Time, Points and Space - Towards a Better Analysis of Wildlife Data in GIS

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)
vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich
von

Stephan Imfeld
von
Lungern OW

Begutachtet von
Prof. Dr. Kurt Brassel
Prof. Dr. Bernhard Nievergelt
Dr. Britta Allgöwer
Prof. Dr. Peter Fisher

Zürich 2000

Die vorliegende Arbeit wurde von der Mathematisch-naturwissenschaftlichen Fakultät der Universität Zürich auf Antrag von Prof. Dr. Kurt Brassel und Prof. Dr. Robert Weibel als Dissertation angenommen.

How to Catch Running Animals with GIS and See What They're up to

Abstract

Geographical Information Systems are powerful instruments to analyse spatial data. Wildlife researchers and managers are always confronted with spatial data analysis and make use of these systems for various tasks. One important characteristic of the animals unter investigation is their locomotion. Thus the temporal aspects are important, but unfortunately GIS are almost ignorant concerning the analysis of the temporal domain.

This thesis is trying to provide a new perspective on how to analyse moving point objects within GIS. A conceptual shift is performed from a space centered view to a way of analysing spatial and temporal aspects in an equally balanced way.

For this purpose the family of analytical Time Plots was developed. They represent a completely new approach of how to analyse moving point objects. They transform the data originating from an animal's movements into a representation with two time axes and one spatial axis that allows for an effective recognition of spatial patterns within the data. Some of the easier Time Plots make use of the *Temporal Data Frames* concept, another analytical framework using exploratory data analysis techniques to analyse and search for regularities in temporal point data. It is especially useful in the exploration of temporal aspects such as solar or lunar cycles.

The *Radial Distance Functions* developed and elaborated are a new method to analyse the environment around a point object. They can be thought of as an extension of the second-order functions applied to areal data. They are also extended to be used in a more dynamic way of analysing the movements of an animal in its environment.

The methods developed were applied to synthetic data as well as different animal species including ants, bats, woodstorks, badgers and lynx to test and illustrate their usability.

Due to the fact that temporal aspects become more and more important in GIS, a system similar to the (spatial) coordinate transformation between different coordinate systems needs to be developed for the temporal domain.

Zusammenfassung

Geographische Informationssysteme (GIS) sind mächtige Systeme, um Raumdaten zu analysieren. Wildtierbiologen sind oft mit der räumlichen Datenanalyse konfrontiert und benutzen daher diese Systeme für verschiedene Aufgaben. Ein wichtiges Charakteristikum der untersuchten Tiere ist ihre Lokomotion. Dabei spielt der zeitliche Ablauf ein grosse Rolle. Unglücklickerweise stellen jedoch GIS kaum Analysemethoden für die Analyse der zeitlichen Dimension zur Verfügung.

Die vorliegende Arbeit versucht nun, eine neue Perspektive aufzuzeigen, wie bewegende Punkt-Objekte im GIS analysiert werden können. Dazu wird ein konzeptioneller Wechsel von einer raumzentrierten Sicht zu einer Analyse vollzogen, in welcher räumliche und zeitliche Aspekte in gleichem Mass einbezogen werden.

Zu diesem Zweck wurde die Familie der *Time-Plots* entwickelt. Sie stellen einen komplett neuen Zugang dar, um bewegende Punktobjekte zu analysieren. Die raum-zeitlichen Daten von sich bewegenden Tieren werden zu einer Repräsentation mit zwei Zeitachsen und einer räumlichen Achse transformiert. Dies erlaubt in einer effizienten Weise, räumliche Muster in einem Datensatz zu erkennen. Einige der einfacheren *Time-Plots* nutzen das neue Konzept der *Temporal Data Frames*. Dazu werden Techniken der Explorativen Datenanalyse eingesetzt, um nach Regelmässigkeiten in den zeitlichen Punktdaten zu suchen. Besonders nützlich ist diese Methode bei zyklischen Zeitaspekten wie beispielsweise Sonnen- und Mondzyklen.

Mittels der neu entwickelten und ausgearbeiteten Methode der *Radial Distance Functions* (RDF) lässt sich die Umgebung um ein Punktobjekt analysieren. Diese Funktionen können als Erweiterung der second-order Funktionen auf Flächendaten angesehen werden. Die RDFs werden zudem erweitert, um die Dynamik in den Umgebungsveränderungen bei sich bewegenden Tieren erfassen zu können.

