ON CORRELATING BIRD MIGRATION TRAJECTORY WITH CLIMATE CHANGES

Janaina Oleinik, Jose Antonio Fernandes de Macedo

ABSTRACT

Climate changes are expected to affect bird migration in several aspects including timing changes, breeding and migration orientation. The correlation analysis of several climate conditions (e.g. temperature, wind, humidity, etc) and bird migration trajectory is the key for explaining bird behavior during migration. Moreover, the resulting correlation can be used for predicting new bird behavior according to climate changes. In this paper we propose an integrated solution for correlating bird migration trajectory with climate conditions. This solution is composed by two orthogonal and complementary methods. The first method concerns discovering regions where birds are used to stop during their migration. The second method is based on a machine learning algorithm for classifying bird stops according to climate conditions. A real bird migration scenario was used for assessing the accuracy of the integrated solution.

Keywords: Spatio-temporal Analysis, Trajectory, Quadtree, Climate, machine learning, bird's migration

1 INTRODUCTION

Migration is a natural process, whereby different birds fly over distances of hundreds to thousands of kilometers in order to find the best ecological conditions and habitats for feeding, breeding and raising their young (Caswell *et al* 2006). Migratory birds therefore rarely fly to their destination non-stop but interrupt their trajectory frequently to rest and feed, or to sit out a spell of bad weather. Climate changes affect the patterns of bird migration. For example, according to (Beaumont *et al* 2006) Australia's migratory birds are arriving earlier and leaving later - most likely due to global warming.

In recent years, many international organizations (Vagg *et al* 2006) began to trace the bird's migration using modern methods such as positioning devices and satellites. These methods have permitted to capture information about bird displacement easier and faster than ever. Thus, bird trajectory information is the key for explaining bird behavior. For instance, typical questions that can be answered using these data might be:

1. Which are the regions that the birds are used to stop?
2. How does the weather condition affect the behavior of the birds' migration on one stop?

Indeed these two questions are orthogonal in the sense that the first question aims at finding regions of interest whereas the second question concerns how to correlate those regions with climate conditions. On the one hand, finding regions where birds are used to stop is a problem that can be solved by just regarding the geometrical facet of bird's trajectory, which relies on space and time dimensions. For this problem, a geometrical computational strategy for finding dense regions can be applied (Samet 1989, 1995). On the other hand, correlating birds' stop with weather conditions is a complex task (Zalakevicius 2001, Bouten *et al* 2005). From the data analysis point of view, user's knowledge must be taken into account in order to avoid obvious and incorrect results. The other way around, analysis methods must provide some insights to the user that permits to tune and improve these methods.

Although these questions need two different solutions, the second question is dependent on the results of the first question. Thus, an integrated approach is needed for providing an adequate solution for the problem of analyzing and correlating trajectory raw data with climate conditions.

In this paper we aim at providing a method to accomplish this task. First, we provide an algorithm for analyzing trajectory data and discover dense regions. Second, we exploit a machine learning method to classify and correlate climate conditions with birds' stops. The contributions of this paper are:

1. An algorithm for discovering dense regions given a set of trajectories (represented by a set of time-geography points);
2. An implementation of a supervised machine learning for correlating climate conditions with stop and nonstop class;
3. An integrated solution that permits to correlate bird stops with climate conditions.

The following sections are organized in the following way: in the next section we establish the problem and present some definitions. The integrated solution is explained in the section 3 with separate sections for each concerned problem, dense regions and climate correlation. The experimental settings and results are detailed in the section 4. In the section 5, we discuss related work. Finally, in the section 6, we will conclude the work that we have done and discuss the future work.

2 Problem Statement

2.1 Dense Regions

Trajectories of birds are collected by position devices and represented as a collection of time-geography points (i.e. timestamp, latitude and longitude). Each point of a trajectory represents the moment when a bird stops at specific location on the ground. Since we are interested in correlating bird migration trajectory with climate conditions, it is not appropriate to use each bird stop (trajectory point) separately in this correlation. Thus, the best idea is to find regions where birds are used to stop and correlate those regions with climate conditions. These regions of interest are called dense regions.

