Work Package 2
Deliverables - D 2.4

Electronic data bases of biological information on marine samples (D 2.4)

(Month 30)

SALSEA-WGNAPES data base description

Rationale
The WGNAPES data base format was originally developed for the joint international survey for blue whiting and herring in the Northeast Atlantic in the mid 1990s. It is built over a Microsoft Access platform by the Faroe Marine Research Institute, FAMRI, (formerly the Faroese Fisheries Laboratory, Partner 15), facilitating the collection and organisation of data and ensuring the quality and integrity of the dataset. All data collected during the surveys were entered into the WGNAPES database (logbook, catch or trawl data, biological data, plankton, hydrography data, and also acoustic data –not shown here).

With the scope of storing the SALSEA cruise data (described in WP2) from the national cruise databases held by the individual partners (1,2 and 15) in a shared database, several options were discussed. The involved partners then chose the WGNAPES. This database format was considered suitable to expand with data from WPs 4 and 5, and could be linked up to the database created by the genetic Work Package 3. The group decided that FAMRI (15) at Faroes should test and adapt the WGNAPES format to SALSEA requirements for storing and sharing the biological and physical data collected on the expeditions. The SALSEA WGNAPES builds on the original WGNAPES data base, but is now designed to store all biological and pysical data generated through SALSEA Merge.

Description
The "SALSEA _WGNAPES_data" is the main database with logbook, biology and sampling data for each fish. All scientific data and analysis on each post-smolt is linked up to this main database, even if the data are stored in separate databases. In the project two workpackages were working in parallel, one on the scale growth analysis, and one on DNA assignments. These two datasets were stored in separate databases. The "SALSEA _WGNAPES_scales" database holding data on circuli growth analysis (SALMON_Scales and SALMON_Circuli_growth tables), and the "SALSEA_WGNAPES_dna" database holding data on the DNA assignments (SALMON_dna table). These two additional databases are linked up to the main "SALSEA_WGNAPES_data" database. (Cfr Flow chart below)

The SALSEA WGNAPES database is located on a SALSEA share point site hosted by Marine Laboratory Scotland.

Use of the databases
To be able to combine and use the data in all databases the link must be properly designed. A "unique key" is the solution to create useful queries from the database. Basically the original design was used to create a unique key by combining several parameters in each table, repeat those in each table, and force referential integrity in the "key" between the tables and databases. Thus all three databases have a common set of keys that make up the unique link between them (they are the first few parameters), and all "offspring" tables have one or more parameters to uniquely link each record up to the "parent" table. E.g. each fish in the "Biology" table gets a unique key by using Country, Cruise, Species keys plus FishNo.
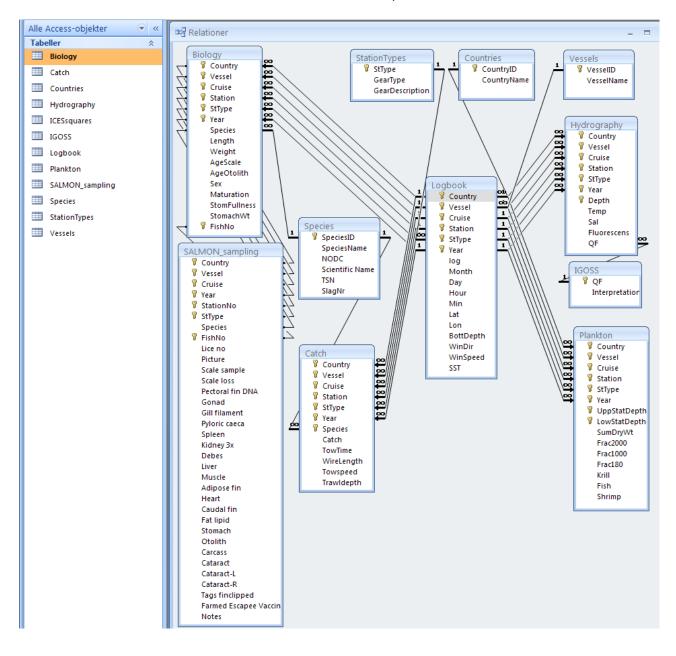
Country, Vessel, Cruise, Station, StType, Year (Logbook and all other tables), Species (Biology, Scales, dna, ...), FishNo/GrowthNO (Biology, Scales), FishNo/SampleID (Biology, dna).

