Identification of Waves in IGC files

Prof. Dr. Alfred Ultsch Databionics Research Group University of Marburg





Mountain Wave Project

Databionics

 Databionics means copying algorithms from nature
 e.g. swarm algorithms, neural networks, social networks

 Swarms e.g. bee hives, ants etc.
 are extremely good at optimizing food sources from every changing and unknown environments

Databionics

Hermann Trimmel: why look at IGC- Data cemeteries:

• it's just dust in there!

He is right!

However:

Swarms are extremely good at optimizing food sources from unknown, ever changing and even harsh environments

Uni Marburg ultsch@informatik.uni-marburg.de

Swarm (social) calculations

- Swarm = all pilots producing ICG-files
- "Food" = lift (climbs)
- There should be patterns of optimal "food source" discoveries
- Eg: north Bavaria/Thüringen during good thermal conditions (prob. of thermal climbs during 5 days)



Alfred Ultsch Uni Marburg ultsch@informatik.

Identification of Waves in IGC files

- Mountain Wave Project (MWP)
- MWP flights in the Andes of South America
- See Map: Locations of 2.500 climbs found in ~ 100 IGC-files of flights
- Which of these Climbs are Thermals, which are in Waves?



Construction of a Climb-Classifier for the Andes flights

Problems to solve:

A) Identification of Climbs in IGC files



= Climb (time series analysis: details reported elsewhere)

B) Can Waves be distinguished from Thermals ?

Wave/Thermal Classifier

- Construction of an unsupervised Classifier
 - (=no ground truth available)
- based on decision rules
- Check rules for plausibility

⇒Unsupervised rule based classifier for wave vs. thermal climbs in the Andes

Method to construct unsupervised rule based classifier

- Model (normalized) variables as a Mixture of Gaussians (GMM)
- Optimize Model using Expectation maximization (EM)
- Verify results using suitable density estimation (PDE -> Ultsch 2003)
- Use Bayes Decision on GMM
- Combine decision Rules with and/or
- Check for plausibility

Rule based Classifier for Climbs in Andes

Example of a particular rule:



R1: if distance flown in a climb >2.8 km then climb is a wave

Result:



Computer Program for identification of wave climbs in Andes (colors = wave strength classes) [Heise/Ultsch 2008]

Question: How good is this classification?

Remember: no ground truth available!

Is the Classifier correct?

Approach:

- Construct classifier for flights in the Alps with same approach
- Get the ground truth for the Alps
- Measure the performance against ground truth and
- Compare with supervised methods to construct classifiers

Estimate correctness for Andes

Get flights containing wave and thermal climbs

- Philip Ohrndorf
- Student of Geography &
- Databionics & Pilot



- selected 160 Flights out of all the flights of 2007/2008/2009
- i.e <200 out of 200.000 flights
- From Online Contest (OLC)database
- Obtaining this data is cumbersome!

 OLC allows only small downloads(10) per day

Uni Marburg Udy ultsch@informatik.uni-marburg.ae

Alfred Ultsch

Ground Truth

 in more than 1.5 months of full time work Philip Ohrndorf selected and hand classified 160 IGC files





Measuring Correctness

- 1. Data was randomly split into two sets
- 2. 80% = training set
 - 20% = test set
 - On the training set the Rule Based Classifier was built without reference to expert classification (ground truth)
 - Correctness (total Accuracy) was measured on test data set
 - Steps1-5 repeated 100 times, mean and s.d. calculated:

Result:Correctness 76% +-2%

Comparison to other Classifiers

- Other Classifier constructions use the ground truth (supervised classification)
- Naïve Bayes classifier = Golden standard
- CART = Example of supervised Rule Based Classifier

Alfred Ultsch **CONSTRUCTION.** Uni Marburg ultsch@informatik.uni-marburg.de



CART's decision Variables

Up to linear correlation: the same variables were automatically selected as in Unsupervised Andes classifier:



Can we do better in the Alps ?

Use an Artificial Neural Network
Logistic sigmoid, back propagation
MLP with 3 hidden layers
Input = 8 - 12-32-64 - 2 =Output



Can we do better in the Alps ?



Performance of this ANN: 98+-1% Almost perfect!



Uni Marburg ultsch@informatik.uni-marburg.de



Accuracy

100

Conclusions

- The construction of an unsupervised Rulebased classification of IGC-flight data from the Alps results in an 76% correct classifier.
- This is comparable to the Golden Standard (Naïve Bayes)
- Supervised Rule Construction leads to the same variables & same performance
- Problem might be easier for the Andes: better bimodal distributions (= other use of waves)
- => It is not unreasonable to assume that our Wave Classifier for the Andes functions in the 80% correctness range

Outlook

For the alps an almost perfect classifier could be constructed (ANN)



- This requires, however, ground truth (Experts who classify the data)
- If this is not available unsupervised classification can only be improved when data from other sources are used:
- Geographic data : height above terrain
- Meteorological data: wind speed & direction



They are a new type of computations:
Social networks = swarms

Provide a new look "from above" => the whole is more than the sum of its parts



OLC's- data access policy

However • 2 serious bottlenecks

OLC's- data access policy

- No substantial scientific theories on swarm computing available
- sorry we all (science) are rather ignorant of this type of computing!
- Needed theory and tools (e.g. U-Matrix on Self Organizing Neural Networks) to discover

Emerging Structures

Did we find something interesting in the past in the data cemeteries?

- Yes!
 - **E**.g:
 - In > 20.000 climbs in "the same" weather:
 - climb strength in thermals is
 - NOT Gaussian distributed, but
 - Square-root-Gausssian distributed (extremely significant p-value)[Ultsch03]

Thank You for Your attention!

Questions?

Vertiklalgeschwindigkeiten der Climbs in Leewellen in m/s

- 0,14 1,48 1.48 - 2.27
- 2.27 3.25 3.25 - 4.73

4.73 - 7.04

Α Unimary

ultsch@informatik.uni-marburg.de ultsch@informatik.uni-marburg.de

Contact: Ultsch@ulweb.de