A dense region is a geographic area with a high concentration of moving objects (birds' stops) in a time interval. In the practical viewpoint, we can represent a dense region with a geometry defined by aggregation of data captured from bird's trajectory.

Generally speaking, a typical query for finding dense regions must cope with the following issues:

- What is the search space? This question concerns the candidate regions where we can find agglomeration of moving objects. For example, in the Figure 1, this could be all Europe, Africa or just the dark area corresponding to the winter-feeding area.

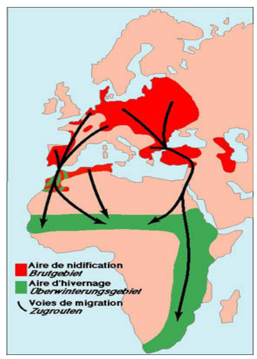

Figure 1 White Stork Migration

- How concentrated a region should be? For choosing a region as dense it must be clear how dense it is expected to be. By density, we mean the ratio between the number of moving objects and the area of the region.

- When do data arrive? The analysis of trajectory data could be upon on historical data or continuously arriving data (e.g. stream data). In the first case, the analysis aims to discover behaviors or patterns, which could be used in further analysis. Analyzing data at each timestamp they arrive aims preview concentration area in a determined ahead timestamp based on speed and direction of data captured. In some scenarios, the first case can be abstracted to the second one.

- Which is the dense region size and shape? For calculating the density of a region it is necessary to know the shape and size the area of the region. In a search space we can look for squares, circles, etc that are dense according to the definition. Or we can relax the form and look for arbitrary polygons. The minimum size of a region is important for avoiding report imprecise areas.

Thus, a dense region can be defined as follows:

***Definition 1 (Dense Region)**:* *In a search space, populated with historical data corresponding to bird migration trajectory, (spatio-temporal data) we search for regions* *respecting a threshold on the number of points by area (density) with a minimal of number of points. Those areas will be reported as dense region.*

2.2 Climate correlation

Climate conditions are spatio-temporal phenomena since meteorological information are always associated with some spatial and time extent. The spatial extent can be defined as a point or polygon. For the sake of simplicity, in our work both dense and climate regions are represented as rectangles. Thus the determination of correlations between dense and climate regions is facilitated.

Weather conditions are comprised of several phenomena:

• Wind: Birds have a phenomenal understanding of winds and appear to adjust their time of migration and their altitudes to maximize tail winds and to maintain a preferred direction despite winds blowing from varying angles. Then, to study the wind condition, such as the direction, air speed, is the key element to understand how the migratory bird choose the right time to depart, how they choose the trajectory, stops and how different birds react to wind condition.

• Temperature, moisture and barometric pressure: Birds are urged northward in the spring by rising temperatures, increasing moisture, and decreasing pressure. Opposite conditions—falling temperatures, decreasing moisture, and increasing barometric pressure—urge them south in the fall.

• Rain: Migrating birds tend to avoid storms and flying against the wind.

• Temperature: Many birds follow a temperature gradient as they return to nest in the spring. Birds vary in sensitivity toward temperature and other environmental conditions.

Since all those phenomena have some influence in the bird migration, we can state the problem of correlate bird migration with climate changes as follows:

Defintion 2 (Climate Correlation): *Given an area where birds are used to stop (dense region) and the weather conditions for this area (wind, temperature, precipitation) in some time instants, express quantitatively the weather's influence in the stop.*

3 INTEGRATED SOLUTION

In this paper, we propose an integrated solution for correlating climate conditions with bird stops during migration. The first part of our solution concerns discovering the regions where birds are used to stop. In order to achieve this result, we have developed an algorithm for finding dense regions on raw trajectory data. The second part of the solution concerns correlating dense regions with climate conditions using a machine learning tool. To this end, we have implemented a machine learning that classifies a tuple of climate conditions into a stop or nonstop class.

