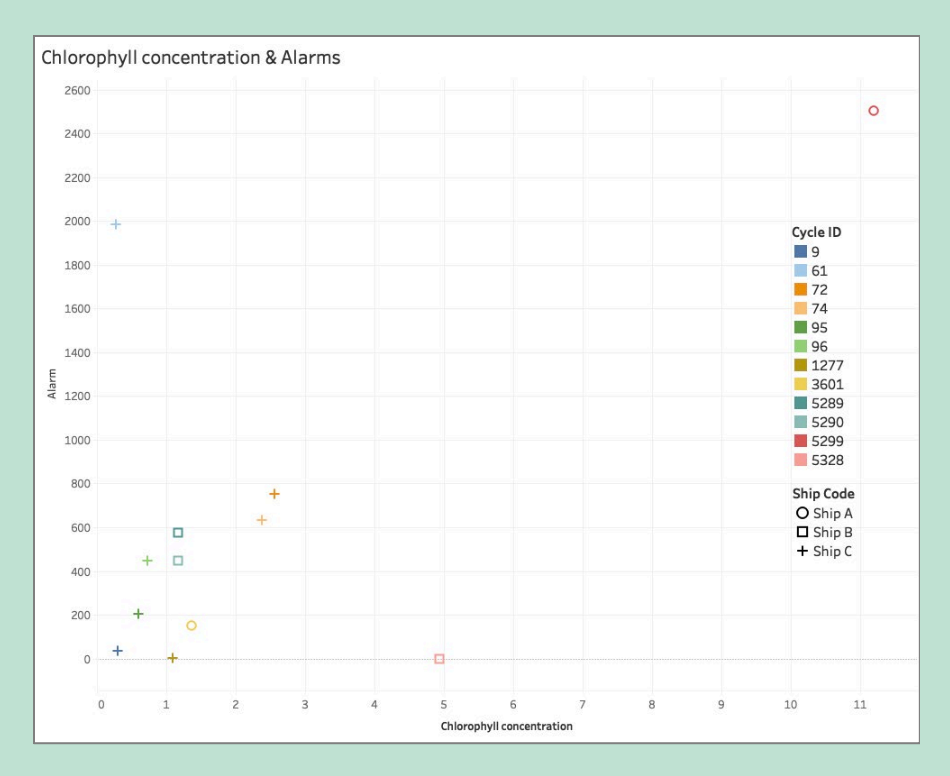


FIGURE 29. Example of a scatter plot where the number of alarms over the measured chlorophyll concentration is displayed. The different colours stand for the cycle ID and the different symbols showing the different ships