Die hier entwickelten Methoden wurden zuerst an synthetisch generierten Daten und später auch an realen Daten einer Ameise, einer Fledermaus, Amerikanischen Nimmersatten, Dachsen, und Luchsen getestet und ihre Verwendbarkeit aufgezeigt.

Angesichts der Tatsache, dass zeitliche Aspekte in GIS immer wichtiger werden, wäre es sinnvoll, ein System zu entwickeln, welches ähnlich der (räumlichen) Koordinatentransformation die Transformation zwischen verschiedenen Zeitsystemen erlaubt.

Contents

List of Figures

List of Tables

1

Introduction

1.1 Problem Statement

This thesis is an interdisciplinary work between Geographical Information Science and Wildlife Biology. It is focusing on methodological aspects that are arising when both disciplines meet. The aim of this work is the development of new analysis methods for movements of animals in the environmental space.

I would like to start off with a basic idea: Animals, in contrast to plants, once upon a time have invented locomotion. The definition of locomotion includes two fundamental concepts: Space and Time. Spatial data, the material Geography is working with, is a very complex matter. In the past few decades powerful instruments have been developed for handling and analyzing such data. The diffusion of these geographical information systems (GIS) into disciplines outside Geography started in the late 80ies and early 90ies. Today we are at the beginning of a new era for GIS where they become almost omnipresent and even available in standard office computer software packages.

Wildlife researchers are always confronted with spatial data analysis, so it is not astonishing that they also started to make use of these systems. Impressive applications were developed and results presented, but a major problem still exists in the handling of wildlife data in GIS. Although powerful in analyzing the spatial domain, GIS are nearly ignorant concerning the handling and analysis of the temporal domain. Thus, major efforts today on this topic are directed towards the representation of time within GIS.

The work presented here aims to go two steps ahead of representational issues. The first step is to extend the understanding of time as a single dimension with a perception of time as a congregation of 'multi-faceted' temporal aspects. The second and main step that goes beyond representation is the development of powerful analysis methods that incorporate both spatial *and* temporal aspects in an equally balanced way. Hence, the main question posed in this work is the following:

How can we include temporal aspects in the analysis of the spatial behavior of an animal within a GIS environment?

The intentions underlying the development of new concepts may be better explained by presenting the current state and its shortcomings. I shall try to do this in the form of the following list of common misconceptions and shall then explain why they need to be overcome.

1. Animal movements are random walks.

2. A movement can be described by measuring the shape of the path.

3. The spatial aspects of the environment can be summarized by a single value.

4. A certain analysis of the data is performed even if it takes a large amount of time.

5. A temporal GIS is a GIS which includes an adequate representation of temporal aspects.

6. Changes in the use of space by an animal can be adequately analyzed by creating separate maps of the data for each calendar month of the observation period.

The following shows some of the misconceptions and errors in reasoning in the above statements.

(1) Random is often confused with the finding that underlying patterns cannot be identified. Animals obviously have clear patterns in their behaviors, but sometimes these patterns are too complex to be identified with simple methods.

(2) The temporal dimension is left out. Without time no changes are possible. This means that two things are vital when analyzing movements: space *and* time.

(3) Summarizing information either means concentrating or loosing information. The art of analysis is to make complex phenomena easy to understand.

(4) Most of today's data in wildlife biology (and probably many other fields, too) are not analyzed to its full potential due to a lack of time. The ease and speed with which an analysis can be performed is crucial for its application in the everyday life of a researcher. If one needs to calculate the time to sunrise for 9000 observations by hand, no one would do it. If it can be done by pushing a button, it is performed whenever there is a chance to gain new insights into the data.

(5) The adequate representation of temporal aspects in GIS is only one step towards a Temporal GIS. The representation of a line and a polygon in a computer software does not make it a GIS until analysis functionality is added to the system. They are mostly called graphics, drawing or cartographic programs instead. The same is true for time aspects. The representation itself is only the first requirement. To make full use of the information, the analysis part of a spatio-temporal information system needs to be developed, a part that is largely missing today.

