ICES Survey Protocols – Manual for *Nephrops* Underwater TV Surveys, coordinated under ICES Working Group on *Nephrops* Surveys (WGNEPS)

Volume 65 I May 2021

ICES Survey Protocols – Manual for *Nephrops* Underwater TV Surveys, coordinated under ICES Working Group on *Nephrops* Surveys (WGNEPS)

1st Edition

Authors

Helen Dobby • Jennifer Doyle • Jónas Jónasson • Patrik Jonsson
Ana Leocádio • Colm Lordan • Adrian Weetman • Kai Wieland

ICES
CIEM

International Council for
the Exploration of the Sea

Conseil International pour
l'Exploration de la Mer

Contents

I   Background

Underwater television (UWTV) surveys for *Nephrops* are considered to be independent of the diel and seasonal behavioural patterns that affect the availability of the species to standard fishing gear (trawl surveys); this therefore makes UWTV surveys unique in providing direct comparisons between years and fishing grounds. This survey protocol presents guidelines for carrying out a standardized UWTV survey, providing the principal fishery-independent information required to determine the stock status of and formulate advice on *Nephrops* stocks. This ensures that the potential benefits of the survey are realized.

The Working Group on *Nephrops* Surveys (WGNEPS), formerly known as the Study Group on *Nephrops* Surveys (SGNEPS), is the co-ordinating expert group for *Nephrops* UWTV and trawl surveys. Since the group's first meeting in Barcelona in 2013, its objective has been to provide international coordination for both types of surveys in the North Atlantic. More specifically, WGNEPS has focused on planning, protocols, quality control, design, and survey development issues.

UWTV surveys are used to provide abundance estimates for *Nephrops* based on functional units (FUs) in ICES areas as well as in a preliminary and exploratory way in several geographical subareas (GSAs) in the Mediterranean (Figure I.1). These surveys measure abundance by counting all *Nephrops* burrow systems in patches of suitable mud habitat as in general there is a relationship between burrows and sediment type (Campbell *et al.*, 2009). Their characteristic entrance shape and other signature features allow *Nephrops* burrows to be distinguished from those of other species. These burrow features have been described in detail (Marrs *et al.*, 1998) and more recently at ICES workshops (Annex 5 in ICES, 2008; Section 4 in ICES, 2016b). Correction factors such as edge effects, burrow detection and identification, and burrow occupancy take into account the potential bias in the methodology and are specific for each *Nephrops* FU (Table 5.1 in Leocárdio *et al.*, 2018). These bias correction factors are fundamental in providing absolute abundance estimates for stock assessment.

UWTV surveys were first carried out at the Fladen Ground in 1992 by Marine Scotland Science (MSS). Since then, the number of *Nephrops* stocks routinely surveyed in this way has increased over time, with 20 *Nephrops* grounds being subject to surveys up until 2018 and a further three (FUs 5, 10, and 33) subject to trial surveys (Figure I.2). Table I.1 list the countries involved in surveys by each *Nephrops* FU and GSA. In the last decade there has been significant progress towards establishing a consistent, efficient, and effective method for assessing and advising on the status of *Nephrops* resources in European seas using UWTV surveys. Across institutes there are differences in survey design and the operating procedures employed. This document summarizes the essential operational requirements and protocols for carrying out a standard UWTV survey.

Table I.1. List of countries involved in UWTV surveys in each *Nephrops* functional unit (FU) and geographical subarea (GSA).

Country	Institute	FU/GSA
Iceland	MRFI	FU 1
Denmark & Sweden[1]	DTU-AQUA and SLU	FUs 3–4
UK England	CEFAS	FUs 5– 6
UK Scotland	MSS	FUs 7–13, 34
UK Northern Ireland[2]	AFBI	FUs 14–15
Ireland	MI	FUs 16–17, 19, 20-21, 22
France	IFREMER	FUs 23–24
Spain	IEO	FU 30
Denmark	DTU-AQUA	FU 33
Italy & Croatia[1]	CNR-IRBIM and IZOR	GSA 17

[1] Survey conducted in collaboration.

[2] Survey conducted in collaboration between UK-Northern Ireland, Marine Institute and UK-England.

Figure I.1. *Nephrops* **UWTV survey areas. The map shows** *Nephrops* **functional units (FUs) in ICES areas (numbers in blue) and geographical subareas (GSAs) in the Mediterranean (numbers in red). The term 'data-limited stocks' is abbreviated to 'DLS'.**

Figure I.2. UWTV surveys in the different *Nephrops* FUs and one GSA over time (up to 2018). For each FU/GSA, the blue dot indicates the first year of the survey, the light grey dots indicate the years in which surveys were not conducted, and the grey line shows the survey series.

1   Survey objectives

Primary objective of the UWTV survey

The main objective is to generate quality-assured estimates of *Nephrops*' absolute abundance within defined areas with a coefficient of variation (CV) or relative standard error of less than 20%, as recommended by ICES Study Group on *Nephrops* Surveys SGNEPS (ICES, 2012).

Secondary objectives of the UWTV survey

The use of UWTV survey platforms to collect habitat and other ecosystem data was first documented at ICES Workshop on the Use of UWTV Surveys for Determining Abundance in *Nephrops* Stocks throughout European Waters (WKNEPHTV, ICES, 2007). As long as it is not detrimental to the main focus of the survey, and depending on factors such as time, equipment, and staffing levels, a UWTV survey could be used to collect additional biological, environmental, and anthropogenic data. Some of the most common additional survey activities are listed below; although detailing these activities is beyond the scope of this document, some reference material is given for further information.

- Collection of multibeam and sediment data to improve the definition of *Nephrops* habitat. Particle size analysis of sediment samples and multibeam backscatter data can augment information on survey area definition (ICES, 2016b).

- Monitoring of various biological parameters (e.g. sex, maturity) of *Nephrops* and provision of length–frequency distribution (LFD) time-series data for *Nephrops* if combined with beam trawl or bottom-trawl sampling. Catch length–frequency distributions from trawl operations during UWTV surveys may provide signals of recruitment (Aristegui *et al*., 2018).

- Collection of hydrographic and environmental parameters (e.g. temperature, salinity, turbidity, oxygen levels). Data from sledge mounted conductivity, temperature and depth sensors (CTD) can, for example, be viewed as an emerging time-series.

- Monitoring of trawl marks and marine litter as evidence of anthropogenic activity on *Nephrops* grounds to comply with Marine Strategy Framework Directive (MSFD) and OSPAR Convention requirements. Image datasets can be viewed to report such features by other end users.

- Full integration of benthic and ecosystem monitoring requirements under the MSFD and OSPAR into existing UWTV surveys.

2   Survey sampling design

UWTV surveys are focused on suitable habitats for *Nephrops*. Prior knowledge of *Nephrops* grounds based on available information such as sediment distribution maps, VMS data, or local knowledge is required. Descriptions of alternative methods for calculating the spatial extent of *Nephrops* stocks are available in ICES stock-specific benchmark reports.

UWTV surveys in development may have to be adaptive and add more stations in order to map the extent of the *Nephrops* patch if the extent is previously unknown. Either for exploratory purposes or in cases of limited time for dedicated UWTV surveys, any spare time available during groundfish surveys (potentially during hours of darkness) could be used for deploying UWTV for *Nephrops*.

A stratification of the survey area may be considered if the FU extends over a wider area and if there are expected differences in mean densities that will require different levels of sampling intensity. Stratification may also be considered for logistical reasons, for example in collaborative surveys involving more than one country and/or vessel.

There are two main UWTV survey design approaches currently in use: grid and stratified random design (Table 2.1). A grid can comprise either equidistant points spread over the survey area from a random starting point such that the grid is slightly offset each year or fixed points spread over the survey area. Stratified random design can be based on both sediment data and VMS data linked to logbook data with buffers to prevent clustering.

In both approaches, sampling intensity is dependent on burrow densities and the distribution of these densities over the survey area. It is essential that the sampling effort, i.e. the number of stations per stratum or total survey, is sufficient to meet the target level of precision (an overall CV or standard error of less than 20%).

Table 2.1. Summary statistics for UWTV *Nephrops* surveys giving recent average numbers of stations, ground area, density design, and CV.