To query the database it is necessary to create a new empty Access database and link the various tables to be used from the main "SALSEA _WGNAPES_data" into this new "query" database. All

three databases should be in the same directory to make the link process simpler. However, to use the "scale" or "dna" databases, it should be enough to open copies of the databases. In those the necessary tables from the main "data" database are already linked up to facilitate selects (extracts).
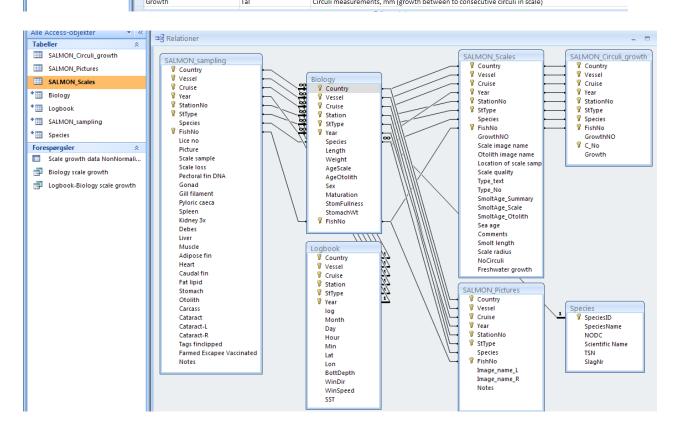
When data are extracted/selected from the database to be imported to an excel spreadsheet, it is important always to include *all* the key parameters in the select, in order to be able to link the excel data back to the database, if needed.

Below is a description/flow chart of the relationships between the tables in the three databases. Further, it is possible to see which tables are linked from the main database into the two additional databases (scales and dna). An arrow (left arrow) pointing to the table name in the "table window" in Access indicate that the table is linked (and thus holds the original data from the main database).

SALSEA _WGNAPES_data: Table relationships (Access). This is the main database with Logbook and Biology data for each fish. The two "scales" and "dna" databases link up to the current "data" database.
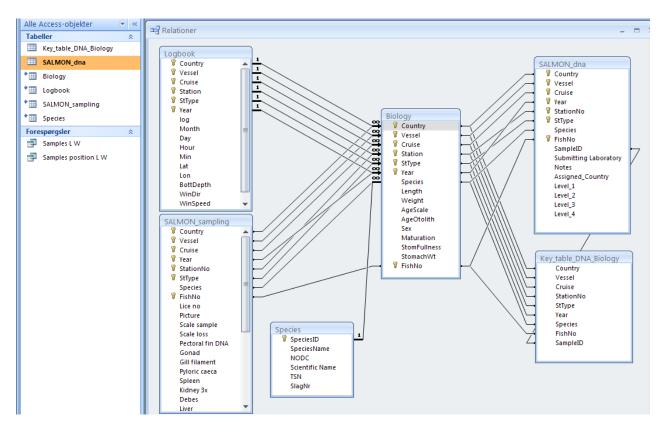
SALSEA _WGNAPES_scales: Relationships between the SALMON_Biology table and the added scale-growth tables (SALMON_Scales, SALMON_Circuli_growth).