(6) The most popular way of analyzing temporal data in wildlife research is to plot the data on a separate map for each observation period. Calendar months presumably have no meaning to animals. By applying such a frame-set to the data, the researcher implicitly (and probably unintentionally) assumes that the changes he or she is looking for overlap with these time periods. If the changes in the data are shorter than the calendar months, they will not be detected. In the case the changes are occurring at a larger time range or

in between them, there is a potential danger that fluctuations in the data are considered to be significant differences.

Here, a completely new approach was developed. A conceptual shift is presented from a space-centered view to a time-centered view, where time can even become the dominant dimension in the analysis (chapter 5).

With this thesis I would like to pursue the following objectives. First I would like to provide a better understanding of time and temporal aspects. It should lead to improved and more accurate analysis methods for wildlife research data within geographical information systems. In the context of GIS it will show that *temporal geographical information systems* not only need a sophisticated data model to represent temporal aspects, but in addition that new and improved analytical methods need to be developed to adequately analyze the data. This work is intended to provide new analytical methods to achieve better insights into changes in the spatial behavior of animals. It will also create analytical instruments to perceive and define biologically 'meaningful' phases in a wildlife researchers data. Even though most examples and statements are made primarily with a focus towards wildlife research, the methodology is of course applicable in any context concerned about spatial movements of point objects.

1.2 Overview

Space can be classified in different ways. The most familiar way is the classification in zero, one, two and tree dimensions representing point, line, area and volume. Animals are often represented as point objects in GIS. For the purpose in consideration here this is being considered adequate, so in the following I will focus on point objects. The term 'animal' used throughout the text can most of the times be considered as an example for a 'moving point object' in general. I assume that the readability and understanding is made easier when using a concrete example, so I will mostly use the term animal instead of 'moving point object'.

Geographical Information Science has a variety of facets. *Database issues* are fundamental to any GIS. A lot of research in this area is currently being conducted on object-oriented systems, interoperability and the integration of time (e.g., Yearsley and Worboys, 1995; Vckovski, 1997). Temporal GIS research can be grouped into three topics: database, query and analysis. In contrast to the first two topics, the last one has not received much attention yet. Chapter 3.4 will give a short overview of temporal GIS research.

Spatial analysis, on which this research is concentrating, is at the core of GIS applications. Basic functionality can be found in all 'off-the-shelf' systems, but as the diffusion of GIS into other research disciplines increases more analysis functions are needed. Especially in biological research a lot of the acquired data is in the form of point data, e.g. locations of plants or animals. In spite of the relatively large literature on the analysis of point data (not to be confounded with point measurements of continuous phenomena), GIS provide little analysis methods for this type of data. Chapter 3.3 will give an overview over the currently available analysis methods for point objects. In the past few years there has been a lively discussion going on about the integration of statistical analysis of spatial and non-spatial attributes. Several ways of coupling GIS with statistical software have been provided in the literature as a proof of concept, mostly using GIS as a visualization front end to the statistical package, but true integration is still lacking (MacLennan, 1991; Lippert-Stephan, 1996).

In chapter 2 the characteristics of points and time specific to wildlife research are elaborated. After that an overview of the current analysis methods is provided in chapter 3.

In part II concepts and new analytical methods are developed. The main focus is set on the integration of various temporal aspects in the analysis of point objects.

A relatively new research field in GIScience is addressing problems of how to handle uncertainty in GIS with its different aspects from boundary location accuracy to classification uncertainty (Aspinall and Pearson, 1995). It would be tempting to include such aspects in the development of new analysis methods. Unfortunately, as this is still a very difficult aspect to handle for geographic information, it would not be realistic to try including it in a first approach for analysing moving point objects. This will need to be addressed in later studies.

In the following I will examine some issues of the main data type for this study, the point object.

I

Background

2

Fundamentals

2.1 Types of Points

For a lot of GIS-technicians points come in very handy in representing a diversity of issues. Lanterns and road signs are easily represented as points in GIS. Hill peaks, trees, badger setts and observations of animals can be handled as points as well. In transportation the locations of ships, trucks or emergency vehicles are tracked by means of GPS where the data are measured and stored as points with two coordinates defined.