Name	FU/ GSA	Area of ground (km²)	Stations per 1000 km²	Design	CV/relative standard error (based on 2017 survey)
South of Iceland	FU 1	6353	13.2	Randomized grid	3.4%
Kattegat & Skaggerak	FUs 3 & 4	13490	12.8	Random stratified	3.75%
Botney Gut & Silver Pit[1]	FU 5	1000	43	Fixed grid	Na
Farn Deeps	FU 6	2750	39.3	Fixed grid	3.00%
Fladen Ground	FU 7	28153	2.5	Random stratified	6.90%
Firth of Forth	FU 8	915	56.8	Random stratified	10.00%
Moray Firth	FU 9	2195	25.1	Random stratified	15.50%
Noup[2]	FU 10	409	12	Random stratified	Na

Table 2.1 (continued)					
Name	FU/ GSA	Area of ground (km²)	Stations per 1000 km²	Design	CV/relative standard error (based on 2017 survey)
North Minch	FU 11	2908	14.4	Random (VMS-based)	7.10%
South Minch	FU 12	5072	8.1	Random stratified	13.60%
Clyde (not including Sound of Jura)	FU 13	2081	18.3	Random stratified	7.30%
Irish Sea east	FU 14	1043	34.5	Fixed grid	10.0%
Irish Sea west	FU 15	5275	23.7	Randomized isometric grid	3.1%
Porcupine Bank	FU 16	7130	8.8	Randomized isometric grid	5%
Aran Grounds	FU 17	1202	25.8	Randomized isometric grid	4%
SW & South of Ireland	FU 19	1973	20.8	Random stratified	12%
Labadie	FUs 20 & 21	10014	8.6	Randomized isometric grid	4%
The Smalls Grounds	FU 22	3063	13.1	Randomized isometric grid	5%
Bay of Biscay	FU 23 & 24	16164	14	Randomized grid	9.9%
Gulf of Cadiz	FU 30	3000	21.3	Randomized isometric grid	8.70%
Off Horn's Reef	FU 33	5737	12.2	Random stratified	9.6%
Devil's Hole	FU 34	1753	10.8	Fixed stations (VMS-based)	Na
Pomo (Jakubo) Pits	GSA 17	4800	9.6	Random stratified	Na

[1] No surveys after 2012.

[2] No surveys after 2014.

To minimize interannual variability it is recommended to keep the survey timing fixed from year to year. Sampling is not restricted to a particular time of day, meaning that 24-hour operations can be carried out.

For some survey areas (the Irish Sea and the Celtic Sea) tide timetables need to be consulted to ensure that spring tides are avoided. Spring tides can cause soft muds to disperse in the water column, in turn causing clarity to deteriorate. This may mean it is more difficult to record high-quality images.

3   Observation methodology

3.1   Protocol for sampling gear and instrumentation

This section sets out the basic general arrangement of the UWTV sledge and sensors required to capture primary spatio-temporal image data with a calculated field of view (FoV). Details of UWTV systems in use by national institutes have been documented by WGNEPS (Annex 5 in ICES, 2016a).

3.1.1   Sledge design

The design is based on the Scottish sledge, where the sledge frame should be wide enough (typically around 1.6 m) so that any sediment clouds will not obscure the field of view. Wide sacrificial skis, a lightweight frame, and appropriate flotation should be used to avoid the sledge sinking when deployed on soft sediment (ICES, 2007).

The sledge setup should be robust enough so all instruments are secured in position, but it should also be flexible enough to allow the balance to be adjusted. Further details of the different sledge designs used are available in ICES (2007). One example of a sledge design can be found in Annex 1.
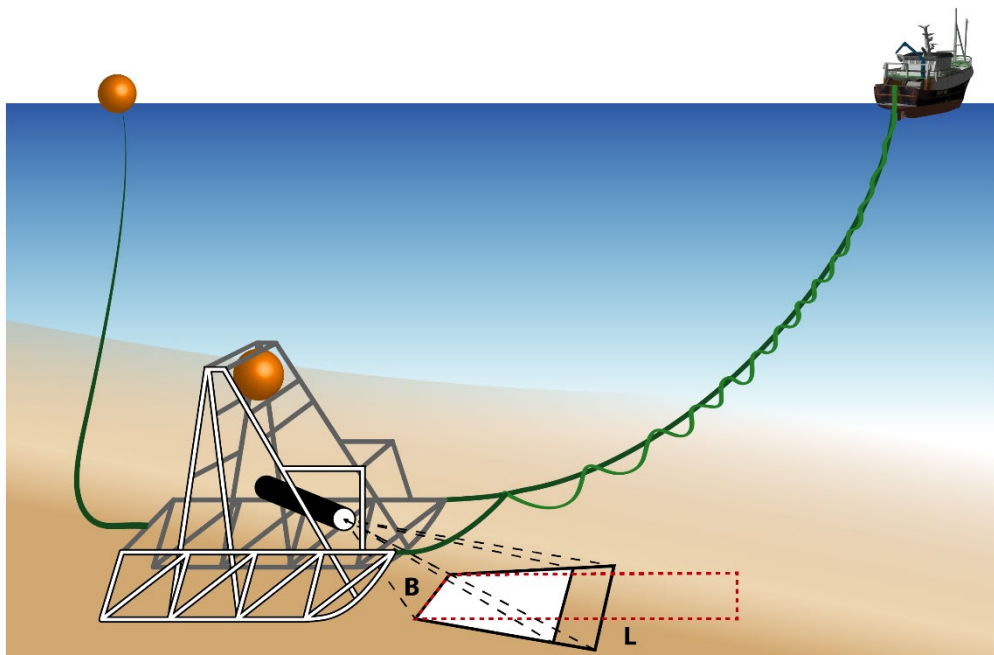

Figure 3.1. Basic principle of sledge design with floatation and angled camera setup showing the field of view (B) and distance of TV track (L). (Photo © K. Mutch, Marine Scotland Science.)

3.1.2   Camera settings

The camera should be mounted on the frame such that it is facing forwards and angled so that area in view is unobstructed by the sledge skis.

Table 3.1. Camera set up specifications by institute.

Analogue camera system		
Institute	Camera mounting angle (º) relative to the seabed	Height centre camera lens to deck (cm)
Marine Science Scotland	37	90
MI-Ireland	40	75
AFBI-NI	45	75
DTU-AQUA	60	70
SLU	45	92
CEFAS	40	97
IFREMER	40	75
CNR-IRBIM	45	80
IEO	45	80
MFRI-Iceland	45	105

3.1.3   Lighting

The light system used must evenly distribute light over the entire field of view. This can be achieved using variable voltage light sources; historically these sources have been filament lamps but more recently LEDs are used, which are considerably more robust. To improve the image data the strength of illumination should be altered to minimize turbidity and particle reflection. Cost, power availability, and sledge setup may limit the lighting array; however, four light units arranged in a square formation with two in front of the camera and two behind tend to provide the most even light dispersal within the viewed area. These light units should be mounted on the sledge on jointed brackets so that adjustments can be made in order to provide the best use of the light sources. Power adjustments between the pair of front and rear lights might be expected to find the optimal illumination. Four light sources also provide multiple shadows that alter as the sledge passes over objects of interest. These shadows assist the reviewer in correctly identifying these objects.

From experience, LED light units with output adjusted to 40% of power creates the optimal lighting scenario for use with analogue camera systems. High definition camera systems require more powerful light units.