Figure 2 illustrates our integrated solution. First, trajectory raw data is processed by the dense region algorithm that stores selected regions in a database. These regions are used as input training/test set by the machine learning-SVM (Kecman 2001). Finally, the trained machine learning can be used to predict future bird stops.

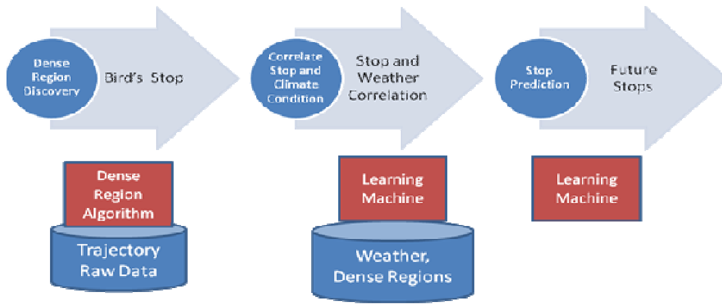

Figure-2 Integrated Solution

3.1 Dense Region Algorithm

The objective of dense region algorithm is: "Given a set of trajectory points, find locations that accumulate a number of moving objects above a given threshold". For example, Figure 3 illustrates a dense region, represented by a rectangle, within a set of bird trajectories.

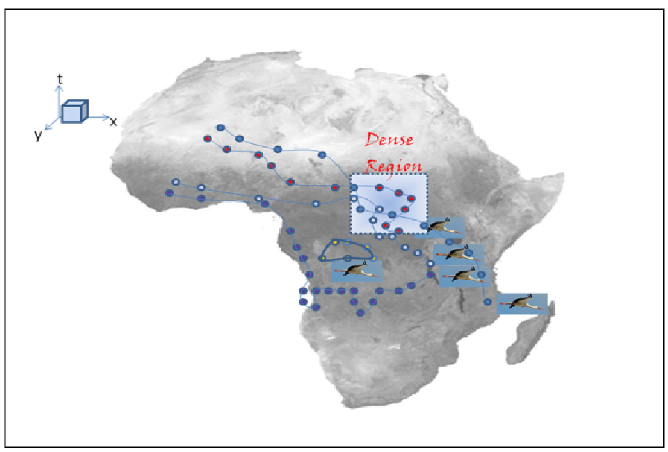

Figure-3-Bird trajectories

Since finding dense regions in set of points is a typical geometrical problem we have chosen PR Quadtree (Samet 1989, 1995) structure as an adequate solution for our problem because it is sensitive to positioning of object. Besides, the shape of the tree is independent of the insertion order of the points. The PR Quadtree is good for search within specified distance of given record and permits to choose the number of points accepted inside a node.

According to this problem, we describe in Listing 1 a pseudo code algorithm for creating a PR Quadtree and in Listing 2, we describe a pseudo code of algorithm for inserting a point in the PR Quadtree. Listining 3 checks if a region can be a dense region.

Generally speaking, the algorithm starts creating a root node for the PR Quadtree, which represents the whole geometric space. Then, the algorithm starts inserting each point of trajectory set in the root node. When the number of points in a root node exceeds a predefined threshold, the algorithm split the root node creating new four leaf nodes representing each one a quarter of the whole space. During the split operation all points presented in the root node must be moved to their respective node tree according their position. Afterwards, the algorithm continues inserting in each leaf until another splits occurs.

Listing 1 describes the main algorithm function named FindingDRinPRQuatree. This function initializes the tree and called a recursive functions for inserting all the points in ListPoints. In InsertPointQT points are added to list of points of a node representing the area where it fits in. If this node is internal, point is added to it and also propagated to the correct child node. The input data for this function are:

- Search_Space is a data structure containing the minimum and maximum value for X, minimum and maximum values for coordinate Y. It means, points that limited the search space.
- ListPoints: it gives all the points representing the moving objects movements observed.
- Thresholdpoints: a value indicating the minimum number of points necessary to consider an area as a dense area.
- Thresholdsize: a side length of a square area. With this value the minimum region area is calculated.