Obviously a variety of things are handled by the same representational feature class. In a mathematical sense points are zero dimensional objects with no extent. In all of the cases above this means that the original elements are reduced until their size converges to zero. This representational 'equifinality' of all these objects might mislead to the temptation that they could be analyzed in a similar way. In the following I will try to clarify the reasons and properties why this is not the case.

There are many ways to approach a classification of point data. In this work I will distinguish the four following aspects often implicitly or explicitly inherent to a point object in a GIS:

1. Generator/Generation Mechanism
2. Object type
3. Sampling method
4. Representation

Researchers often wish to know the generation mechanism which created a certain spatial pattern found. This mechanism is most often unknown and can often only be inferred from the resulting patterns. Since it is seldom known in advance, there is not much use in classifying point objects upon this first aspect.

The second aspect can be clarified by asking the following two questions:

2a. Does a single object change its location?
2b. How long does the object exist?

The ordering of the questions does not really matter. Figure 2.1 illustrates the two questions and a possible classification scheme. The distinction between moving and static objects seems to be trivial but is nonetheless important as can be seen below. The question on the life span of an object splits up several types. Objects that only exist for a fraction of a second can be distinguished from ones that exist for a longer time, for multiple times or even infinitely. Objects that change their location cannot have a life span of 0, because they must exist at least at two different locations which for zero dimensional objects requires the existence at two different times.

Cressie (1993) only considers temporal point objects which do not change their location. He distinguishes only two types. They are called *space-time shock point processes* and *space-time survival point processes*. For the first type, events occur instantaneously over both time and space. Conversely, for the latter type, events are born at some random location and time, and then live for a random length of time. These are the only temporal point patterns considered by Cressie (1993). The other types of point objects considered here (figure 2.1) were missing, probably due to the fact that his work is concentrating on examples originating from the field of botany.

In the present work I will concentrate on objects that change their location and exist for a certain period of time.

One of the factors which is often 'hiding' the type of object is the sampling scheme applied when collecting data (figure 2.2). For example in wildlife telemetry locations of animals are often taken at certain intervals such as days or weeks. They are then recorded as point objects with two coordinates and a time identifier. In a subsequent analysis they are often considered in an equal manner as if they were objects existing only at the observed locations. The locations in between samples are ignored even though the animal must have passed some of the area between the samples.

The sampling scheme needs to define at least three aspects which can be evaluated by the following three questions:

1. What is the temporal sampling scheme?
2. What is the spatial sampling scheme?
3. What kind of boundary is defining the sampling area? (infinite/none, bordered)

The temporal sampling scheme is of concern here. When applying different schemes (figure 2.2) the resulting data and their representation is changed accordingly. In the past most of the data in spatial studies on animals were applying some kind of a snapshot sampling scheme. This resulted in data sizes which could still be handled. The researchers were mostly concerned about temporal autocorrelation of the data in order to be able to apply certain statistics that require independence among observations. They almost never considered the accurate description of an animal's movement as a matter of interest (e.g., White and Garrot, 1990), although such an approach seems to be reasonable in the case of moving objects. We might need to apply a so-called lateral thinking (de Bono, 1972) to overcome these limitations and start developing new analytical methods.

It is difficult to judge whether the available statistical methods required the sampling scheme outlined above or that the sampling methods used stimulated the development of the statistical methods. Anyhow, the fact is that today's analytical methods for spatial point objects concentrate on snapshot oriented temporal sampling schemes. Most if not all of them ignore the temporal domain