3.1.4   Field of view setup

The field of view (FoV) can be described as the horizontal position on the viewing monitor over which any part of a burrow opening must pass to be included in the abundance count. In most cases the FoV tends to be at the bottom of the viewing monitor, i.e. closest to the camera and the narrowest point of the trapezium created by the angle of the video camera (see Figure 3.1). There may be reasons to relocate this FoV due to obstruction from equipment, the camera angle, or turbidity, and this relocation can be accommodated as long as the FoV is calibrated and well defined. Either an altimeter or lasers can be used to define the FoV. Point or fan lasers can be mounted on the sledge either side of the camera at a set distance apart, with the laser's point of contact on the seabed clearly visible on the viewing monitor. This provides a fixed FoV measurement that is not affected by the height of the camera off the seabed at any one time. The height of the camera is affected by variations in the tow speed or environmental conditions, effectively creating an elongated rectangular viewed surface area. Alternatively, an altimeter

can be mounted vertically to the sledge at a fixed, set height in relation to the camera. At regular intervals the altimeter registers the height of the unit off the seabed (and in turn the relative height of the camera off the seabed). This distance varies as the sledge rises off the seabed (due to increased tow speed, for example) or sinks into extremely soft sediment. This creates a variable FoV widening and narrowing along the length of the tow path, which is accommodated in the abundance calculation scripts. An FoV of greater than 1 m width may not allow sufficient definition for detecting and identifying burrows. In high density grounds a narrow FoV may be required, whereas in low density grounds a wider FoV may be more appropriate; experience has shown an FoV of between 0.7 m and 1 m is the most appropriate for this work.

3.1.5 Laser setup

When the FoV is defined by lasers, these have to be mounted vertically on the sledge and parallel to each other. The light emitted by the lasers must be clearly visible on the bottom edge of the monitor screen. The distance between the individual lasers is crucial for calculating the FoV.

The screen grab below (Figure 3.2) shows a green stripe laser setup with a field of view of 81.5 cm which clearly defines the countable strip of seabed.

Figure 3.2. Screen grab of video data with field of view defined by stripe laser (81.5 cm).

3.1.6 Track duration

According to ICES (2007, 2008) analysis of count data has shown that the mean and variance of burrow density per track stabilizes after around five minutes provided that the sledge covers at least either half the required distance within the full ten minutes of the run or approximately 100 m. To allow detailed examination of the seabed vessel speed should be approximately 0.7 knots, which results in a surveyed track of approximately 200 m at each station. It is recommended that each track should be of at least ten minutes duration, although this can be longer in poor viewing conditions.

The variability of distance covered per run could be reduced by changing the recording protocol from one of fixed time to fixed distance as measured by either an odometer wheel or DGPS dynamic positioning system.

3.1.7 Estimation of vessel and sledge distance over seabed

It is essential to accurately calculate the exact distance the sledge travels over the seabed during the period when image recording is taking place. This data should be presented in metres. As a

minimum, vessel global positioning system (GPS) data can be collected as proxy data for each track. For the sledge itself, this can be done by using various methods such as an ultra-short baseline (USBL) system or an odometer device mounted to the sledge.

An odometer has two components – an arm and a wheel. The arm is attached to the rear of the sledge but has a hinge which allows the wheel to remain in contact with the seabed either in undulating or rough seabed conditions or if the sea conditions are creating any lift that is transferred to the sledge. The wheel has a circumference of 1 m and, using an attached magnet and pickup on the arm, its every rotation, and therefore distance travelled, is recorded shipside. To avoid spurious or inaccurate data being recorded during sledge deployment or prior to recording being started (i.e. whilst the sledge is settling on the seabed), the arm is attached to a motor that allows the wheel to be lowered only when the sledge makes contact with the seabed, and the distance travelled can be reset to zero shipside at any time.

The USBL and odometer approaches are not affected by the winding in or out of the cable to control the sledge during the TV track. It is important to correct the vessel position if using this to calculate the distance of the TV track. This distance is calculated by recording the amount of cable paid out or in at the end of each one-minute block. It is beneficial to use record both vessel GPS and USBL/odometer data in case of sensor failure.

An example of a UWTV station sheet to record metadata during a track is given in Annex 2.

3.1.8 Quality control of navigation data

For collection of GPS navigation data for both sledge and vessel, the sledge tracks should be checked and the sledge position data used only where it is of sufficient quality. Otherwise data from vessel GPS can be used, as long as it has been corrected to account for any cable adjustment.

Figure 3.3 is an example quality control plot from a particular survey station showing that the tracks for both sledge and vessel are similar and that data from the sensors are of high quality received with no erratic signal. In contrast, Figure 3.4 shows sledge data for another survey station that is very poor quality due to it being received with an intermittent signal. In this situation vessel GPS position would be used to calculate distance over ground.

Figure 3.3. R-tool quality control plot for station 287 of the 2014 FU 20–21 survey. The sledge track is in blue and the vessel track in black. The red dots indicate the one-minute block of the sledge track. The plot shows high-quality navigation data from both sledge and vessel.
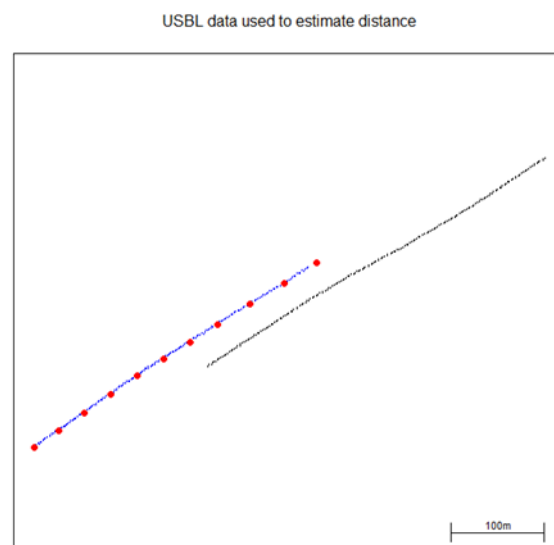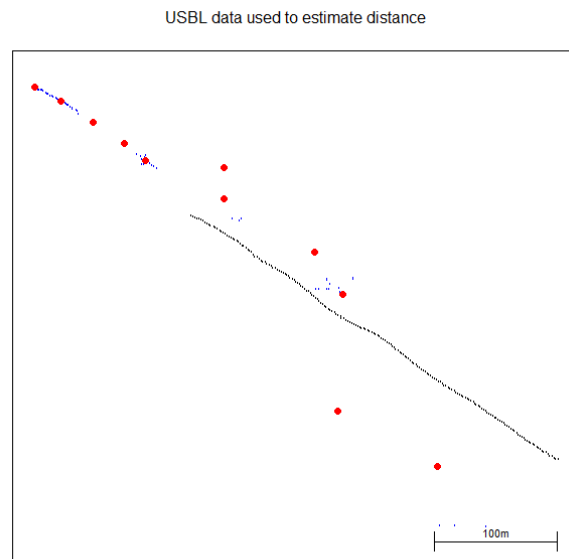
USBL data used to estimate distance

Figure 3.4. R- tool quality control plot for station 163 of the 2015 FU 20–21 survey. The sledge track is in blue and the vessel track in black. The red dots indicate the one-minute block. The plot shows very poor-quality navigation data from the sledge sensors.

3.1.9 Recording, storage, and review of image data

It is critical to synchronize navigation and image files using time and date stamps to ensure correct georeferenced data. This can be done for example, by means of image overlays, time-related file naming, and EXIF data annotations.

Image data may be recorded by many types of devices such as DVD recorders, DV tape recorders, hard disc drives, and network attached storage (NAS) drives in various formats. Whichever is used, the aim should be to achieve the highest possible quality. Image data should be backed up regularly during the UWTV survey. There is also software available to make automatic backups of files and folders.

The type of review monitors used will also depend on the camera system in use. Analogue (standard definition) TV signal is best reviewed on CRT monitors, whereas high definition can be reviewed on laptops and flat screen monitors.

3.1.10 Operational procedures

Deployments differ greatly between systems used. If the sledge is hauled in through the ramp, it is important to keep a steep angle when deploying and recovering it so its skids gently climb the ramp. The ship speed should be maintained to allow for a safe and controlled deployment of the sledge. The recommended sledge speed is approximately 0.7 knots to allow detailed examination of the seabed. From experience the length of the cable/water depth ratio should be approximately 1.8/the depth of water. The precise length of the cable depends on the platform but also on prevailing surface conditions and currents that may vary during a tow, with generally more cable required to counteract a vessel's motion in poor weather. If excess cable is paid out, it should be noted that this can drag on the seabed creating dust plumes that obscure the camera's view.

It is important to maintain constant ground contact during the TV track. This can be facilitated by adjusting the cable length or vessel speed or by adding weight to the sledge.