When a region (represented by a node) reaches the *Thresholdpoints* it should be split. In this moment, our algorithm checks for a possibility of a dense region including the threshold number of points in a area centered in the region being splitted. If the points are concentrated there, the node is marked as dense region and split task continues. Otherwise, we check for a minimum size, *Thresholdsize*, if it is reached then the node will be not split but can continue to receive points.

The found dense regions are marked in the structure. To list all of them we must read the tree.

Listing 1. FindingDRinPRQuatree

```
INPUT: Search_Space, ListPoints, Thresholdpoints, Thresholdsize
OUTPUT: Tree with nodes annotated as dense regions.
   1. RT_NODE = new QTNode (Search_Space); Initialize tree
   2. FOR EACH POINT P IN ListPoints:
         a. RT_NODE = InsertPointQT(P, RT_NODE);
   3. RETURN RT_NODE
```

Listing 2. InsertPointQT

```
INPUT : Point P, Node N
OUTPUT : Node N
   1. IF (N.Number_of_points < Thresholdpoints) THEN
   2. Add point to Node's list of points. Increment points counter
   3. Copy points to respective child node
   4. ELSE
   5. IF  (PmaxX - PminX )/2 > Thresholdsize THEN
         a. CheckCandidateDR(N); We can continue to split…
         b. Create the new leaf nodes;
         c. Create the new node replacing the leaf node splitted.
   6. ELSE Don't split anymore.
         a. Add point to Node's list of points. Increment points counter
         b. Copy points to respective child node
   7. SetDRtag(N,"yes"); Mark the node as a dense region
```

Listing 3. CheckCandidateDR

```
INPUT: Node N
   1. CR =Compute the region (Thresholdsize) centered in N area.
   2. Nr_points =Search into N.PointsList for points belonging to CR
   3. IF number_points > Thresholdpoints THEN
         a. SetDRtag(N,"yes");
```

The time complexity a PR Quadtree algorithm is $\Theta(n \log4 n)$ (Pemmaraju *et al* 1994), where n is the total number of points and log4n is the PR Quadtree's depth. In terms of space complexity, obtained $8n (\log4 n - \lceil \log4 n/2 \rceil + 8n/3 - 1/3)$ as the worse case number of nodes necessary to build a PR Quadtree with N points.

Analyzing the complexity in terms of modifications done to the algorithm, we found that checking for candidate dense region is a constant step, since the number of points inside a list is limited to the *Thresholdpoints*. On other side, we have included a stop point into split process (build). In this way, worst case space and time complexity are the same of the original *Quadtree*. In the best case we could have just one node with all points inside into it.

3.2 Support Vector Machine for Climate Correlation

Machine Learning algorithms discover the relationships (pattern) between the variables of a system from direct samples of data. These algorithms originate from many fields: statistics, mathematics, theoretical computer science, physics, neuroscience, etc.

We have decided to use a Support Vector Machine (SVM) since we can formulate our problem as a classification problem. More clearly, given a region affected by certain weather conditions, this region can be classified as a stop or nonstop region according to these conditions. One of the supervised machine learning-SVM generally works well for the linear or non-linear classification (Kecman 2001).

In this sense, we choose to implement the SVM by using the machine-learning library (Collobert *et al* 2002) from IDIAP Research Institute in C++. We used Gaussian Radial basis function as the kernel function and σ variance is the regularization parameter.

With the purpose of implementing our method, we have executed the following steps:

- Data filtration: data were organized in a database and filtered for the study of a determined region. Figure 4 illustrates the data schema used. The climate class is specialized into four subclasses, namely: temperature, wind, air pressure, and precipitation. This schema facilitates the addition of new climate information by adding new subclasses to the climate hierarchy.