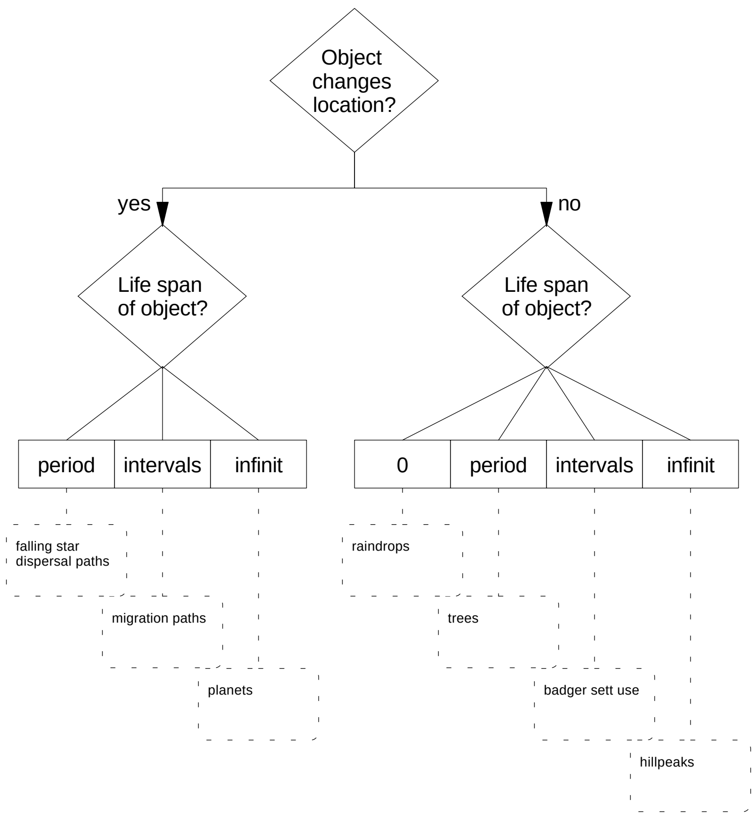


Figure 2.1: Classification of Point types.

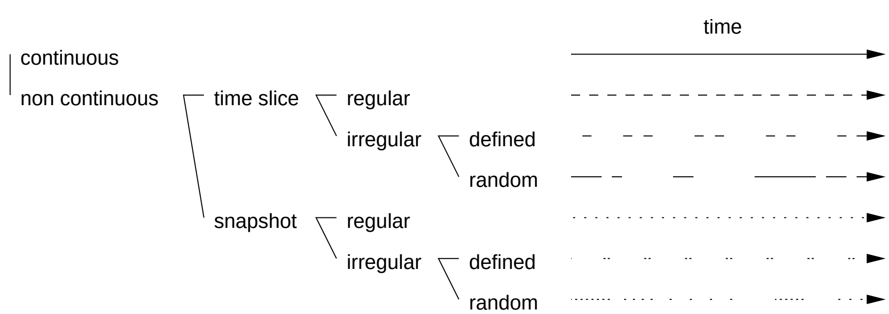


Figure 2.2: Classification of temporal sampling schemes.

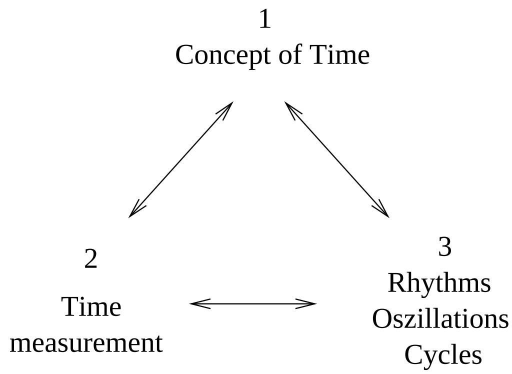


Figure 2.3: The three interrelated aspects of time.

and hence make no distinction between the analysis of static and mobile objects. Even in newer text books on spatial statistics (e.g., Arlinghaus, 1996) chapters on sampling methods completely neglect the need for defining temporal sampling schemes, obviously due to the fact that no established methods for analyzing such data exist.

We need to break up this interdependence of development of analytical methods according to the available data and vice versa sampling data according to the available statistical methods.

2.2 Time and Rhythm

Like many fundamental concepts, time is difficult to define in words, but most of us refer to it and use it whenever we look at a clock. We measure time in seconds, minutes, hours, days, and years. The internationally agreed definition of the second is the time it takes for 9,192,631,770 cycles of a frequency that resonates with an atom of cesium.

What is 'time'? The first answer that often comes to mind is something similar to an immaterial linear continuum where only one point on that continuum called present exists. There are no start and end points similar to space. Being asked what time it is one gets answers like 12 hours and 23 minutes past midnight, a measurement technique which restarts counting every 24 hours. Starting from these three contrasting views of time I would like to shed a little light on several aspects of the perception of time.

Figure 2.3 illustrates these three fundamental aspects mentioned above. They are very much interrelated: (1) Time as a fundamental concept, its measurements (2) and the applications of oscillations or rhythms (3).

