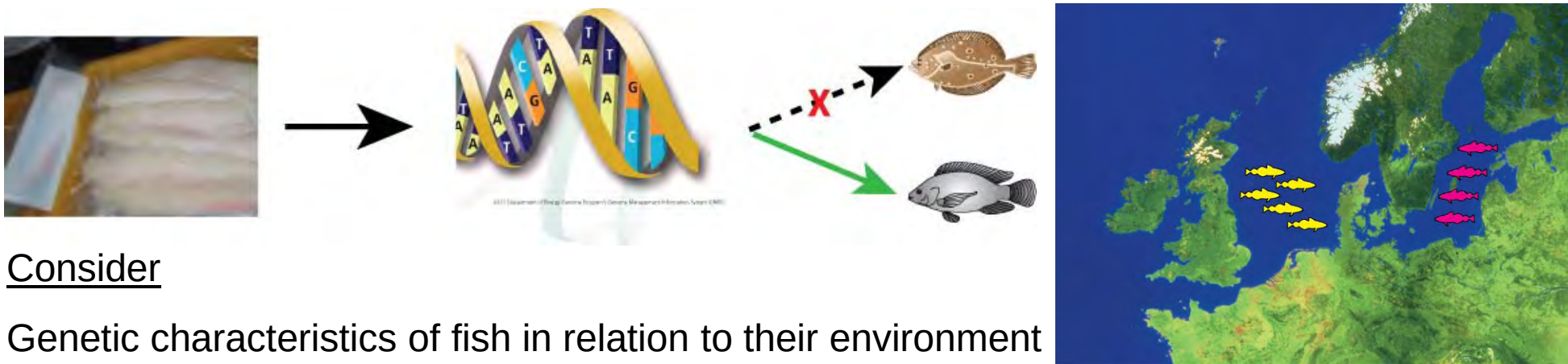




Consider

Genetic characteristics of fish in relation to their environment

Why?

Traceability of fish and fish products, environmental monitoring, species management

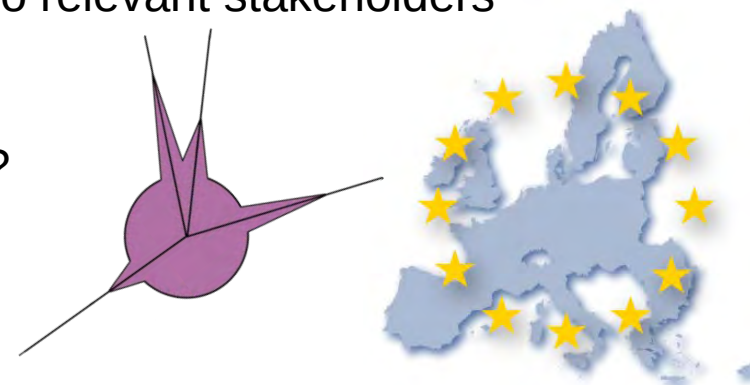
Problem

Illustration and communication of highly scientific data to relevant stakeholders

Solution

Geovisualization ... What characteristics are important?

What symbology is suitable?



Geovisualization Challenges of Seascape Genetics

Eoin Mac Aoidh & Jann Martinsohn

European Commission Joint Research Centre
Maritime Affairs Unit,
Ispra,
Italy

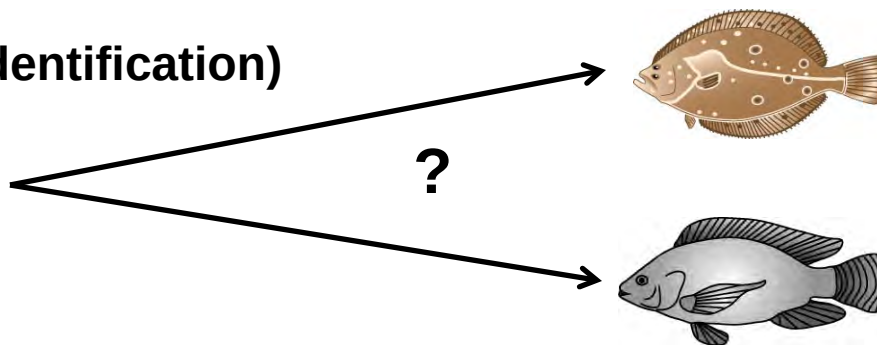


- Context: Seascape genetics
- Case study description & geoviz objectives
- Existing seascape genetics geoviz solutions
- Proposed symbology
- Conclusion



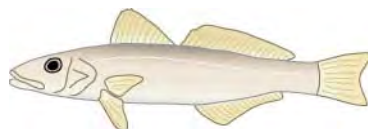
What?

(Species identification)

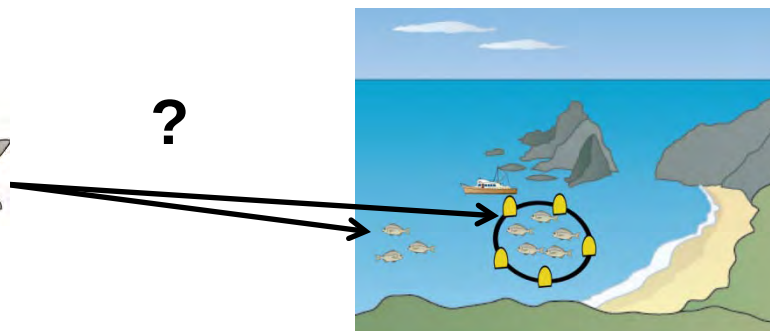


Where From?

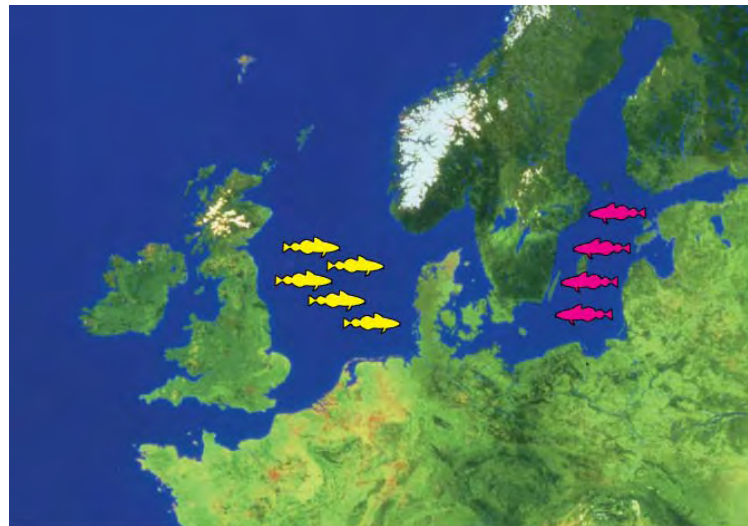
(Origin assignment)



Wild or Cultured?