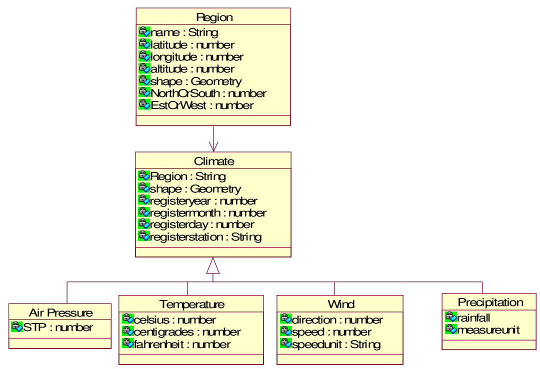

Figure-4- Definition of Climate object and its sub objects

- Find correlated climate elements with stops on migration trajectory: after filtered the correlation is tested by using a statistical function (Stastistical Correlation 2007).

- Specify SVM inputs: from the data stored in the database schema, text files are generated to be the input for the machine learning.

- Record Training, test and validation results: the training/test data sets are composed by attributes which are features correlated to a stop and the corresponding class (target). In our case, there are two classes corresponding to the learning aim: stop or nonstop class. The attributes are composed by the filtered climate data. Then the formal definition of input sample for the machine learning in the training and test phases is defined as:

  $\{(X\mu,t) | X\mu=(x0, x1, x2, x3, x4), t \in \{1,-1\}, \mu=1,2,..,p\}$

  x0: the threshold;

  x1: the precipitation value;

  x2: the wind speed value;

  x3: the air pressure value;

  x4: the temperature value

t = 1: stop class; t = -1: nonstop class

p is the number of samples.

The output of the machine learning is the classification of the given Xμ, that is 1(stop) or -1(nonstop). We separate the samples set into 3 equal sub datasets (train, test and validation) by random.

4 EXPERIMENTAL STUDY

4.1 Data sets

In our study, the data from bird migration was collected by satellite tracking, which recorded White Stork's trajectory stops during 1998 to 2006. In total, we count on 1907 registered locations. The data corresponds to 19 birds observed. Table 1 illustrates how trajectory raw data are organized.

Table-1 Raw data Structure

Bird	X-Coord.	Y-Coord.	Date
Felix	52.305	15.839	25.08.2001
Felix	51.046	17.421	26.08.2001
Felix	49.758	19.914	27.08.2001
Penelope	39.366	31.111	19.03.1999
Penelope	42.237	27.54	23.03.1999

The first column contains the name of bird, the second and third, the pair of coordinate points representing latitude and longitude respectively, and the fourth, the date. Original trajectory raw data were converted from latitude/longitude reference system to (x,y,z) reference system according to ECEF (ECEF 2007) (earth-centered, earth-fixed) Cartesian coordinates. However for the computation of dense region we have just taking into account x and y coordinates. We have computed dense regions for trajectories of whole data without selecting specific time interval.

Concerning climate data, we have downloaded the global surface climate daily data set from the National Climatic Data Center of USA (http://cdo.ncdc.noaa.gov/CDO/cdo) that is recorded by specific weather station and select the interesting fields, showed in Table-2.

Table-2 Climate elements definition

Climate Element (per day)	Unit
Mean temperature	Fahrenheit
Mean station pressure	Mb
Mean wind speed	Knots
Precipitation amount	0.01 inches

Data from discovery dense regions and weather data have been inserted into the object table respective of database schema. In this process, the unknown data (original value is noted as 9999.9 or empty) was estimated by the mean value of the known value.

For the sake of simplicity, we have selected to our test only the region of Barcelona. We defined a rectangle that is centered in Barcelona with length of edge 4 (degrees) in the latitude (+2 in east and -2 in west of Barcelona) and longitude (+2 in north and -2 in south of Barcelona). The stops that are in this rectangle are treated as in the region of Barcelona.