If lasers are fitted to the sledge, it is important to remember to turn them on during the descent and to turn them off, as well as dim the sledge lights, on the ascent after the track is completed. This is important to remember for safety purposes if the lasers are not water-activated.

To assist in the recovery of a sledge that has become separated from the towing cable, the sledge should be equipped with a locating system. This can be in the form of a safety buoy that is either attached to the sledge and towed behind it or housed on the sledge and released either by a surface-activated acoustic hydrophone or a predefined timer mechanism. To allow the buoy to be hooked up during recovery, a second small flotation should be attached to it along with around two fathoms of rope. In addition, a transponder can be mounted on the sledge; this transponder remains passive until activated acoustically from the surface. Although this does not provide a physical link to the sledge, it does accurately locate the system, allowing for the use of alternative recovery approaches (e.g. grappling, recovery trawl, and ROV).

Occasionally rocks and small boulders can be encountered when the sledge is in motion. If the system allows for multiple cameras, the addition of a forward-facing camera may enable a better view of such oncoming obstacles. Winding the cable onto the winch while increasing the ship speed would lift the sledge quickly up off the bottom to avoid damaging it.

It is useful to set up a slave monitor from a bridge sounder which can aid the driver in flying the sledge over any rocky outcrops encountered. If there is a possibility to run a multibeam backscatter such as EM2040 Dual RX in the background, this can also show when hard ground will be encountered on a TV track.

If a survey station is on "hard" ground (meaning unsuitable *Nephrops* habitat), it is recommended to record and save the information on seabed type, preferably including image and navigation data, in order to have a record of a station with zero burrow density. This zero-density data on unsuitable *Nephrops* habitat will aid in area definition on future surveys.

3.2 Counting methodology

3.2.1 Burrow identification

The main characteristics which define a *Nephrops* burrow system have been described in detail by the workshop reports (ICES 2008, 2016b).

In summary the main features are:

- a crescent shaped entrance;
- a delta of excavated material;
- where visible, a shallow angle of descent;
- entrance apexes which either face each other in a simple U-shaped system or converge on one central point in a more developed system; several entrances form a T-shape in some systems.

3.2.2 Counting procedures

A standard operating procedure for burrow counting should be produced by each institute. This should include details of how many minutes are to be counted from the video, where to count on the screen, and the removal of minute counts where image quality deteriorates. A warm-up session (or familiarization session) is to accustom the reviewer to the specific characteristics of the area from where the image data was recorded. Details of this session should also be included in the operating procedure.

The warm-up session involves selecting a random ten-minute video recorded in the same area as that being reviewed before being re-examined in the normal way. The difference is that the abundance values are not taken into consideration for the purposes of the assessment. This allows reviewers to familiarize themselves with the type of image data they are about the assess prior to generating the 'final' count data. This process should be carried out every time a reviewer takes a break from counting for more than one hour.

Counting procedure:

1. On completion of the training process, survey counts must be conducted as blind and independent counts; this should be done by a minimum of two counters. Data sheets should be separate for each counter.

2. The number of *Nephrops* burrow systems will be counted by each minute block.

3. A warm-up count is required for the first minute of each video, and this count is logged. Then a minimum of seven good quality minutes (in which there is sledge ground contact, speed, and visual clarity) should be counted.

4. If counting is resumed after a break of more than one hour, counters should be refamiliarized by reviewing a full ten-minute video before they restart counting (using random image data from the same area).

5. What to count and not count:

- *Only count* burrow systems and partial burrow systems that pass off the bottom flat edge of the monitor and within the FoV.

- *Count* a burrow where any part of the crescent shape of its entrance touches the bottom of the inside FoV (Figure 3.5).

- *Do not count* a burrow system that passes off the side edge of the screen or outside the field of view.

- *Do not count* collapsed or inactive burrow systems.

6. If image quality deteriorates during a one-minute block (e.g. through turbidity, speed, sledge gliding, or technical issues such as lighting,) the portion of uncountable seconds must be recorded. A valid one-minute block must contain at least 30 countable seconds.
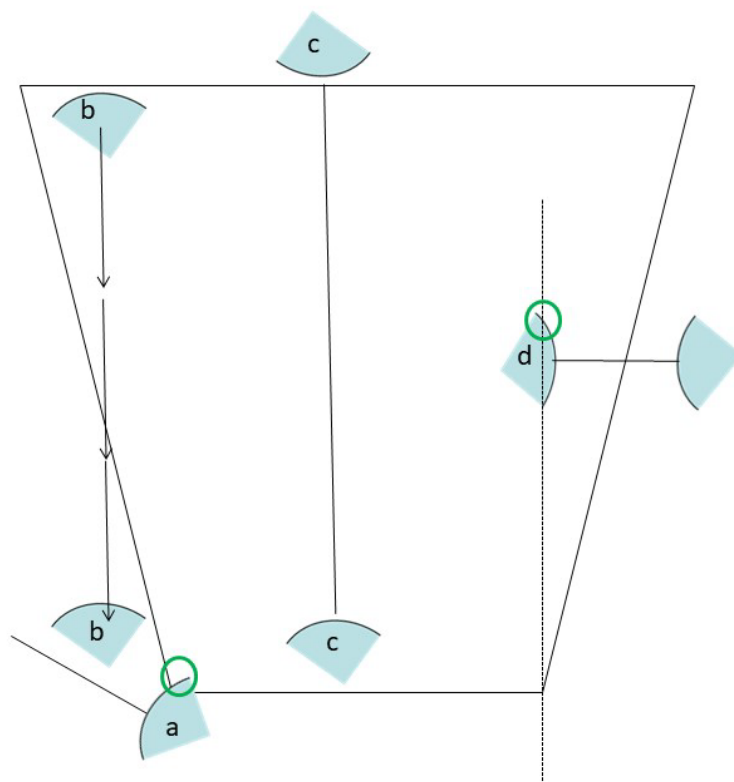
Figure 3.5. An example of different *Nephrops* burrow systems where: (a) any part of the crescent shape of a burrow entrance touches the bottom of the inside FoV and is counted, (b) a burrow passes off the side edge of the screen or FoV and is not counted, (c) a burrow is counted and another burrow from the same system comes later into the FoV and has already been counted, and (d) any part of a burrow entrance falls into the FoV and is counted (d). It should be noted that a slight change in course of the sledge may cause a burrow entrance to fall out of the FoV, in which case that system is not counted. (Photo © Marine Scotland Science)

3.2.3 Protocol for verification of burrow count data

Lin's concordance correlation coefficient (CCC) (Lin, 1989) is used in three ways for the verification of count data from UWTV surveys: firstly, to develop and test the data from a reference set; secondly, as a tool to check the standards of reviewers' counting skills against a reference set; and thirdly, to screen the survey data using an agreed threshold to check the pairwise data. Lin's CCC methodology measures the ability of counters to exactly reproduce each other's counts on a scale of 1 to –1, where 1 is perfect concordance (i.e. a pairwise plot with all points lying along the 1:1 line). This method is suitable for datasets that contain a minimum of ten observations and Poisson distribution such as burrow counts.

3.2.4 Compilation of reference image datasets

Reference sets and counts were introduced and developed at the first ICES *Nephrops* burrow workshop WKNEPHBID (ICES, 2008), where there was a need to standardize and test counting ability. A reference set is a selection of high-quality images (video or stills) representative of burrow densities encountered, the various species interactions on the seabed, and also small-scale topography of a specific survey area. The aim of a reference set is to test the counting skills of each reviewer against the accepted reference counts once the reviewers have received adequate training as noted in Section 3.2.6 for that specific area. Once the reviewer has passed the Lin's CCC threshold they are permitted to review the survey data.

The following lists the standard specifications for compiling a reference set; the list was revised at the recent WKNEPS workshop (ICES, 2018a).