- Genetic characteristics-environment relationship
- Genetic differences in fish from region A to region B
- Visualize the genetic characteristics in relation to:
Land mass, currents, temperature, salinity, chlorophyll, etc.



- **Consumer protection**
 - Food traceability & labelling schemes
- **Legal & economic**
 - Control & Enforcement to combat Illegal fishing (€10-20bn p.a.)
- **Wildlife preservation & sustainable fishing**
 - Endangered species management
- **Scientific research**
 - Spatio-temporal environmental analysis

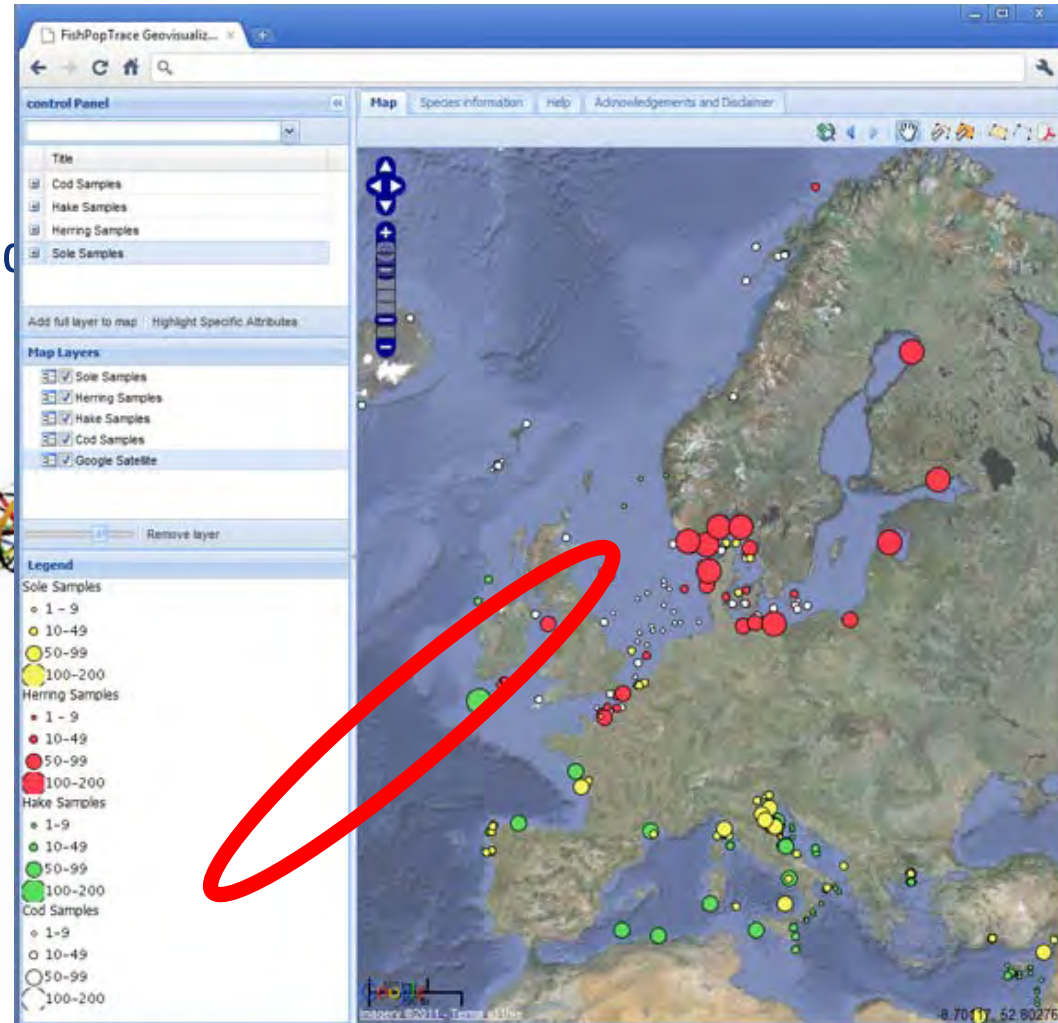
Cross-domain applications

Requires effective transfer of genetic structure information from the scientific domain

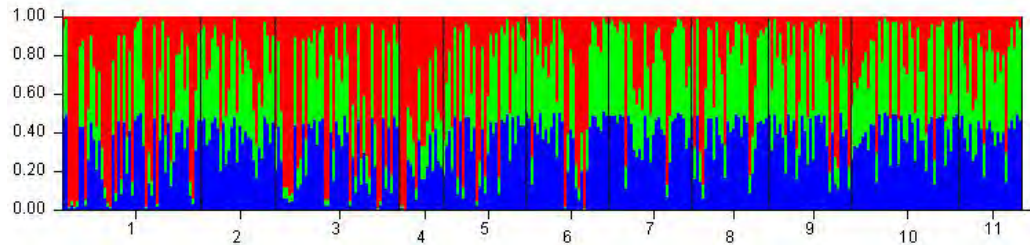
Geovisualization



- Over 7,500 fish specimen collected
 - Cod, hake, herring, and sole.
 - Locations all around Europe
- Analysed using SNPs
 - Sites in the genome where individuals have different variants (A, C, G or T)



- Specimen are clustered according to frequencies of specific SNPs identified in their genome



- Plotting the *geographic locations* at which specimen in the same cluster of frequencies were sampled can reveal strong geographic correlations

Cod in the Baltic Sea have similar frequencies of specific SNPs to each other, but different frequencies to cod in the North Sea.



- Geovisualize:

1. Cluster membership of sampled specimen

To which genetic cluster does a specimen sampled at location X belong?
What other sampled locations show fish with similar characteristics?

2. Probability of misassignment to a cluster

What is the probability that the specimen could be incorrectly assigned to a cluster?

Which cluster is the misassignment likely to have come from?



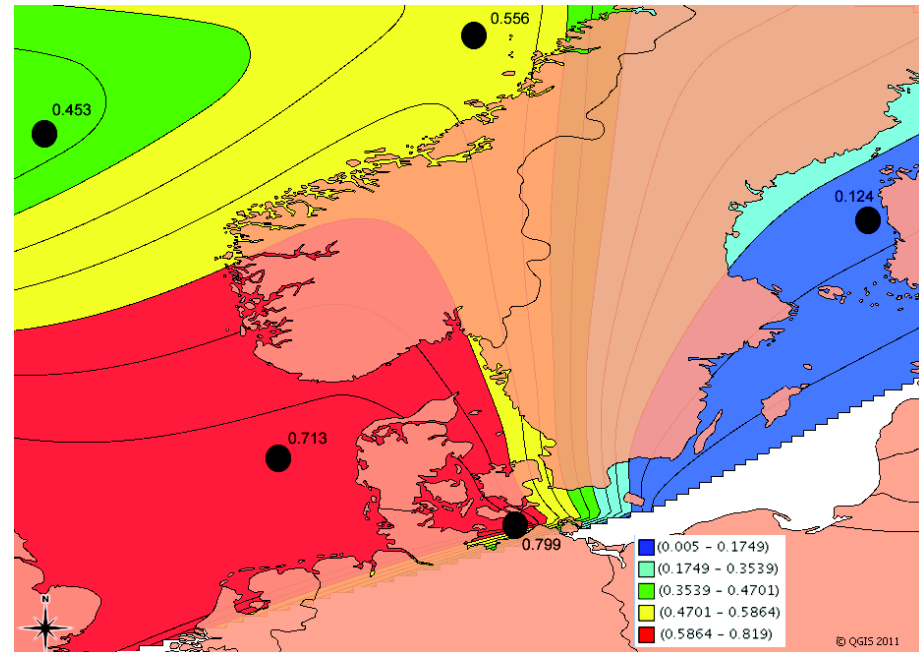
- Web-based geovisualization for communication and analysis (OpenGeo Stack)
- Symbology should be suitable, simple, yet expressive
 - Suitable for available data
 - Simple to understand
 - Simple to render