Tabeller
- SALMON_Circuli_growth
- SALMON_Pictures
- SALMON_Scales
- Biology
- Logbook
- SALMON_sampling
- Species

SALMON_Scales

Feltnavn	Datatype	Beskrivelse
Country	Tekst	Post code (or ICES list)
Vessel	Tekst	Call sign (or ICES list)
Cruise	Tekst	Unique cruse identifier
Year	Tal	Four digits
StationNo	Tekst	National station number, link to Station in BIOLOGY
StType	Tekst	CTD, WP2, MIK, TRAWL, PDYP, PBLÅ, PTRAWL, …, link to StType in StationTypes
Species	Tekst	3 character code, libk to SpeciesID in SPECIES
FishNo	Tekst	Fish/record number, link to Recnr in BIOLOGY
GrowthNO	Tekst	Sample number used for growth analyses (key)
Scale image name	Tekst	Name of the file where the image of the scale is stored
Otolith image name	Tekst	Name of the file where the image of the otolith is stored
Location of scale sample	Tal	A number between 1 and 22, indicating the location of the fish where the scale has been taken. 999 = ur
Scale quality	Tal	A number between 1 and 3, indicating the quality of the scale for detailed growth analyses, focusing on
Type_text	Tekst	Wild, farmed (escapee from the fish farm industry), stocked (for stock enhancement or sea ranching), re
Type_No	Tal	Similar to Type_text, but in numbers: wild = 1; farmed = 2; stocked = 3; reared = 4; unknown = 999
SmoltAge_Summary	Tal	Smolt age (integers only) derived from all information available. In cases of discrepancy between scale
SmoltAge_Scale	Tekst	Smolt age derived from scales. If uncertain, describe the uncertainty
SmoltAge_Otolith	Tekst	Smolt age derived from otoliths. If uncertain, describe the uncertainty
Sea age	Tal	Number of winters at sea
Comments	Tekst	Any comments relevant to age and growth analyses
Smolt length	Tal	Backcalculated smolt length from the scale
Scale radius	Tal	Total length of the scale from origo to the scale edge
NoCirculi	Tal	Number of circuli in the marine part of the scale

SALMON_Circuli_growth

Feltnavn	Datatype	Beskrivelse
Country	Tekst	Post code, link to Country in BIOLOGY
Vessel	Tekst	Call sign, link to Vessel in BIOLOGY
Cruise	Tekst	Unique cruse identifier, link to Cruise in BIOLOGY
Year	Tal	Four digits, link to Year in BIOLOGY
StationNo	Tekst	National station number, link to Station in BIOLOGY
StType	Tekst	CTD, WP2, MIK, TRAWL, PDYP, PBLÅ, PTRAWL, …, link to StType in StationTypes
Species	Tekst	3 character code, libk to SpeciesID in SPECIES
FishNo	Tekst	Fish/record number, link to Recnr in BIOLOGY
GrowthNO	Tekst	
C_No	Tal	Circuli number, C1 is the first circuli growth in seawater
Growth	Tal	Circuli measurements, mm (growth between to consecutive circuli in scale)

SALSEA_WGNAPES_dna: Relationships between the SALMON_Biology table and the added DNA assignments table (SALMON_dna).

Feltnavn	Datatype	Beskrivelse
Country	Tekst	Post code
Vessel	Tekst	Call sign
Cruise	Tekst	Unique cruse identifier
Year	Tal	Four digits
StationNo	Tekst	National station number, link to Station in BIOLOGY
StType	Tekst	CTD, WP2, MIK, TRAWL, PDYP, PBLÅ, PTRAWL, ..., link to StType in StationTypes
Species	Tekst	3 character code, libk to SpeciesID in SPECIES
FishNo	Tekst	Fish/record number, link to Recnr in BIOLOGY
SampleID	Tekst	Sample number used in local database (key)
Submitting Laboratory	Tekst	Name of the laboratory submitting the dna sample
Notes	Tekst	
Assigned_Country	Tekst	Assigned country
Level_1	Tekst	regions
Level_2	Tekst	regions, 2- level
Level_3	Tekst	regions, 3- level
Level_4	Tekst	regions, 4- level

SALSEA _WGNAPES database table description

Parameters in bold indicate **primary key variables**, and used together they form a **unique key** from the logbook to the other tables. Table names highlighted in yellow.