1. All institutes should produce at least nine reference videos each having a duration of ten minutes. The data should be selected to cover low, medium, and high *Nephrops* densities as well as species complexes likely to be encountered in each area. Videos should have generally good quality visibility and have been taken during a station running at optimum speed and with continuous ground contact.
2. Each institute should collate data from their most recent surveys.
3. Upon completion, reference sets are to be made available to all institutes in either DVD format (with each run contained in a separate chapter) or media file format such as MP4.
4. Stations are to be titled one through nine.
5. Reference counts are to be developed and tested using Lin's CCC methods following the guidelines set out by WKNEPS (ICES, 2018a). The R-scripts for developing reference counts can be found on the ICES WGNEPS GitHub site.
6. The reference sets should reflect the current survey conditions and ground characteristics. It is recommended to update the reference sets every time there is a change in gear setup (e.g. FoV, camera position, lasers), any substantial change on the ground, such as in topography, or change in population structure reflected by burrow systems size.
7. There should be sufficient reference sets per area to ensure that they cannot be learned.

3.2.5 Developing reference counts for a survey area

Once a reference set has been collated, count data from this set are to be evaluated using Lin's CCC methods, following the guidelines agreed at the most recent *Nephrops* burrow identification workshop (ICES; 2018a). A decision tree for using Lin's CCC statistics to screen these count data for intra-reviewer and inter-review consistency and then finally to check the robustness of reference counts is available from this workshop. The first step is to test the Lin's CCC values when comparing counts from an individual reviewer. This step is termed the 'intra-performance check' and sets out to ask whether the reviewer can reproduce the same counts. If a reviewer's count data do not pass the set threshold of 0.5 at this step, they are dismissed. Step two is to test the CCC values when comparing the mean counts of each reviewer; this is termed the 'inter performance check'. A threshold of 0.5 is used here along with a rule where if a reviewer fails more than 50% of their pairings then they are dismissed from the process. Reference counts are calculated as the mean of all the remaining reviewers with no weighting applied. It was agreed at the workshop that it was acceptable to calculate reference counts in cases when only two reviewers passed steps 1 and 2. This is similar to the process for the verification of survey count data. Finally, step three is to test the robustness of the reference set using all the reviewers' counts. The count data that passes all the steps in this process will be used as the accepted reference count data. This method of developing reference counts using Lin as a screening tool has been completed for multiple survey areas with a range of densities and interspecies interactions and proved to be a transparent, independent, and robust process.

The R-code for generating the reference counts can be found on ICES WGNEPS GitHub site.

For new and developing UWTV surveys, reference sets from established survey areas which are similar in terms of densities and benthic species interactions can be used while local expertise is being developed.

3.2.6 Training material and procedure

In line with SGNEPS recommendations (ICES, 2010) all scientists must be trained/re-familiarized using training material prior to counting the survey footage. When counting at sea this training should take place in the days prior to embarking on the survey.

1. All institutes must produce training material (which includes the reference sets) for areas covered by their UWTV surveys. It is a minimum requirement for the training material to include two minutes of annotated image data covering the range of density and visibility and information about other burrowing species encountered as well as a photographic guide of signature features of *Nephrops* burrows (ICES 2008, 2016b; Leocádio, *et al.*, 2018)

2. A standard operating procedure (SOP) for training in *Nephrops* burrow identification should be produced by each institute and implemented on commencement of a survey. The SOP provides guidance specific to each institute regarding:
   - pre-survey training;
   - the procedure for using reference footage and associated results;
   - familiarization or warm-up sessions;
   - the correct procedures during sledge deployment to ensure all required data is recorded;
   - how to correctly identify *Nephrops* burrows, including clear photographs of *Nephrops* burrow systems as well as examples of other non-*Nephrops* burrow systems to highlight the differences;
   - reviewing techniques;
   - examples of the correct recording templates to use;
   - if appropriate, abundance keys for non-*Nephrops* species.

   This SOP should be followed by all the survey team.

3. New members of staff should have individual training with experienced counters.

4. After covering the training material, counters must go through the reference set and validate their counts using Lin's CCC (with a minimum threshold of 0.5) prior to counting the survey footage (ICES, 2008; Leocádio, *et al.*, 2018). The R-code and example files for checking counter performance against a reference set can be found on ICES WGNEPS GitHub site.

5. Staff and protocol exchange between institutes is recommended where possible. On collaborative surveys such as those covering FUs 14 and 15 this exchange should occur annually.

For those collaborative surveys where it is not possible to have staff exchange on board it is recommended that a minimum of 15 % of stations are reviewed by each participating institute.

3.2.7 Quality control of survey count data

After all counts are completed, additional independent or consensus counting must be done in order to correct discrepancies in the data.

For an independent counting procedure Lin's CCC method (with a minimum threshold of 0.5) should be used to check which stations will need to be reviewed again. If the paired count data for a station does not pass then a third reviewer will need to count this station. From experience when a station does not pass the threshold on the third review it is recommended that the scientist in charge checks the counting process and video quality. Examples and R-code can be found at ICES WGNEPS GitHub site.

Consensus counting has been used in the past, but it is only recommended for training purposes or in developing surveys.

It is recommended for surveys on very low density grounds (e.g. < ten burrows per minute) such as FUs 1 and 16 that all burrows are timestamped or annotated by two reviewers independently. Annotated burrow counts are compared and when they differ they are checked and a consensus agreed upon by both reviewers, preferably involving a third reviewer.

If there are substantial unexplained fluctuations between abundance estimates (e.g. more than 40% change) from previous years or within a previous year, it is recommended to review a random 20% selection of the most recent survey stations. If the partial review counts differ by more than 20% from the survey counts, a full review of the most recent survey should be considered. This is a rule of thumb to aid in determining that there are valid interannual variations not caused by counting trends.

Quality control plots should be used to verify the data collected on a station by station basis. Figure 3.6 is an example of a quality control plot used to screen the survey counts with Lin's CCC analysis. In this example, station 25 requires a third counter to review the image data where the paired count data do not pass the threshold limit of 0.5 (left panel) and where the points do not lie on a straight line and also the trend for the reviewers is in opposition for minute blocks 1, 4, 6, and 7 (right panel). Station 2 passes the minimum threshold where the data lies close to the line (left panel) and the counting trend is very similar (right panel).
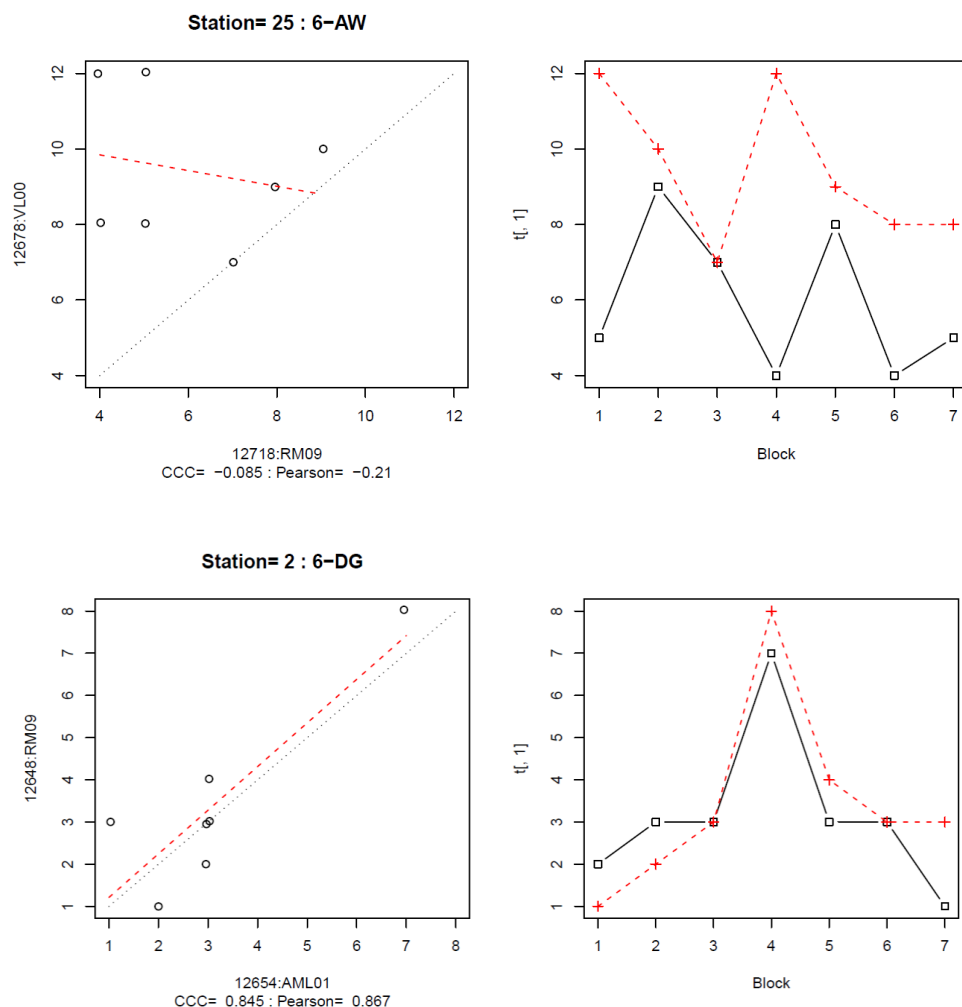
Figure 3.6. Quality control plots of survey counts using Lin's analysis. The left panel shows a pairwise plot of the count data and the red dashed agreement line and the output value of Lin's CCC and Pearson correlation. The right panel shows a line plot of the count data by minute block for each reviewer, where the solid and dashed line each denote a specific reviewer.