Why not use existing symbologies?



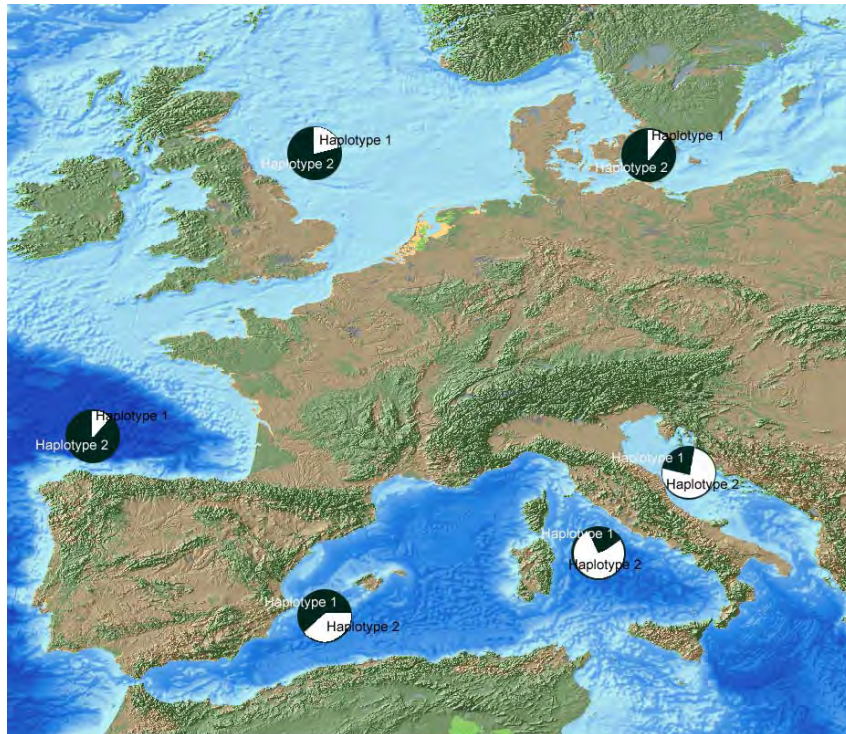
- Continuous surface maps are unsuitable
 - Data is not continuous
 - Too many unknown areas we cannot accurately interpolate
 - Maps could be misleading causing problems
 - E.g. for control and enforcement use

- Complex clipping, interpolating and rendering around land boundaries



- Point based representation is more appropriate

- Frequencies of genetic characteristics often visualized as geolocated pie charts (point based)



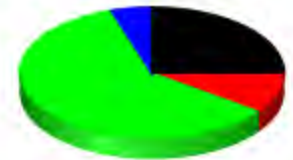
Haplotypes present at sampled locations

Same concept applies for visualizing ratios of cluster membership

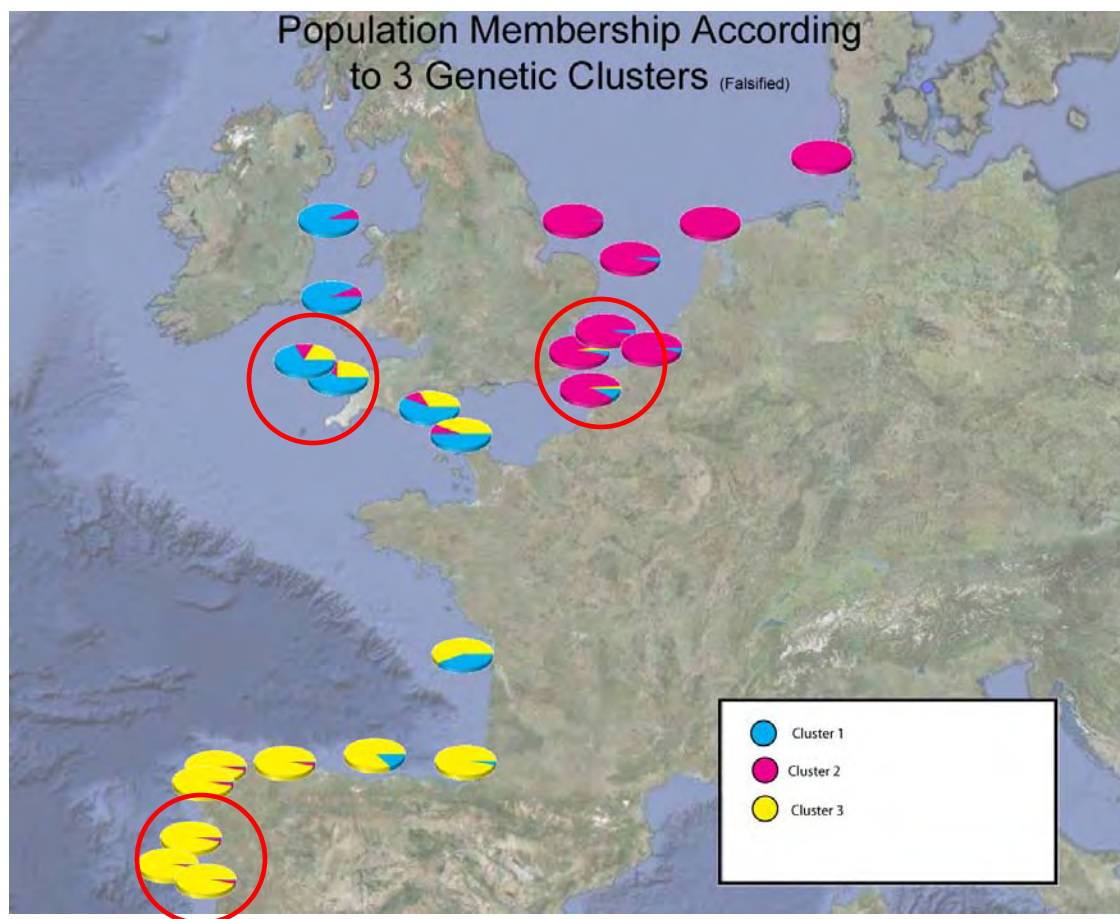
Assume 4 genetically different clusters of the same species