Time in the Hindu philosophy is perceived as a wheel instead of being conceived of as linear, and inherent in the concept of a spinning wheel is the notion of rhythm (London, 1997).

In the GIS literature time is often considered as a third (or fourth in 3D systems) dimension in addition to space. This resulted to the often used term of the *space-time-cube*. At first this looks very appealing, but it might be misleading. It mixes the continuous spatial aspect, where two or more spatial points can co-

Table 2.1: Time anchors used in various disciplines and situations.

Anchor	Examples
'Big bang'	assumed start of the universe
01.01.0000	Start of the Year of Birth of Christ
01.01.1900	Start time in the DOS/Windows operating system
01.01.1904	Start time in the Macintosh operating system
02.12.1967	my birthday
01.01.1970	Start time in the Unix operating system
Solar year	
Lunar eclipse	indigenous peoples
Month	
Day	starts at midnight (of time zone)
Hour	
Minute	
Second	
Present	Anchor in Palaeontology

exist, with the temporal aspect, where only one time-point can exist at a time. In addition there is only one direction possible on the time axis pointing from the 'past' to the 'future'. I would propose a new term called the *space-time-elevator* which symbolizes the two above-mentioned differences between space and time.

I will first go through various aspects of time measurement. Then the use of time in our language is discussed. In the last section about the conception of time I will provide an overview of the handling of time in computer science.

2.2.1 Measurement of Time

In our every day's use of time we make a clear distinction between time (e.g. 5 o'clock) and date (e.g. 1.1.1999). By 'date' we mean the actual time measured in days passed since the date 1.1.0000. With 'time' we use a cyclic measurement of 24 hours (= 86400 seconds = 794243386928000 oscillations of a cesium atom). This repeating measurement makes it easy to refer to a different time as for example when we make an appointment the same time tomorrow.

Since time does not have a start nor end point, we need to define some anchor point or multiple anchor points to which we can refer in our measurements. One was mentioned above, the date 1.1.0000, often referred to as the year of birth of Christ. But there are several others. In palaeontology the time span considered is so huge that the term 'present', however undefined it is, can be considered as a time anchor. Here the occurrence of certain species in sediments is also used as relative time anchors. The 'big bang' is sometimes considered as an anchor or starting point e.g. in geology, physics and astronomy. For different time scales different anchors and resolutions have been defined, starting from astronomical and geological time to paleontological and biological time until the finest time intervals considered in atomic oscillations and processes. Table 2.1 shows some of the common time anchors.

As it is the case with spatial data, data with temporal characteristics can have both exact measurements or boundaries or less precise specifications. At first, the range of temporal accuracy seems to be much larger going from sub-second measurements to an accuracy of several hundreds of million years in geology. But the corresponding values for spatial measurements also range from micrometers to maybe several thousand kilometers. Comparing the order of magnitude for time of about 10^{16} to the one for space of around 10^{10} there is only a difference of a factor of about 1:1 million. This should not cause a computational problem.

Originating in the different time scales used in the various disciplines several 'time systems' have been defined. In today's geographical research we often come across terms such as *time zone, local time, UTC or standard time, solar time* and more specific terms as for example *GPS time*. They can be considered as the equivalent in time for the different coordinate systems used for describing space. An event recorded in one system (e.g. time zone 2) may have to be translated twice if it will be used in another application: once for the (spatial) coordinate system, and once for the time system. The second translation is an aspect completely disregarded by today's GIS technology and will need attention in future temporal GIS research.

2.2.2 What About the Fuzzy and Abstract Temporal Terms?

The 'wish' of every computer is to handle well defined 'crisp' data. Our everyday use of time contrasts these needs in an obvious way. Terms like era, age, epoch or season are most of the time ill defined. Sometimes crisp definitions were proposed, but they often have two drawbacks which prevent a wider application. Firstly, numerous people have tried to provide crisp definitions, resulting in a variety of statements. Secondly, a lot of temporal aspects simply cannot be defined crisply due to their nature. For example the four seasons summer, autumn, winter and spring are defined as three-month periods. They are often used as a temporal raster in the analysis of biological data. The intuitive definition of these seasons is somehow related to temperature, weather and vegetation conditions which are different in each season. But they are not constant over the years. The use of three month periods for the definition of seasons becomes even more difficult when traveling from north to south over the globe. The seasons are different in Europe than in Australia or in the Tropical Zone and have smooth transitions in between these areas.