There were 5 records found from 2001 to 2004 (Table-3). Each stop represents a constant population of storks.

Table-3 Stops for birds

STOPYEAR	STOPMONTH	NUMOFSTOPS
2001	9	1
2002	9	3
2003	2	1
2003	9	1
2004	9	1

As we have few number of sample data, it is important in guarding against testing hypotheses suggested by the data (the type III error). Cross-validation (Kohavi 1995) is carried out. We run the machine learning 5 times for different training test and validation data partitions then took the average over the 5 runs.

4.2 Experimental results

Figure 5 shows the results of dense region discovery algorithm execution using different configurations. Tests 1 and 2 have different number of birds. As expected, in Test 1 we have found more dense regions than Test 2. In Test 3, the size of region is bigger regarding Test 1 and Test 2, thus we have found less and more sparser regions.

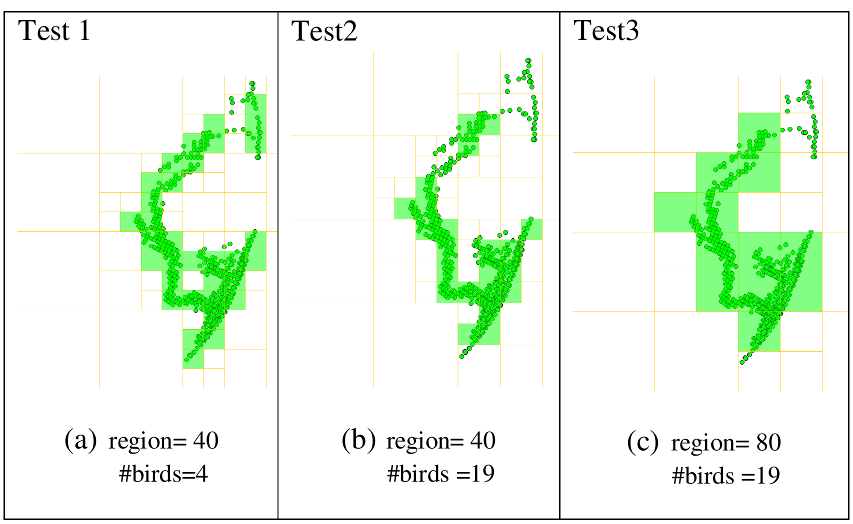

Figure-5 Discovered Dense Regions

The selection of discovered dense regions to be used in the next step of our solution is user dependent. By this we mean, the size of the region and number of birds within it must have a meaning for the user's application.

Then, we calculate the correlation between each mean value of climate condition variable (by month and year) and the number of stops (Table 4). These results were obtained by using the *corr* function(Statistical Correlation 2007) of Oracle 10G that calculates the Pearson's correlation coefficient, which is a common measure of the correlation between two variables X and Y.

Table-4 Correlation between Climate objects and stops of birds

Climate object	Correlation
Temperature	0.218534449
Wind	-0.78507745
Air Pressure	-0.15408494
Precipitation	0.26680673

The correlation values show that:

- The wind has the large negative correlation with number of stops.
- The air pressure has the small negative correlation with number of stops.
- The temperature and precipitation in this period (2001 – 2004) has small positive correlation with number of stops.

On the other hand, the average of misclassification over 5 runs (Table 5) shows that the average misclassification is less than 20%, so we can confirm that at least 80% predictions are correct. Then the pattern that SVM found by training is reliable.

Table-5 Average number of misclassification over 5 runs by using SVM

# of run	Test Stop	Test Nonstop	Validation Stop	Validation Nonstop
1	1	2	1	2
2	1	0	2	1
3	2	0	2	0
4	1	2	1	0
5	2	2	1	2
Total: Avg	16.25%		15%	

4.3 Discussion

Regarding common problems described in (Ni *et al* 2007), we discuss how our solution copes with them, as follows:

- Answer loss: In our tests with real data we did not face this problem because bird's trajectory is constrained implicitly by some specific path. This property guarantees that there is no discontinuity among stop points. However, this is only true if our data set has no loss of information, such as positioning devices lost signal during some time interval.