Specimen at location (x,y) have genetic material originating from the

Green (60%), Black (25%), Red (10%), and Blue (5%) clusters

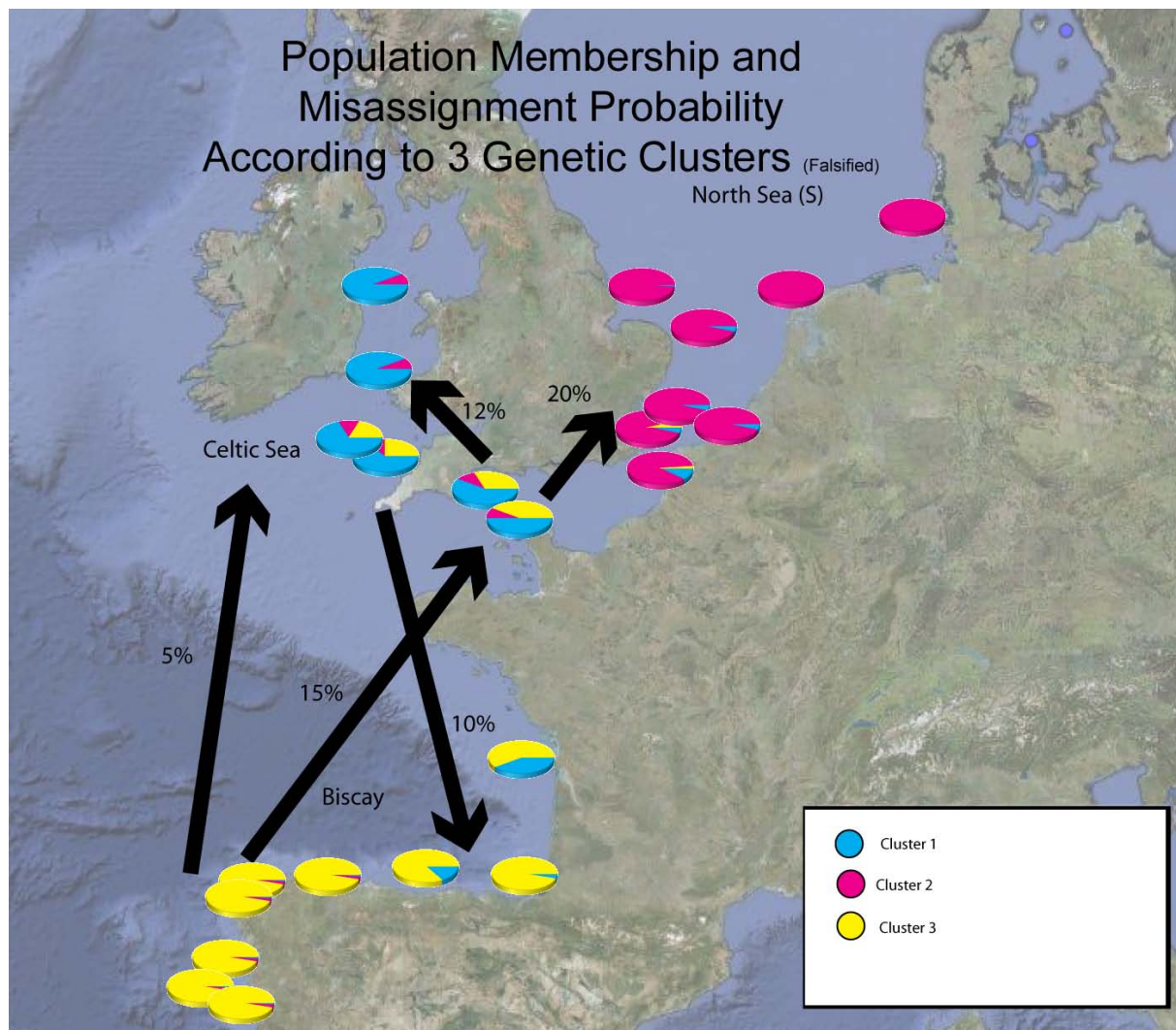


- Population cluster membership of common sole plotted using pie charts (FishPopTrace analysis slightly falsified)



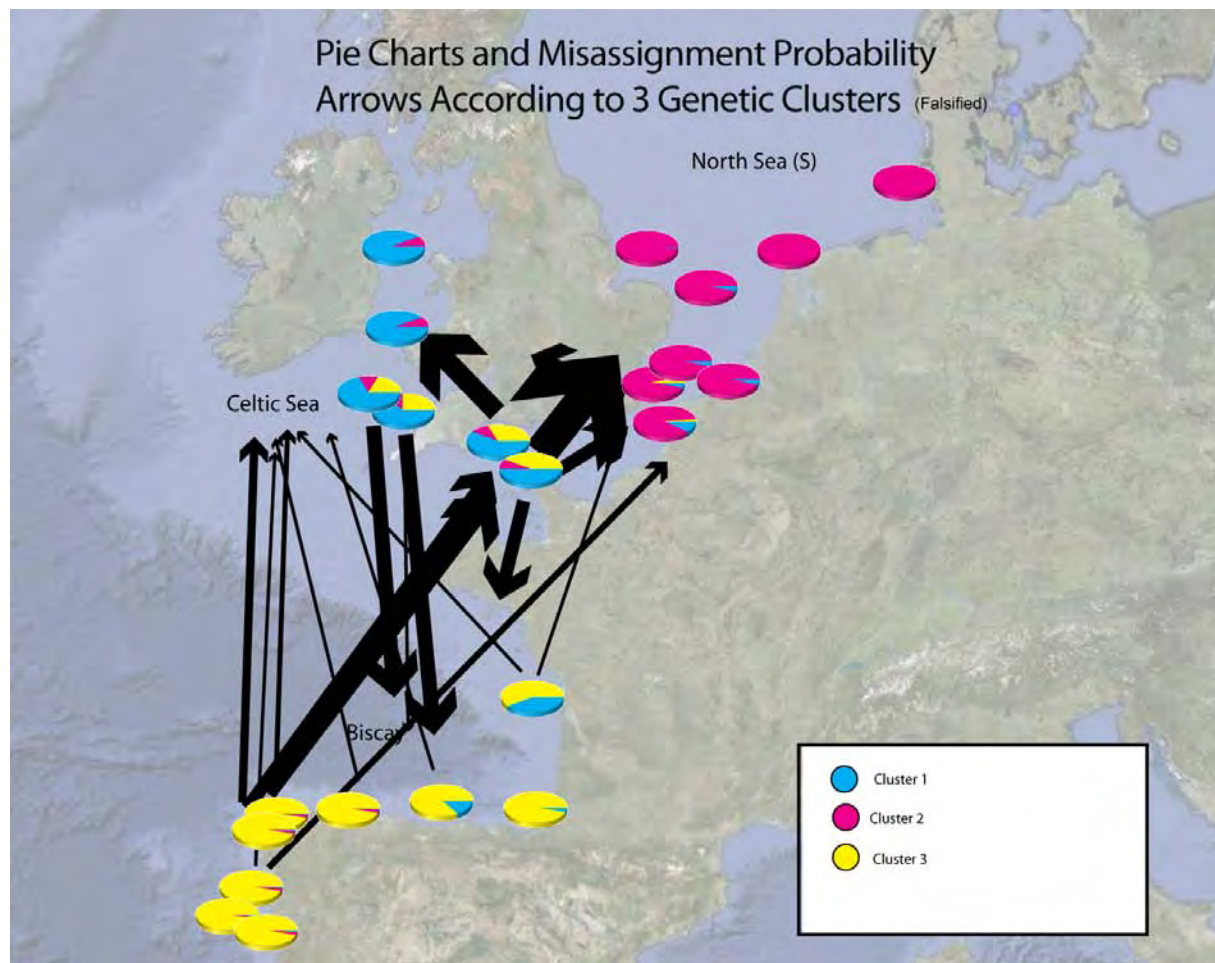
- Small ratios sometimes occluded by other pies
- Pie charts must be regularly shaped
- Variables Restricted to:
slice size, colour, pie radius