Logbook:

Country	Post code, 2 chars according to countries table
Vessel	Call sign, 2 or 6 digits acc. to Vessels table
Cruise	Cruise identifier
Station	National station number
StType	Geartype/activity: one line per activity at the same station: National definition of station type
Year	YYYY (4 digits)
Log	Value from the acoustic log (Nm)
Month	MM
Day	DD
Hour	HH, timeUTC 0-24
Min	MM
Lat	Decimal degrees, negative latitude south 0° "0.0000"
Lon	Decimal degrees, negative longitude west of 0° "0.0000"
BottDepth	Bottom depth (m)
WinDir	Compass degrees
WinSpeed	m/s

Hydrography:

Country	Post code, 2 chars according to countries table
Vessel	Call sign, 2 or 6 digits acc. to Vessels table
Cruise	Cruise identifier
Station	National station numbers
StType	Geartype/activity: National definition of station type
Year	YYYY (4 digits)
Depth	Depth of measurement (m)
Temp	°C (at least 2 decimals)
Sal	Salinity (psu, at least 3 decimals)
QF	Quality of salinity data: 0-5 (IGOSS quality flags)

Country	Post code, 2 chars according to countries table
Vessel	Call sign, 2 or 6 digits acc. to Vessels table
Cruise	Cruise identifier
Station	National station numbers
StType	Geartype/activity: National definition of station type
Year	YYYY (4 digits)
UppStatDepth	Upper station depth (m)
LowStatDepth	Lower station depth (m), if only one depth then same as upper
SumDryWt	Plankton mg dry weight/m^2 in each interval
Frac2000	Size graded values, 2000 my sieve
Frac1000	1000 my sieve
Frac180	180 my sieve
Krill	From 2000 my sieve
Fish	-"-
Shrimp	-"-

Catch:

Country	Post code, 2 chars according to countries table
Vessel	Call sign, 4 or 6 digits acc. to Vessels table
Cruise	Cruise identifier
Station	National station numbers
StType	Geartype/activity: National definition of station type
Year	YYYY (4 digits)
Species	Species code: HER, WHB, MAC, SAL, …
Catch	Kg
Towtime	Minutes
Wirelength	(m)
TowSpeed	Knots
Trawldepth	(m)

Biology:

Country	Post code, 2 chars according to countries table
Vessel	Call sign, 4 or 6 digits acc. to Vessels table
Cruise	Cruise identifier
Station	National station numbers
StType	Geartype/activity: National definition of station type
Year	YYYY (4 digits)
Species	Species code: HER, WHB, MAC, SAL, …
Length	cm with one decimal (dot as decimal sign)
Weight	G
AgeScale	Year from scale readings
AgeOtholit	Year from otolith
Sex	Empty means not sexed, 1= Female, 2= Male, 0= not possible to determine sex
Maturation	Maturation scale: Herring 1-8, Blue whiting 1-7
StomFullness	Stomach fullness, visual scale 1-5 (ICES)
StomachWt	Weight of stomach with content (g)
FishNo	Serial number identifying the fish (alphanumeric)

SALMON_sampling (added table linked to the "basic" Biology table in Salsea).