- Overlapped regions: This problem can occur with our current solution if we report areas before split. We check for a dense region in the center of the area being splitted, if it is found, it is reported. After, the sub-areas of this area might continue to "receive" points. Thus, they could reach the Threshold points and also be reported.

- Lack of local density: this problem is possible to happen with our current implementation since we have defined a dense region with a spatial extent with fixed size. To solve this problem we could use the same approach discussed in (Ni *et al* 2007). In this sense, we will have similar point density definition. By reporting all dense points, a dense region is constructed.

Although we have optimistic accuracy prediction results, we have to be caution. The size of samples data is small and we cannot be sure that it is extensible for new samples of data. Another reason that we cannot be totally optimistic is the correlation between stops and three of climate conditions (temperature, air pressure, and precipitation) are not strong. As the recorded trajectory data duration are only limited from 2001 to 2005 in Barcelona, the global warm in the short term has no great influence on the behavior of bird migration. It's not the result that we wanted. We expect in the future years, more trajectory data will be collected and change the situation.

5 RELATED WORK

Discovering regions that are dense in number of moving objects within a trajectory is a known problem discussed in the literature (Ni *et al* 2007, Jensen *et al* 2006, Hadjieleftheriou *et al* 2003).There are some solutions related to our idea of implementing a specific algorithm to respond this query. In general, these solutions are not adequate since they present some undesired behaviors (Jensen *et al* 2006, Hadjieleftheriou *et al* 2003) or because they do not find the requirements of our problem. For example, (Ni *et al* 2007) do not look at historical data, it just takes into account data arriving at the present moment of the analysis in the intention of discover the trajectory next step. Here it is true that we can make the abstraction of a time interval as a timestamp and employ it. But on the other side, this solution tries to predict the trajectory just based on the trajectory data itself (point and direction), instead, we are interested in explore semantic data, the weather, to solve the problem.

Other approach commonly used is data mining clustering algorithms (Carneiro *et al* 2008). These methods are useful for discovering threshold, but not for reporting clusters above a given threshold which was stated as a requirement in our problem definition.

Relevant related work on analysis of the correlation between weather and bird migration use predefined regions as candidates for dense regions (Zalakevicius 2001). Conversely, in our proposal we do not have a-priori knowledge about such regions, which are discovered dynamically.

In (Chernetsov *et al* 2005)they are concentrated on the correlation between density of total birds (many kinds of birds, the population of the birds are more interesting for the research of the density) of the weather condition. But they have similar weather condition modeling.

The difference from our work is that we work only on one kind of bird. So the population will not change for this kind of bird (or rather small, the number of birds will be almost constant). We choose the fix region for only one kind of bird to study influence of the local weather condition on the choice of stop.

Some mathematicians are working on the prediction of migration path (the whole trajectory) by the stochastic models that are built on the observation of trajectory (Sheldon *et al* 2007).

6 CONCLUSIONS

In this work we have propose an integrated solution for correlating climate condition s with bird migration trajectory. More specifically, given a data set of migration trajectories and climate conditions information, our solution computes dense regions where birds are used to stop and describes the correlation between these regions with climate conditions. We have obtained significant correctness (80%) for the prediction of bird stop give a climate condition.

Our approach can be easily extended in order to investigate more climate conditions such as the cloudiness, sunshine length, and pollution levels basing on our extensible model. This extension can be done by creating a new subclass of climate condition, filter the data, find the correlation and construct new training set for the machine learning.

Regarding the algorithm for finding dense regions, we can extend the solution adding a new phase for discovering the lost areas discussed in 4.3. We can sketch this extension by traversing the tree taking into account only the leaf nodes not tagged as dense regions or empty nodes.