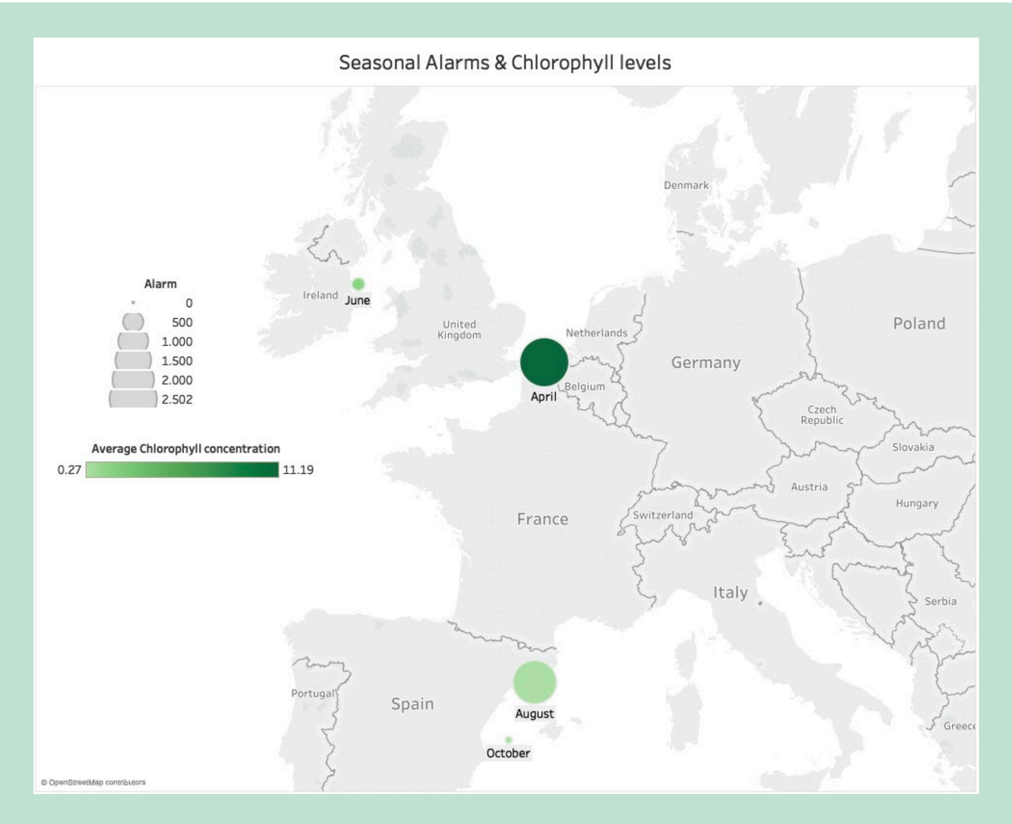


FIGURE 30. Example of map displaying alarms in correlation with Chlorophyll concentration and the time the ballast operation was performed.

Already with the introduction of 20-30 vessel's datasets over one year and with the utilization of data from different BWMS technologies the pattern recognition will be more precise and provide answers to complex questions regarding the operability and performance of BWMS as a function of a wide range of parameters. With the design of the BWMaster^{TM} database and the database enrichment mechanisms, the foundation has been laid to be equipped for the potential market for comprehensive, complex data analysis.

5. Conclusions

The project AitBall was granted under the Ecoinnovation programme in December 2012 to support the development of the monitor for ballast water. A hardware change during the project delayed the system but in 2018 the data architecture, the data flow and management, and describing the algorithms to handle our own data and Third Party data are all complete and the AitBall project is therefore finalised in February 2019. We have carried out all activities and met all five subobjectives:

- Assessment of measurement techniques
- Selection of parameters and development of simple algorithms
- Specification of data requirements and architecture for a cloud-based platform
- Establishing a database with data management and analytical capacity
- Determination of model for mathematical learning

The last bullet has been interpreted as development and implementation of our big data pattern analysis tool, as presented overleaf.

We are also proud to report that a company “Ballast Water Monitoring A/S” since 2013 has commercialized the bw monitor hardware and currently manufactures and sells the system to BWMS manufacturers and shipowners worldwide.

Monitoring of Ballast Water Quality

This report summarises the efforts of LITEHAUZ in collaboration with Medico Chemical Lab, to develop an automated monitor of ballast water quality. The instrument should reliably indicate the performance of a ship's system for treating ballast water under the IMO's Ballast Water Management Convention, and do so directly in the ballast pipe already during ballast water uptake, without addition of chemical or interaction from crew.

The ballast water monitor is based on two well-known and robust technologies – light scattering and fluorescence. Viable organisms are detected according to a fluorescence signal from chlorophyll, which is detectable directly in the ballast water stream, and an indication of the presence of particles, including larger organisms, is provided by a laser diffraction signal related to the density of particles. Furthermore, algorithms control and manage the monitor as well as analyse incoming data in real-time, allowing the BWMS operator to assess the performance of the ballast water management system, and take immediate action in case of malfunction.

Since 2013, the company "Ballast Water Monitoring A/S" has commercialized the ballast water monitor hardware and currently manufactures and sells the system to BWMS manufactures and ship-owners worldwide.

The Danish Environmental
Protection Agency
Haraldsgade 53
DK-2100 København Ø

www.mst.dk