**Sub-group meeting ICES database and StoX 17/01/2018**

**Data base Submissions**

* Data submission process is straight forward
* 2016 data – All check if data is correct and upload corrected data
  + Download 2016 data and delete the data off ICES data-base – deadline 31 January.
  + New correct 2016 data uploaded for 30 April
* Norwegian data – will get into the system but waiting for developers to process XML database files. Cecilie will see if they can meet the deadline of 30 April for 2017 and 2016 data.
* Old data should be uploaded over the next couple of years as we progress. A commitment in the HERAS group to upload retrospective data but not with a firm deadline as instituts might struggle to follow developments. Also it will differ between countries how far back we can go reliably. Need to include text in report on commitment to upload data back in time.

**Validation**

* Some frustration with the validation process for us novices. ICES datacentre are aware of this and are working to make the system more user friendly for non-experts.
* Sometimes hard to see where the message comes from because line numbers refer to the lines in the XML files that the data is transformed into before validation. ICES is aware and working fixing this.
* Some of the validation rules are maybe not optimal for all surveys. User groups are encouraged to inform data centre if some rules are not relevant for their survey.
* Also, if a survey groups thinks of some good rules to implement to get in touch with data centre and they will help.
* ICES datacentre encourage users to get in touch when getting cryptical error messages so they can be aware and try to make them more user friendly. They can also help chasing down what the error is.
* Some specific issues discussed
  + Problematic that trawling distance errors give a critical warning/error
  + Validation based on trawl position distance from transect – not good idea as a general rule. In some of the HERAS survey strata trawling is carried out away from the transect, as the area is sometimes revisited for night time /day time trawling. Using the map as a visual check after by the uploader would is a better check.

**Map**

* Quite a lively discussion as this is a much liked feature. Quite a few suggestions for improvements/added features:
  + Ability to select species and have hauls with and without the species from the survey displayed in different colours to indicate where the species was caught and not. Show all hauls in survey, if a species is selected show haul with that species in a different colour.
  + Bubble-plot option was discussed – showing Sa values for selected species as bubble plot.
    - However: Agreed that there is scope for misunderstanding if someone not familiar with the surveys use this functionality to display for example herring distribution and are not aware that herring is “hidden” also in the NASC categories MIX and CLU.
  + Ability to download the displayed map as an image would be very useful for reports for example.

**CTD stations – hydrographic data**

* Encourage all cruises to upload data to ICES hydrographic data base and integrate with plotting in acoustic data portal. GER, NOR, DK, IE already upload this data for the HERAS survey. Scotland will look into feasibility.
* Visualize hydrographic data in the acoustic map interface
  + Surface temperature from satellite data for the survey period
  + Chlorophyll maps for the survey period from satellite also
  + Ability to click on a survey CTD position and get a display of the temperature and salinity profile displayed.
  + Seabed temperature map from survey CTD data

**User management**

* Different level of users is a good idea.
* Good thing to have restricted access to delete data not uploaded by yourself.

Biodata for all species – should we upload other species. We did not discuss this. Susan’s opinion: Depends on how consistently catch is fully identified and catchweight obtained for all components for all years. Is this indicated at the survey level in the database for example? (can you assume presence / absence from the available data for all species /countries / years? If not this needs to be specified somewhere in the database. Maybe should state that trawl data for a survey can only be considered “complete” for Survey target species.

**Surveys designation**

ICES to make it a possibility for multiple survey codes associated with a cruise – using the acronyms for surveys. Still need to “split” a physical cruise into two cruise upload files/units if the designation changes half way for example. WESPAS is a good example where the first part of this cruise is just WESPAS, second part however is both WESPAS and HERAS. Solution is to upload part 1 and 2 of this cruise as two cruises where part one is tagged only with WESPAS and part two is tagged as WESPAS and HERAS. The cruise will appear in search for either then.

**Procedure**

Data that is the database should be the most recent and that which is used in the production of indices. We need to commit to this and make sure there is not a level of data manipulation happening (fixes of errors for example) on data outside of the database, but that when errors are discovered a corrected file is uploaded to the database.

**Versioning**

Did not discuss this a lot, but it was generally consider sensible that once a cruise has been used to generate an index for use in an assessment that copy of the cruise should be “locked” and given a version. The cruise can be opened up again if errors have been discovered and fixed, but a copy of the version used in an assessment etc. will be kept as a separate version.