Feltnavn	Datatype	Beskrivelse
Country	Tekst	Post code (or ICES list)
Vessel	Tekst	Call sign (or ICES list)
Cruise	Tekst	Unique cruse identifier
Year	Tal	Four digits
StationNo	Tekst	National station number, link to Station in BIOLOGY
StType	Tekst	CTD, WP2, MIK, TRAWL, PDYP, PBLÅ, PTRAWL, ..., link to StType in StationTypes
Species	Tekst	3 character code, libk to SpeciesID in SPECIES
FishNo	Tekst	Fish/record number, link to Recnr in BIOLOGY
Lice no	Tal	Number of lice
Picture	Tekst	Photograph taken
Scale sample	Tekst	Scale sample taken
Scale loss	Tal	Scale loss in % of body surface (total loss)
Pectoral fin DNA	Tekst	DNA sample from pectoral fin taken
Gonad	Tekst	Gonad sample taken
Gill filament	Tekst	DNA sample from gill taken
Pyloric caeca	Tekst	Pyloric caeca sample taken
Spleen	Tekst	Spleen sample taken
Kidney 3x	Tekst	Kidney smple (3 places) taken
Debes	Tekst	Special sample for disease (ISA) taken (Debes, Faroes)
Liver	Tekst	Liver sample taken
Muscle	Tekst	DNA muscle sample taken
Adipose fin	Tekst	Adipose fin taken
Heart	Tekst	Heart taken
Caudal fin	Tekst	DNA sample of caudal fin taken
Fat lipid	Tekst	Lipid sample taken (fresh fish only)
Stomach	Tekst	Stomach taken
Otolith	Tekst	Otolith taken, serial number in a station (0 means not found)
Carcass	Tekst	Carcass frozen
Cataract	Tekst	Cataract determined: blank= ok, Left, Right or Left & Right
Cataract-L	Tekst	Cataract on left eye: 0= ok, 1 small white spot in lens, 2 30% lens covered white, 3 60% lens covered, 4 whole lens white
Cataract-R	Tekst	Cataract on right eye: 0= ok, 1 small white spot in lens, 2 30% lens covered white, 3 60% lens covered, 4 whole lens white
Tags finclipped	Tekst	Finnclipped or tagged fish
Farmed Escapee Vaccinated	Tekst	Visual determination whether farmed/vaccinated/escaped fish
Notes	Tekst	Comments

FishNo: (alphanumeric), link to FishNo in BIOLOGY table.

Support tables:

Species:

SpeciesID	3 character code (FAO): HER, WHB, MAC, SAL, …
SpeciesName	Species name in English (or sometimes in Faroese)
NODC	10 digit code
ScientificName	Latin name
TSN	TSN code (new universal code)

Values in Species table: (See the SPECIES table in the database for a full list of species and corresponding Latin names, etc)

SpeciesID	SpeciesName
WHB	Blue whiting
CAP	Capelin
COD	Cod
HAD	Haddock
HER	Herring
HOR	Horse mackerel
LUM	Lumpsucker
MAC	Mackerel
MES	Mesopelagic fish
RED	Redfish
POK	Saithe
SAL	Salmon
…	…

Gear:

STtype	Geartype/activity: National definition of station type
GearType	PLANKTON,CTD, or TRAWL (mandatory)
Geardescription	Informative desription of gear

StationTypes

StType	GearType	GearDescription
CTD	CTD	CTD
HYDR-300-HCSBC	CTD	CTD, Rosette, Fluorometer, Light meter
KRIL	PLANKTON	Krill trawl
MIK	PLANKTON	MIK net
MOC	PLANKTON	MOCNESS net
PBLÅ	TRAWL	Pelagic trawl with buoys (blåse)
PDYP	TRAWL	Pelagic trawl without buoys
PTRAWL	TRAWL	Pelagic trawl
TRAWL	TRAWL	Pelagic trawl
TRWL-119-FLF01	TRAWL	Salmon trawl, Aquarium, No groundrope, Vágs doors, 60 m bridles
WP2	PLANKTON	WP2 net

Vessels:

VesselID	Callsign
Vesselname	Vesselname

Values in Vesseltable:

VesselID	Vesselname
SEPI	Argos
TFJA	Arni Fridriksson (old)
TFNA	Arni Fridriksson
TFEA	Bjarni Sæmundsson
LLZG	G.O. Sars (old)
LDGJ	Johan Hjort
OW2252	Magnus Heinason
LHUW	Michael Sars
DBFR	Walter Herwig III
PBVO	Tridens
LMEL	G.O.Sars (new)
OXBH	Dana
EIGB	Celtic Explorer
EIQN	Celtic Voiager
LIVA	Eros
LMQI	Libas
...	...

IGOSS:

QF	Quality Flag
Interpretation	Interpretation

Values in IGOSS table:

QF	Interpretation
0	No control
1	Correct
2	Inconsistent
3	Doubtful
4	Erroneous
5	Corrected