AquaMaps - behind the scene

Josephine “Skit” D. Rius
FishBase Project/INCOFISH WP1
WorlFish Center

INCOFISH WP3 Technical Workshop
Campinas, Brazil
18-22 April 2006

- Description of tables (HCAF, HSPEN, HSPEC)
- How the tables relate
- Rules on calculating the envelope values of environmental parameters
- How the probabilities are calculated

HCAF (Half-degree Cells Authority File)

Contains information on:

a) various cell names (codes) that are used by this and other databases
b) statistical cell properties such (center, limits, and area);
c) membership in relevant areas(FAO areas, EEZs or LMEs);
d) physical attributes (depth, salinity or temperature);
e) biological properties (e.g. pimary production).

Data Providers:

Sea Around Us Project
CSIRO
Kansas Geological Survey
Kristin Kaschner

HSPEN (Species Environmental Envelope)

- Contains information used for describing the environmental tolerance and preference of a species:
  - distribution using FAO areas and bounding box
  - range of values per environmental parameter (min., preferred min., preferred max., max.)

HSPEC (Half-degree Species Assignment)

- Contains the assginment of a species to a half-degree cell and the corresponding probability of occurrence of the species in a given cell.
- Overall probability is the multiplicative equation of each of the environmental parameters (SST, salinity, prim. prod., sea ice concentration, distance to land)

How the tables relate...

Get candidate species and corresponding occurrence data

- Using the Occurrence table of FishBase, get all ‘good’ point records of marine species

where:

‘Good point’ records is a set of at least 10 non-doubtful point records, and should occur in at least 10 ‘good cells’

‘Good cell” occurs in the FAO area OR bounding box in which species occurs

Compute the envelope for each parameter and store in HSPEN

Step 1	Using the good cells, a. take the observed minimum value as the true min b. take the 10th percentile value as the preferred minimum c. take the 90th percentile value as the preferred maximum d. take the observed maximum value as the true maximum
Step 2	Get the extended min. value by: 25th percentile - 1.5 * interquartile range if observed min. > extended min., use extended min. value as the current min. Get the extended max. value by: 75th percentile + 1.5 * interquartile range if observed max < extended max., use extended max. value as the current max.

Step 3	Set a minimum for the preferred range by: If preferred max. – preferred min. < set minimum range then current pref. min. = MidPrefRange – .5 * set min. range current pref. max. = MidPrefRange + .5 * set min. range	where: MidPrefRange = preferred max. + preferred min. / 2 Set min. range varies depending on parameter: SST = 1 Salinity = 1 Primary Production = 2 Sea ice concentration = N.A. Distance to land = 2

Step 4	Set a minimum difference between the current min. and current preferred min. by:
	
	if current pref. min. – current min. < set min. differrence then
	current min. = current pref. min. - set min. differrence
	
	Set a minimum difference between the current max. and current preferred max.
	
	if current max. – current pref. max. < set min. differrence then
	current max. = current pref. max. + set min. differrence
		Set min. difference varies depending on parameter:
		SST = .5
		Salinity = .5
		Primary Production = 1
		Sea ice concentration = N.A.
		Distance to land = 1

Compute for probability and store in HSPEC

1. For each species, compare its envelope values (HSPEN) for each parameter with that of each cell in the HCAF table

2. Compute individual probabilities for each parameter as:
   - 0: if cell value < envelope minimum
   - 0: if cell value > envelpe maximum
   - 1: if cell value is within envelope preferred values

   Cell value – env. Min / pref. min – cell value:
   - if cell value between envelope min. and preferred min.

   Env. Max. – cell value / cell value – pref. max :
   - if cell value between envelope max. and preferred max.

3. Compute overall probaility:

   Depth x SST x Primary Production x Sea ice concentration x Distance to Land

4. Check and note if cell is a ‘good cell’
5. Store the combination Species ID, CellID and probability
6. Pass data to mapper

Stats...

	Number of species	Remarks
Marine species	12498	Marine with occurrence points
With sufficient no. of good cells to compute for envelopes	4546	
With sufficient number of good cells compute for envelopes AND with complete depth range	2909	Expected number of species for which map data can be generated
With map data	691	so far generated