- Generalisation of misassignment trends by grouping similar specimen
- What about automated maps for a Web-based interface:
Clusters cannot be grouped, arrowed and labelled by hand





Remember:

Seascape genetics must also show elements of the seascape

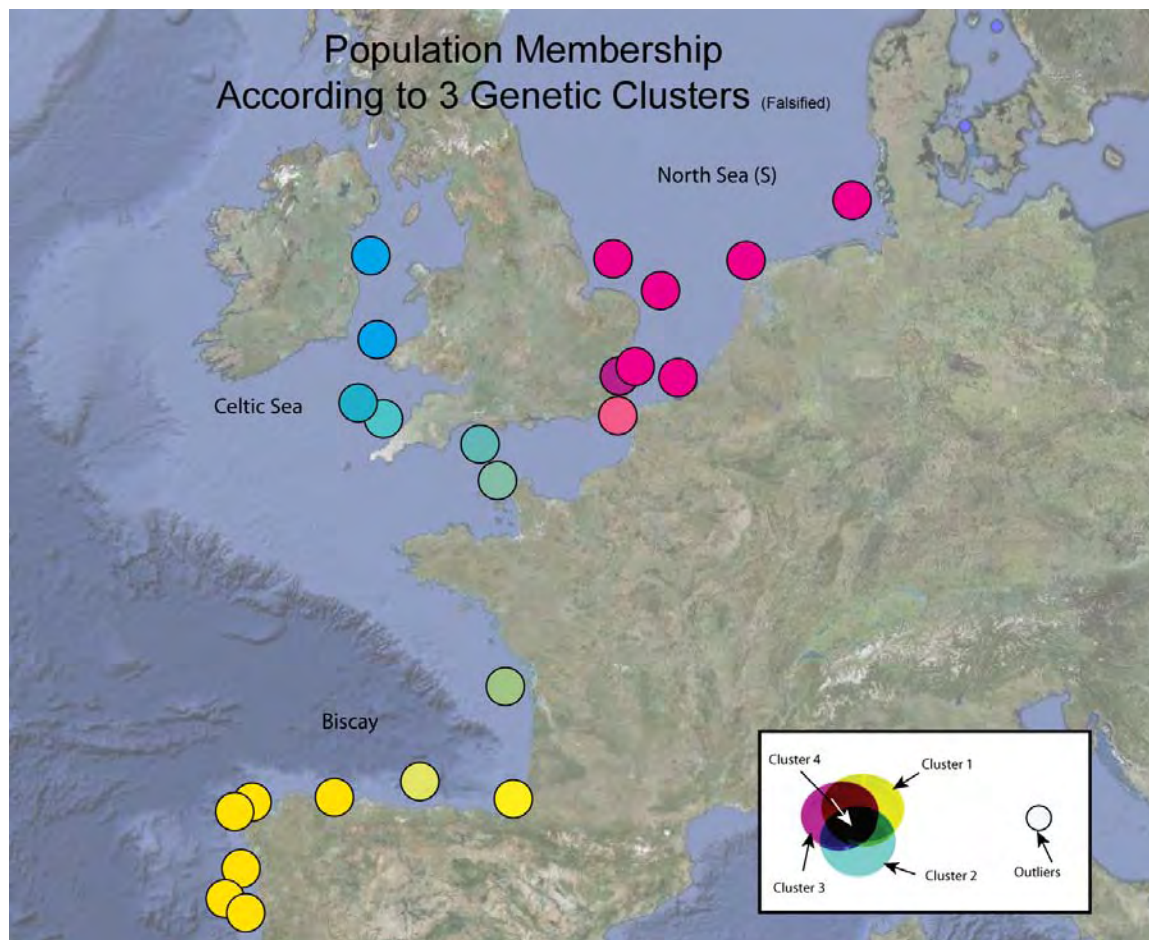
Problems:

- Small segment occlusions
- Grouping / simplifying arrows
- Map occlusion
- Arrows and pies are separate objects

Solution:

Simplify and visualize as a single layer

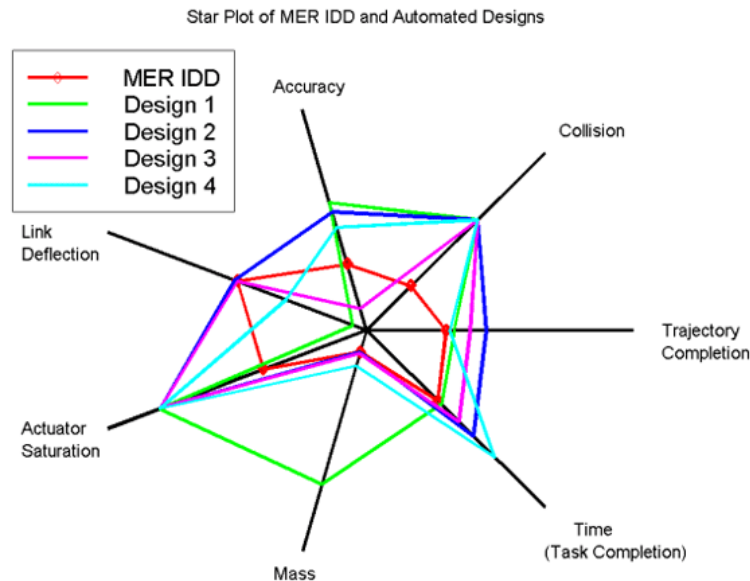




- Ratio of cluster characteristics generates a fill colour
- Each cluster assigned to a colour of the CMYK model. Ratio = Hex
- Notion of cluster membership retained
- Occlusion no longer a problem
- Container shape no longer restricted
- **How about misassignment probabilities?**