7 REFERENCES

Beaumont, L.J., McAllan, I.A.W. and Hughes, L. (2006) “A matter of timing: changes in the first date of arrival and last date of departure of Australian migratory birds”. *Global Change*

Biology 12: 1339-1354.
Bouten, W., Belle, J. V., Gasteren, H. v. , Vrugt, J. A., Shamoun, J. and Buurma, L. (2005) “Predicting Bird Migration: Data-Driven versus Concept-Driven Models”. International Bird Strike Committee.
Carneiro, C., Alp, A., Macedo, J. and Spaccapietra S. (2008) “Advanced Data Mining Method for discovering Regions and Trajectories of Moving Objects: Ciconia ciconia Scenario”, *AGILE Conference* (to be published).
Caswell, H., Moller, A. and Fiedler, W. (eds) (2006). *Birds & climate change (Advances in ecological series)* (2nd Ed.) Academic Press.
Chernetsov, N. and Huettmann, F. (2005) “Linking Global Climate Grid Surfaces with Local Long-Term Migration Monitoring Data: Spatial Computations for the Pied Flycatcher to Assess Climate-Related Population Dynamics on a Continental Scale”. *Lecture Notes in Computer Science. Computational Science and Its Applications – ICCSA*. Springer Berlin / Heidelberg. pp 133-142.
Collobert, R., Bengio, S. and Mariéthoz , J. (2002) “Torch: a modular machine learning software library”, *Technical Report IDIAP-RR 02-46, IDIAP*.
ECEF, Earth-centered, Earth-fixed Reference System Conversion tool. (2007). http://www.jstott.me.uk/jcoord/.
Hadjieleftheriou, M., Kollios, G., Gunopulos, D. and Tsotras, V. (2003) “On-line discovery of dense areas in spatio-temporal databases”. *Proccedings of the 8th International Symposium on Spatial and Temporal Databases* (SSTD).
Jensen, C., Lin, D., Ooi, B. and Zhang, R.(2006) “Effective density queries on continuously moving objects”. *Proceedings of the 23nd International Conference on Data Engineering* (ICDE).
Kecman, V. (2001) *Learning and Soft Computing Support Vector Machines, Neural Networks and Fuzzy Logic Models*, The MIT Press, Cambridge, MA, 608 pp.
Kohavi, R. (1995). “A study of cross-validation and bootstrap for accuracy estimation and model selection” *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence 2* (12): 1137–1143. Morgan Kaufmann, San Mateo.
Ni, J. and Ravishankar, C. (2007) “Pointwise-Dense Region Queries in Spatio-temporal Databases”. *Proceedings of the 23nd International Conference on Data Engineering* (ICDE) Pages 1066 – 107
Pemmaraju, S. V. and Shaffer, C. A. (1994) “Analysis of the Worst Case Space Complexity of a PR Quadtree” . *Information Processing Letters*,Vol. 49, N 5,p 263-267.
Samet H. (1989) *The Design and Analysis of Spatial Data Structures.* Addison Wesley.
Samet, H. (1995) *Applications of Spatial Data Structures: computer graphics, image processing and GIS.* Addison Wesley.
Sheldon, D., Elmohamed, M. A. and Kozen, D. (2007) “Collective Inference on Markov Models for Modeling Bird Migration” .*Neural Information Processing Systems Conference*.
Statistical Correlation (2007). http://en.wikipedia.org/wiki/Correlation
Vagg, R. and Hepworth, H. (2006) “Migratory Species and Climate Change: Impacts of a Changing Environment on Wild Animals” UNEP / CMS Secretariat, Bonn, Germany. 68 pages.
Zalakevicius, M. (2001) “Bird Migration and the Climate: A review of the studies conducted in Lithuania in the context of Climate Change”. Acta Zoologica Lituanica, Vol 11, Nr 2.