3.2.8 Ancillary information

With the available resources (staff and time) it may be possible for national laboratories to agree to the collection of ancillary information during the survey or while reviewing the image data.

This additional data should include quality checks on the TV track performance such as visibility, ground contact, speed of the sledge, cloud and any other interference, all of which could be recorded for each one-minute interval using a classification key.

While reviewing *Nephrops* activity in and out of the burrows, ancillary data could be collected on the presence of trawl marks, marine litter, fishing gears, and the occurrence of fauna. As an example, there is an OSPAR special request to record sea pen species (*Virgularia mirabilis*, *Funiculina quadrangularis, and Pennatula phosphorea*) using a key devised to categorize the density (ICES, 2011). As a minimum, collecting data on the presence or absence data of these sea pen species is achievable.

Annex 3 presents an example of an Excel datasheet that can be used for recording UWTV observations during image review.

4 Caveats

A number of factors are believed to contribute bias to UWTV survey estimates of *Nephrops* abundance. In order to use a survey abundance estimate as absolute it is necessary to correct for these potential biases. The bias estimates are based on simulation models, preliminary experimentation, and expert opinion. The values for specific *Nephrops* grounds are found in Leocádio *et al.* (2018) and are updated during benchmark workshops, e.g. when new information has become available.

4.1 Edge effect

Factors that are significant in determining the scale of edge effects are the width of viewed track, the FoV, and the size of a *Nephrops* burrow system. A formula to calculate the edge effect is proposed by Campbell *et al*. (2009) and given below (Equation 1):

$$Z = \zeta \; \frac{\pi \left(\frac{(L-S)}{\cos(C-0.5V)} \right) \tan(0.5H)}{\pi \left(\frac{(L-S)}{\cos(C-0.5V)} \right) \tan(0.5H) + D} \qquad \text{Eq. (1)}$$

V – Vertical angle (radians)
H – Horizontal angle of field of view of camera lens (radians)
C – Angle camera is mounted at on sledge (radians)
L – Height of camera lens from bottom of sledge (m)
S – Sinkage of sledge into seabed (m)
D – Median burrow length (m) (assuming burrows are randomly oriented with
    respect to the direction of the towed track)
Z – Corrected burrow density estimation
ζ – Original burrow density estimation

Further details on the caveats with respect to burrow detection, burrow identification, and occupancy are given in Leocádio, *et al.* (2018).

5 Survey data analysis

5.1 Data quality control

It is essential that quality checks are continuously undertaken during the survey, such as 1) regular checks of sensors on sledge, 2) monitoring of "live" video quality, 3) monitoring of distance data quality, and 4) FoV and survey count data using R-scripts (see output examples in Leocádio, *et al.* (2018)).

5.2 Geostatistical methods

The spatial covariance and other spatial structuring and geostatistical analysis of the mean and variance can be carried out using RGeostats (Petitgas *et al.*, 2017; Renard *et al.*, 2018). Kriging geostatistical procedure is used for several *Nephrops* stock assessments that use a grid design to provide abundance estimates and estimation uncertainties. FU 17 (Aran Grounds) is fully documented in the stock annex in R markdown document (ICES, 2016c). Other software that can be used for geostatistical methods include Surfer and ArcGIS; these, however, do not provide the kriged estimation variance or CV. EVA software - Estimation VAriance software provides CV on kriged estimates (Petitgas and Lafont, 1997).

The first step in the work up of the survey data is to translate the individual count data into densities (number per m^2) by using the estimate of viewed area (on a minute-by-minute basis). An average density estimate for each minute is then calculated by averaging over all counters. Station density is then calculated as the average of densities over all minutes, weighted by the area viewed within each minute.

5.3 Random stratified methods

Within the workup of data collected under a random stratified survey design, stations are then grouped by strata, and for each stratum the mean density is raised to the area of the stratum and the stratum variance of the calculated mean. The estimates are then summed across the strata to generate an estimate of the overall abundance, variance, and relative standard error (ICES, 2007).

Precision improves as the number of stations is increased, but it also depends on the underlying variability in the density estimates. The minimum number of stations (*n*) required to achieve a particular relative error (r) can be defined as:

$$n = \frac{1}{\frac{r^2 u^2}{z^2 \sigma^2} + \frac{1}{N}} \qquad \text{Eq. (2)}$$

where *u* is the estimated average burrow density and σ^2 the estimated variance (Thompson, 2002). *N* is the total number of stations that would have to be surveyed to cover the entire ground this value is very large and therefore 1/*N* is very small and can be dropped from the equation (The *z*-score for a particular significance level is also required). Equation 2 can be rearranged to give an expression for relative error in terms of number of stations.

5.4 Sampling effort

Data from previous surveys can be used to give an indication of how the number of survey stations should be adjusted to obtain a target precision level in future surveys. Investigations of optimal sampling effort to achieve high precision levels by exploration of variograms and

station spacing for geostatistical kriging method and bootstrapping approaches of stratified random method are detailed in ICES (2007, 2012).

Density estimates should be inspected by strata to identify areas where additional samples could improve precision estimates and vice-versa if there are areas in which sampling intensity could be reduced. The appropriateness of the survey stratification and the boundary definition of the survey area should be assessed on a regular basis.

6   Reporting of results

Reporting from the survey should include, at minimum, the following graphics:

- a bubble plot of adjusted density for each station by year and/or contour plots of the kriged adjusted density estimates by year;
- a violin and box plot of adjusted burrow density distributions by year;
- a time-series plot of total abundance estimates with error bars (indicating 95% confidence intervals) and $B_{trigger}$ reference point if proposed (ICES, 2018b).

Annex 4 contains example of maps and figures and R-script. Example files for checking counter performance against a reference set can be found on ICES WGNEPS GitHub site.

6.1   Survey summary sheet

A survey summary sheet should be provided for each survey. The following may serve as an example:

Country		Vessel Name	
Survey Code		Dates (start/end)	
FU / Ground Name			
Number of staff			
N. of staff exchange			

Objectives

Survey Design	
Number of stations (planned/completed)	
Average Field of view (cm)	
Distance over ground – source used	
% stations with trawl marks	
Deviations from the survey plan (e.g. coverage/weather related problems, technical problems, potential biases, etc.)	
Final adjusted mean density	
Final adjusted abundance and CV	
Overall footage quality (poor, medium, good)	
Reference footage for survey area generated	
Quality control of station counts (Lin's CCC or consensus count)	
Other survey activities (CTD, Trawl, sediment samples, sediment profile images, etc.)	
Map:	

7   References

Aristegui, M., O'Brien, S., Blaszkowski, M., O'Connor, S., Fitzgerald, R., and Doyle, J. 2018. The "Smalls" *Nephrops* Grounds (FU22) 2018 UWTV Survey Report and catch scenarios for 2019. Marine Institute UWTV Survey report. https://oar.marine.ie/handle/10793/1376.

Campbell, N., Dobby, H., and Bailey, N. 2009. Investigating and mitigating uncertainties in the assessment of Scottish *Nephrops norvegicus* populations using simulated underwater television data. ICES Journal of Marine Science, 66: 646–655. https://doi.org/10.1093/icesjms/fsp046.