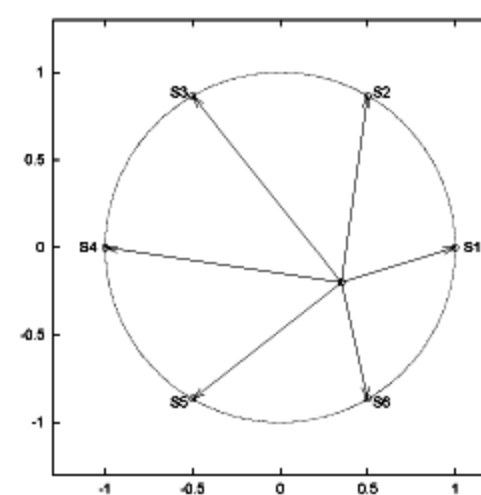


- Popular multivariate representation techniques



Star Plot

Image from www.nasa.gov



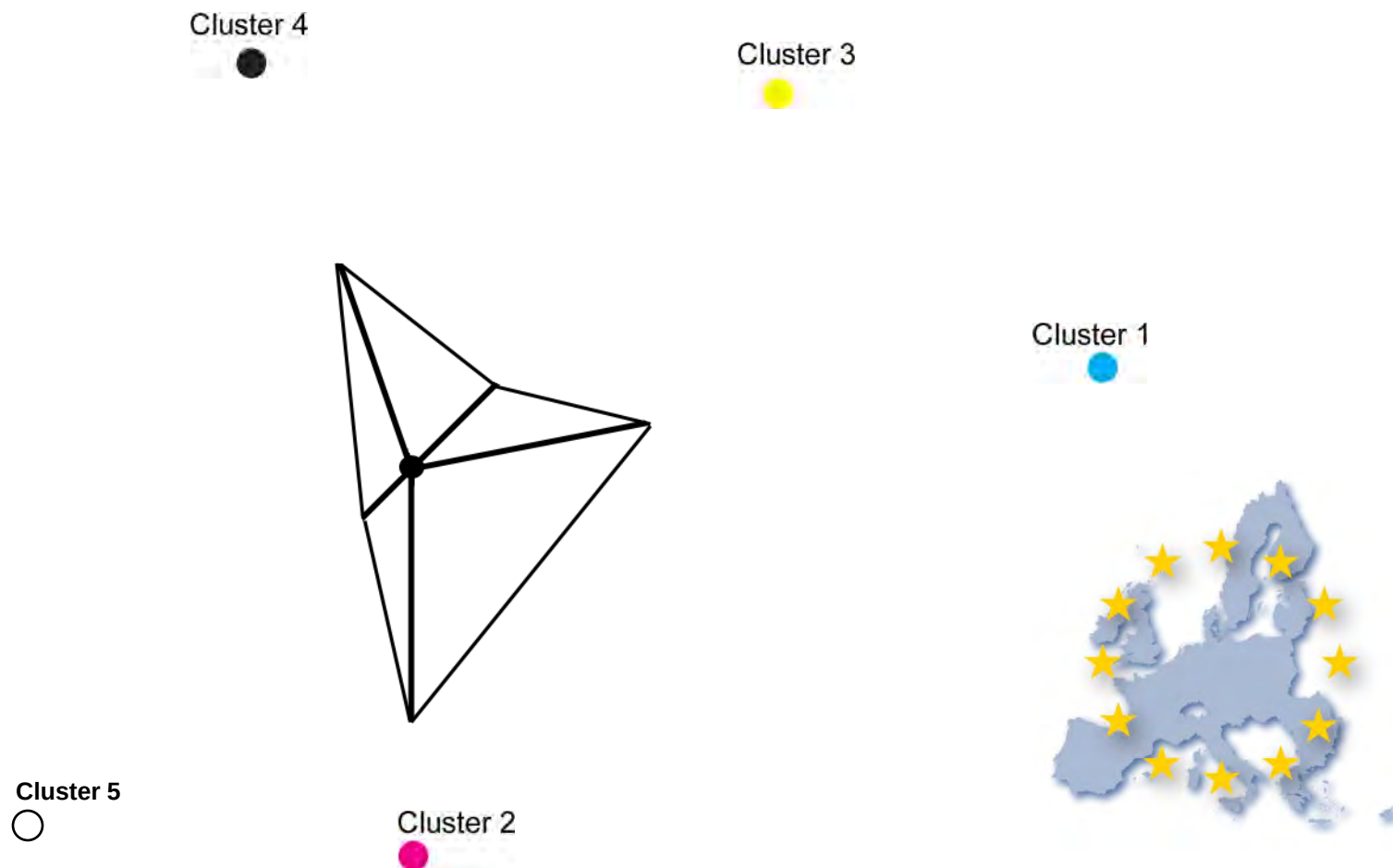
RadVIZ method (Ankerst et. al.)

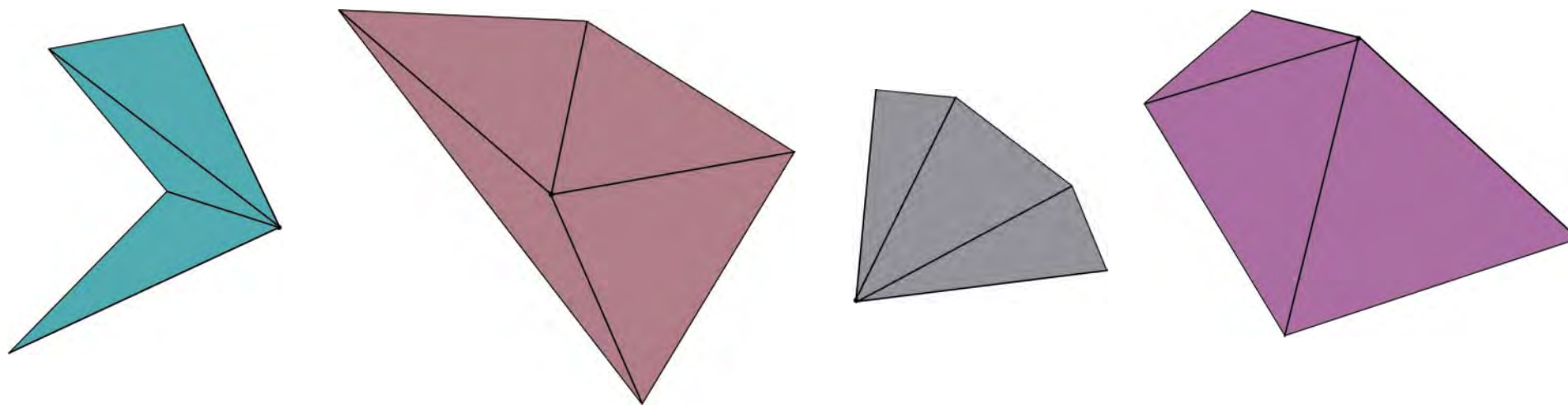
Image from Brundsdon et. al.