Campbell, N., Allan, L., Weetman, A., and Dobby, H. (2009). Investigating the link between *Nephrops norvegicus* burrow density and sediment composition in Scottish waters. ICES Journal of Marine Science. 66. 2052–2059. https://doi.org/10.1093/icesjms/fsp176.

ICES. 2007. Report of the Workshop on the use of UWTV surveys for determining abundance in *Nephrops* stocks throughout European waters (WKNEPHTV). 17–21 April 2007, Heraklion, Crete, Greece. ICES CM: 2007/ACFM: 14 198 pp.

ICES. 2008. Report of the Workshop and training course on *Nephrops* burrow identification (WKNEPHBID), 25–29 February 2008, Belfast, Northern Ireland, UK. ICES CM 2008/LRC:03. 44 pp.

ICES. 2009. Report of the Benchmark Workshop on Nephrops (WKNEPH), 2–6 March 2009, Aberdeen, UK. ICES CM 2009/ACOM:33. 156 pp.

ICES. 2010. Report of the Study Group on *Nephrops* Surveys (SGNEPS), 9–11 November 2010, Lisbon, Portugal. ICES CM 2010/SSGESST:22. 95 pp.

ICES. 2011. Protocols for assessing the status of sea-pen and burrowing megafauna communities *In* Report of the ICES Advisory Committee, 2011. ICES Advice 2011, Book 1, Section 1.5.5.3. 11 pp.

ICES. 2012. Report of the Study Group on *Nephrops* Surveys (SGNEPS), 6–8 March 2012, Acona, Italy. ICES CM 2012/SSGESST:19. 36 pp.

ICES. 2016a. Interim report of the Working group on *Nephrops* Surveys (WGNEPS), 7–8 November 2016. Reykjavík, Iceland. ICES CM 2016/SSGIEOM:32. 69 pp.

ICES. 2016b. Report of the Workshop on *Nephrops* burrow counting (WKNEPS), 9–11 November 2016. Reykjavík, Iceland. ICES CM 2016/SSGIEOM:34. 62 pp.

ICES. 2016c. , June –September 2015, By correspondence. ICES CM 2015/ACOM:38. 86 pp.

ICES. 2018a. Report of the Workshop on *Nephrops* burrow counting (WKNEPS), 2–5 October 2018, Aberdeen, UK. ICES CM 2018/ EOSG:25.47 pp.

ICES. 2018b. ICES Special Request Advice Celtic Seas Ecoregion. EU request to ICES to provide plausible and updated FMSY ranges for the stocks of species inhabiting western EU waters. March 2018. https://doi.org/10.17895/ices.pub.4149.

Leocádio, A., Weetman, A., and Wieland, K. (Eds). 2018. Using UWTV surveys to assess and advise on *Nephrops* stocks. ICES Cooperative Research Report No. 340. 49 pp. http://doi.org/10.17895/ices.pub.4370.

Lin, L. I. 1989. A concordance correlation coefficient to evaluate reproducibility. Biometrics, 45(1): 255–268. https://doi.org/10.2307/2532051.

Petitgas, P. and Lafont, T. 1997. EVA2: Estimation VAriance. Version 2. A geostatistical software on windows 95 for the precision of fish stock assessment surveys. ICES CM 1997/Y:22. 22 pp.

Petitgas, P., Woillez, M., Rivoirard, J., Renard, D., and Bez, N. 2017. Handbook of geostatistics in R for fisheries and marine ecology. ICES Cooperative Research Report No. 338. 177 pp. https://doi.org/10.17895/ices.pub.3717.

Renard D., Bez N., Desassis N., Beucher H., Ors F., Freulon X., Petitgas, P., and Woillez, M. 2018. RGeostats: The geostatistical package version:11.2.6. MINES ParisTech. Free download from: http://cg.ensmp.fr/rgeostats.

Thompson, S. K. 2002. Sampling. Second edition. John Wiley and Sons, New York, New York, USA.

Annex 1: Template of sledge design

Figure A.1. Outline of plans for the original sledge design, variations of which are being used by several institutes today.

Annex 2: Template data sheet for UWTV station

The example sheet below can be used to record 1) metadata on the UWTV track during recording and 2) the cable being paid out or in at the end of each one-minute block to correct vessel distance. Refer to .

Nephrops Underwater TV Survey Data Sheet				
SURVEY CODE:	**DVD Number**		Entered	
			Validated	
Station#:	**Time UTC**24hr			
Date: / /	**Start:**		**Depth** (m)	
Area:	**Time Recorded**			
FU:		**Cable Out**	**SOL**	
			EOL	

Only record if counting for full 10 mins useful in case of data delete

Observer ID:					
Burrow Count:					
***Nephrops* In:**					
***Nephrops* Out:**					
Fish Present:					
Trawl Marks:					

Notes on UWTV track:
Speed:
Visual Clarity:
Ground Contact:
USBL:
Lasers:
Multibeam-BS:
Biological Features:

Min	Valid Minute 30 secs countable	Cable Correction If SHIP nav only
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		

QUALITY CHECK after station complete:		STATION RATING:	
SHIP NAV DATA:			
USBL NAV DATA:		**Excellent**	
FOOTAGE DATA:		**Good**	
		Ok	
		Ok-poor	
		Poor	
		Redo	

Grab Sample				
Depth (m):		**Sediment type**		
Notes:				

Annex 3: Template UWTV data sheet for recording observations

The example Excel sheet below can be used to record counts of burrows, *Nephrops* observed inside and out of burrows, and the number of seconds within a one-minute block where counting is not possible. Ancillary data can also be recorded, such as qualitative records on ground type, visual clarity, sledge speed and behaviour, the presence of trawl marks, and the presence of fish as well as the identification where possible of sea pen species and other marine life.

Nephrops underwater TV survey recount data sheet

DVD Number: _____   SURVEY CODE: _____

Sea pens: O=<1   F=2-19   C=>20   FAQ/PNP/VAM

Station: _____   Date/Time: ____   Area _____

Notes:

ANCILLARY DATA

ID: Min	Burrow Count	NEPHROPS		Seconds not counted		Ground Type	Visual Clarity	Sledge Speed	FISH			Other Life	Trawl Marks
		In	Out	STOP	START				Gadoids	Flats	Others		
1													
2													
3													
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14													
15													

Annex 4: Recommended maps and plots for UWTV survey report

Examples of density maps, krigged contour plots, violin plots, and total abundance estimates plots for UWTV survey reports. The R-code and example files for these plots can be accessed at the WGNEPS GitHub site.
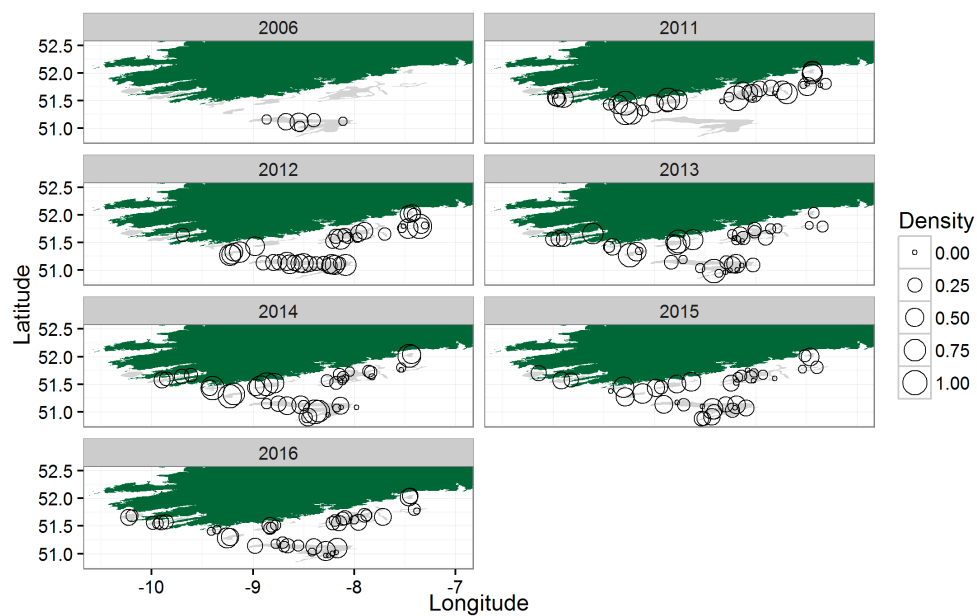

Figure A4.1. FU 19 grounds: Bubble plot of the density for each station by year.