- N equally spaced axes



- Rotate the axes to point in the direction of cluster misassignment, plot probability on each axis





- Difficult to compare (non-uniform shape)
- Size of fill area is misleading
- Direction and probability difficult to discern



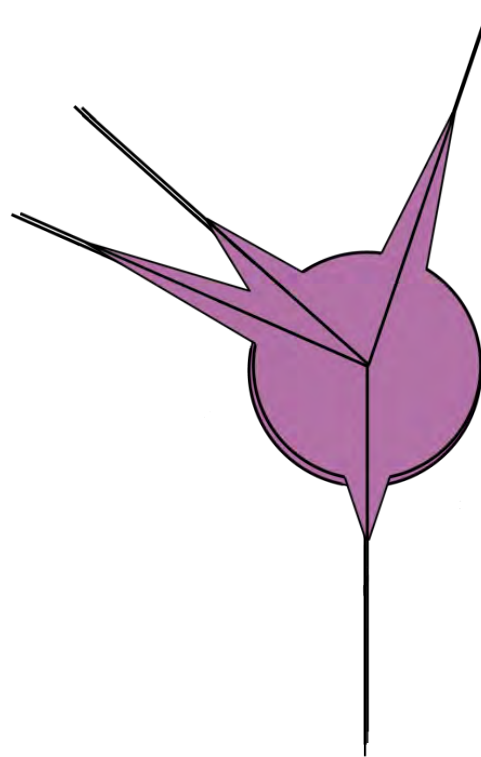
Cluster 3



Cluster 1

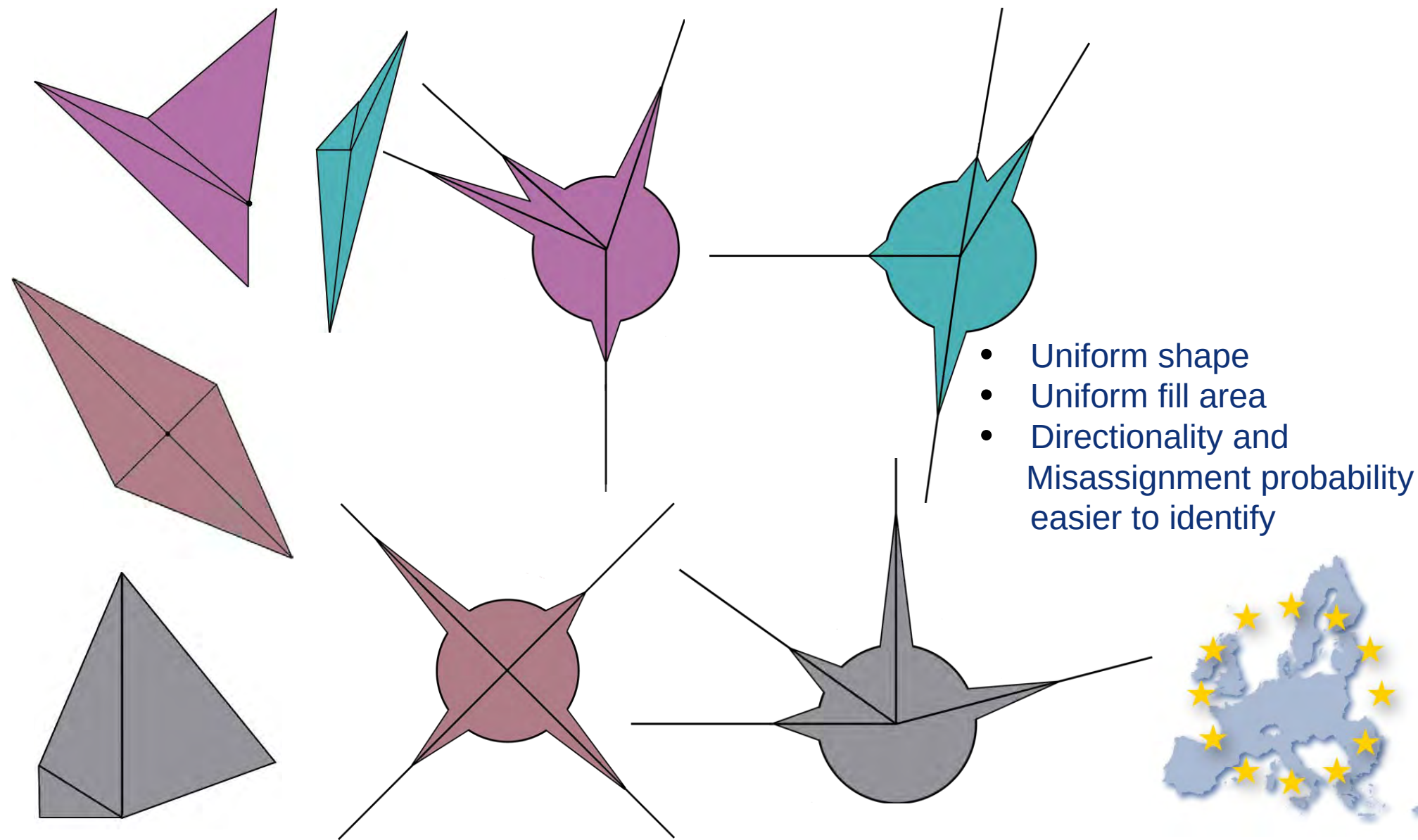


Cluster 4

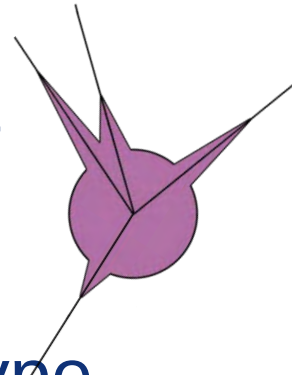


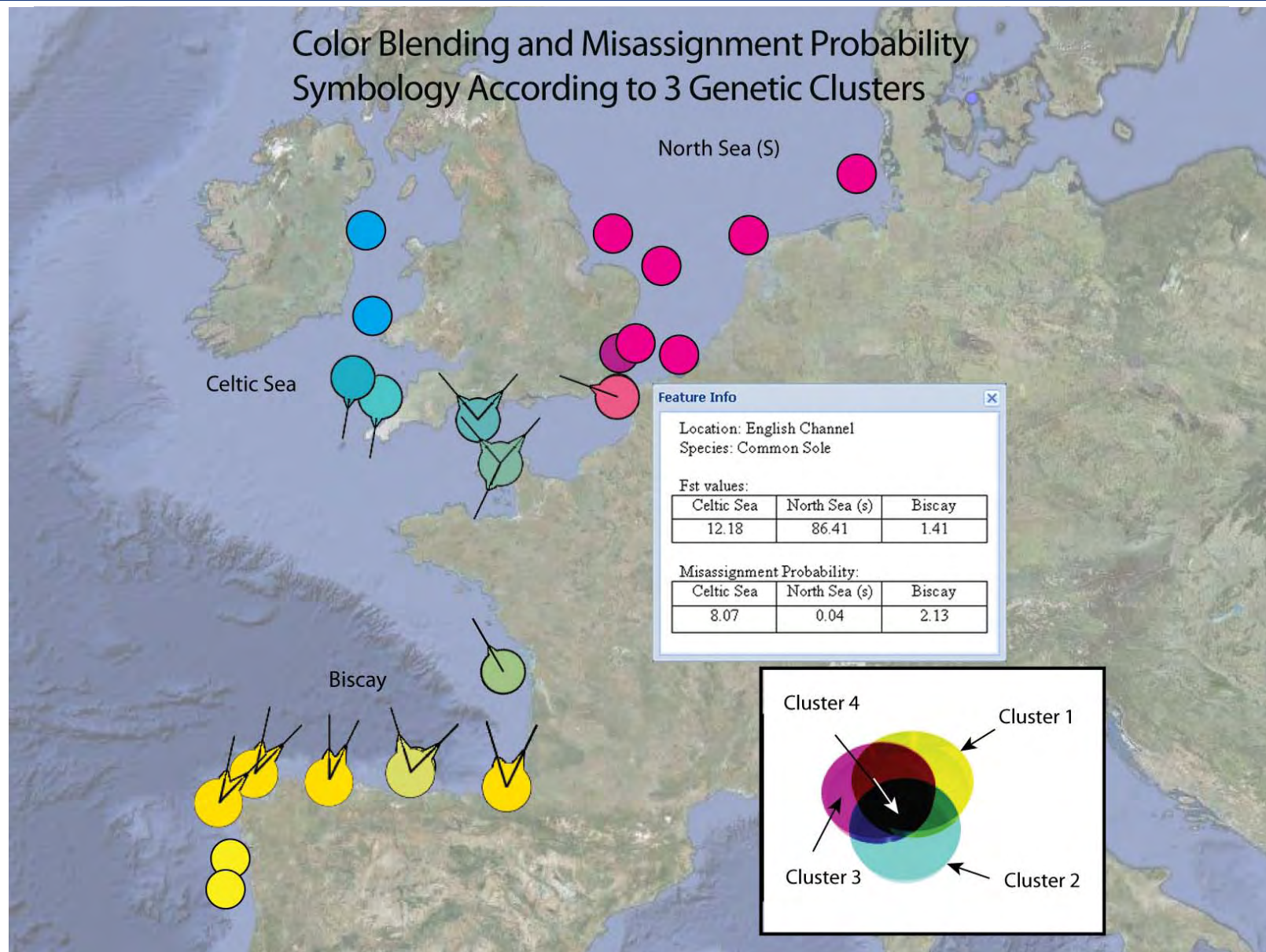
Cluster 2