Figure A4.2. FU 20–21 grounds. Contour plots of the krigged density estimates by year from 2013 (top left) to 2016 (bottom right).
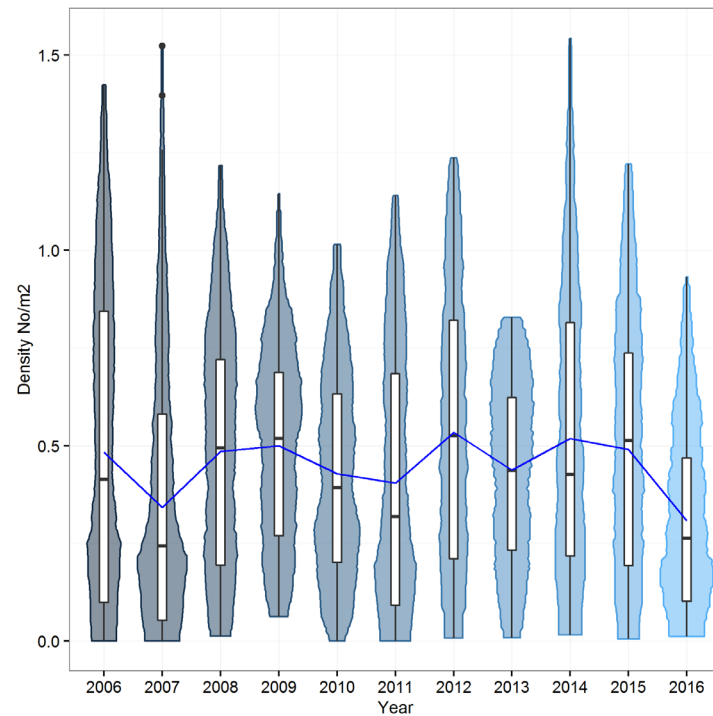
Figure A4.3. FU 22, Smalls Grounds. Violin and box plot of adjusted burrow density distributions by year from 2006–2016. The blue line indicates the mean density over time. The horizontal black lines represent the median values, the white boxes are the inter-quartile ranges, the black vertical lines are the ranges, and the black dots are outliers.
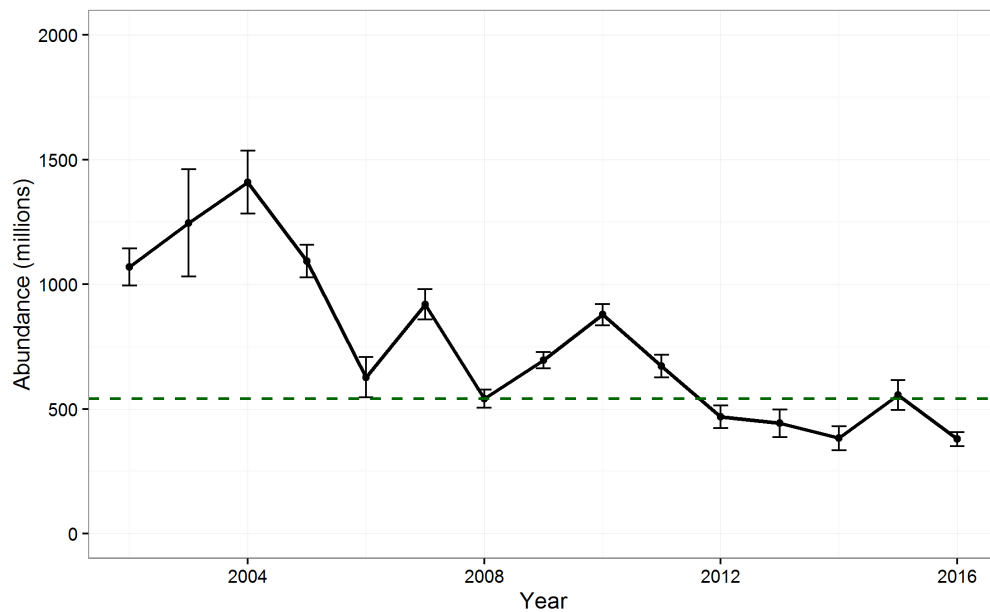

Figure A4.4. FU 17 Aran Grounds. Time-series of total abundance estimates for FU 17 (error bars indicate 95% confidence intervals), and MSY reference point B$_{trigger}$ is represented by the dashed green line.

Annex 5: Abbreviations and technical terminology

AFBI	- Agri-Food and Biosciences Institute
CEFAS	- Centre for Environment, Fisheries and Aquaculture Science
CNR-IRBIM	- National Research Council Institute for Biological Resources and Marine Biotechnologies.
DTU - AQUA	- Technical University of Denmark, National Institute of Aquatic Resources
CRT	- Cathode ray tube
CTD	- Conductivity, temperature, depth
CV	- Coefficient of variationDP - Dynamic positioning system
DVD	- Digital versatile disk
DV tape	- Digital video tape
EXIF	- Exchangeable image file
FoV	- Field of view
FU	- Functional Unit
GPS	- Global positioning system
GSA	- Geographical Sub-Area
ICES	- International Council for the Exploration of the Sea
IEO	- Instituto Español de Oceanografía
IFREMER	- L'Institut Français de Recherche pour l'Exploitation de la Mer
IZOR	- Institute of Oceanography and Fisheries Croatia
LED	- Light-emitting diode
Lin's CCC	- Lin's concordance correlation coefficient
MFRI	- Marine and Freshwater Research Institute
MI	- Marine Institute
MP4	- Digital multimedia container format
MSFD	- Marine Strategy Framework Directive
MSS	- Marine Scotland Science
OSPAR	- The Convention for the Protection of the Marine Environment of the North-East Atlantic (OSPAR Convention)
NAS drive	- Network attached storage
ROV	- Remotely operated vehicle
SGNEPS	- Study group on *Nephrops* surveys
SLU	- Swedish University of Agriculture Science
SOP	- Standard operating protocol
TV	- Television
USBL	- Ultra short baseline
WGNEPS	- Working Group on *Nephrops* Surveys
WKNEPHBID	- Workshop and training course on *Nephrops* burrow identification
WKNEPHTV	- Workshop on the use of UWTV surveys for determining abundance in *Nephrops* stocks throughout European waters
WKNEPS	- Workshop on *Nephrops* burrow counting
UWTV	- UnderWater TeleVision

Annex 6: Author contact information

Jennifer Doyle (editor)

Marine Institute
Rinville
Oranmore
Co. Galway
Ireland
H91 R673
Email: jennifer.doyle@marine.ie

Patrik Jonsson (editor)

Swedish University of Agricultural Sciences (SLU)
Department of Aquatic Resources (SLU Aqua)
Lysekil
Sweden
Email: patrik.jonsson@slu.se

Adrian Weetman (editor)

Marine Scotland Science
375 Victoria Road
Torry
Aberdeen
Scotland, UK
AB11 9DB
Email: Adrian.Weetman@gov.scot

Helen Dobby

Marine Scotland Science
375 Victoria Road
Torry
Aberdeen
Scotland, UK
AB11 9DB
Email: Helen.Dobby@gov.scot

Jónas Jónasson
Marine and Freshwater Research Institute (MFRI)
Fornubudum 5
220 Hafnarfjörður
Iceland
Email: jonas.jonasson@hafogvatn.is

Ana Leocádio

European Commission - Brussels
DG Maritime Affairs and Fisheries
Rue Joseph II 99, 1000 Bruxelles, Belgium
Email: ana.leocadio@ec.europa.eu

Colm Lordan

Marine Institute
Rinville
Oranmore
Co. Galway
Ireland
H91 R673
Email: colm.lordan@marine.ie

Kai Weiland

Technical University of Denmark
National Institute of Aquatic Resources (DTU AQUA)
Hirtshals
Denmark
Email: kw@aqua.dtu.dk