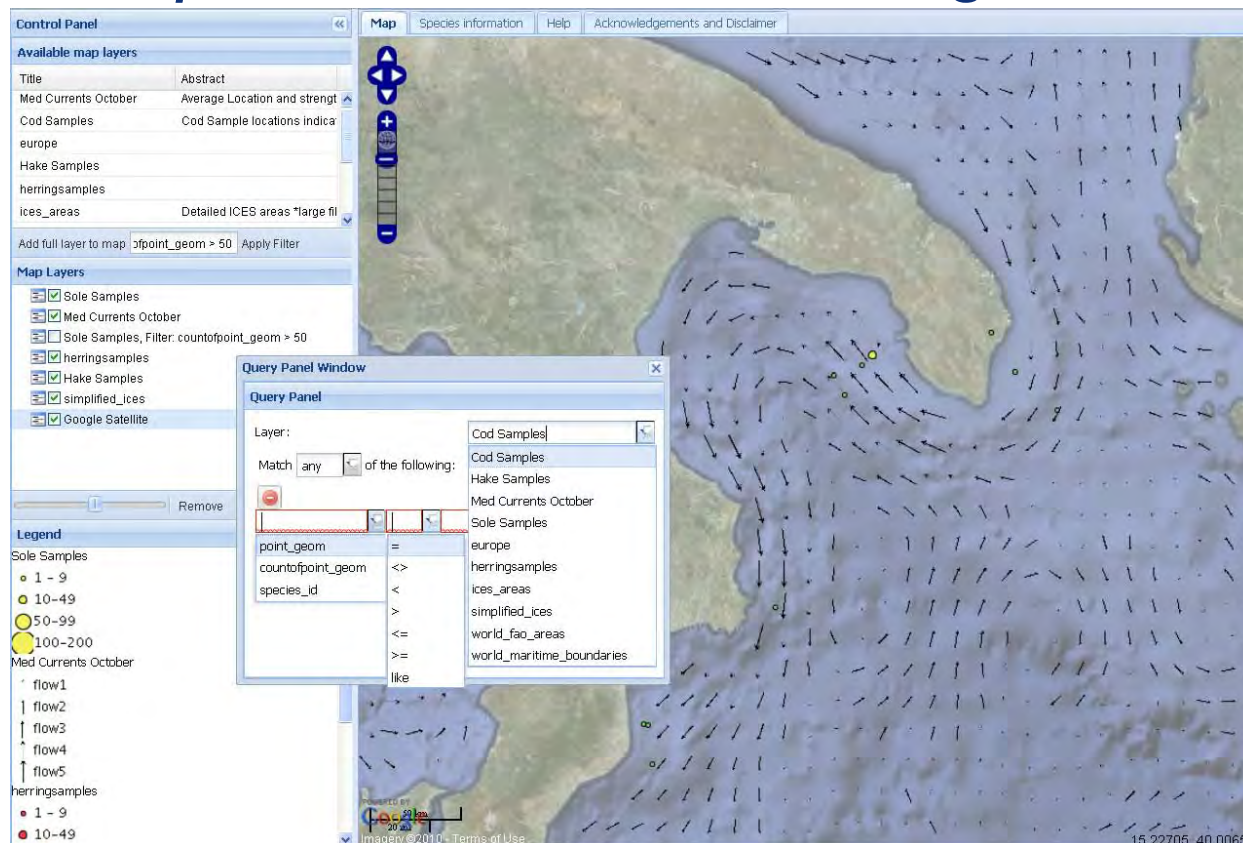


- Replacement for pie charts and misassignment arrows
- Information incorporated into a single symbol type
- A single point-based map layer
- Grouping and drawing of arrows unnecessary
- Automated symbol generation
- Overhead on system and user reduced
- Background space available for seascape variables
- How will it look on a map?





- *Implement* auto-symbol generation technique
- *Incorporate* it into the existing Web-based platform



- *Evaluate*