This is only one example of temporal terms often used in our languages and in research. There are relative terms used to describe temporal relations such as *before* or *after*, *lately* or *recently*. The first aspect of these terms stating that it is in the past or future is well defined, whereas the second aspect that there is a close (temporal) proximity to the reference time is only vaguely expressed. This becomes even more pronounced in terms like *while, during* and *contemporary*.

In biology such terms are also widely used. Breeding season, migration time, generation time or descriptions of the age of an animal like juvenile or adult are only a few examples.

There are at least two ways to improve this situation for computational purposes. The first is to close the semantical gaps in the terminology of temporal aspects. Several authors (e.g., Korte, 1997) discussed various forms of temporal relationships between two objects. Nevertheless, a coherent framework and

Table 2.2: Terminology used for transaction and valid time

Author	Transaction time	Event time	Valid time
Lum et al. (1984)	physical time	logical time	
Lum et al. (1984)	database time	world time	
Snodgrass and Ahn (1985)	transaction time		valid time
Kemp and Kowalczyk (1994)	database time	world time	
Yearsley et al. (1994)	database time	event time	

the corresponding semantics are still missing for efficient communication and application in computers. The second improvement might build on the ongoing research on fuzzy sets, boundaries and objects in the spatial domain (Burrough and Frank, 1996; Schmitt, 1996; Zadeh, 1965). Such concepts are also needed in the temporal domain as illustrated by the examples above but have not been applied yet.

2.2.3 Representation of Time in Computer Science

From a technical point of view there are several aspects to be considered. First we need to distinguish between the time an object was created or updated in the database (*transaction time*) and the times the object exists in reality (*valid time*). The terminology used here is confusing because of the variety of terms in use. Table 2.2 lists some of the frequently used terms and synonyms. Most authors make use of *transaction time*, but are concentrating on event-based models where only the changes are recorded and stay valid until another change occurs. Aside from these two representations, it is always possible to implement a so called user-defined time in databases by storing these values in standard (e.g. integer or floating point) fields (user-defined time). But this is not the approach needed here. The aim is to create a universal representation of temporal aspects in databases.

Newer publications, especially those using an object-oriented approach, accept the need for *valid time* defined as the duration for which an object exists.

Recent discussions on temporal representation propose the creation of instant, interval and period data types with both variable granularity and uncertainty (Kemp and Kowalczyk, 1994). An instant type represents a single point in time (e.g. 1.1.1990 01:00:00h) whereas the period type is defined by two instants. The interval type is an unanchored duration of time. Handling unknown or undefined instants is done using special values representing for example infinity ($-\infty$, ∞). It also would be possible to use a 'time ray' type to represent events (instants) that are finished (or started) without a defined start (end).

Examples for applications of these representational aspects of time can be found in Yearsley and Worboys (1995); Bagg and Ryan (1997).

How is time really implemented in today's computer systems? Traditionally most software support some date type. It is maintained as a number of days passed since a specific date (see table 2.1). A (day-)time field was often neglected and in some systems it is still missing (e.g. ESRI's Arc/Info). In several systems, for example SQL, the SAS system and the Oracle8 system, a DateTime field was

Table 2.3: Examples of possible rhythms that can influence the behavior of animals.

Type	Examples
astronomical	24h day, day-night, moon phase, year, tides (amphidromic points, cotidal contours), sunset, sunrise, altitude of sun above horizon, twilight start, end and length (civil, nautical and astronomical), day length, moon rise, moon set, percentage of moon illuminated, altitude of moon above horizon, moon phase
meteorological	weather (rain showers, wind, freezing, drying, rainy-seasons, snow cover), temperature
biological (a-spatial)	diving times, physiological circadian rhythms (e.g. fruit fly hatching)
biological (spatial)	abies fruiting, rutting season, migration
complex	tides
human	hunting season, working hours, traffic rush hours, week, holidays, vacations, spare time activities, train schedule (30'/first-last), closing time of restaurants or bars, second, minute, hour, calendar month, decade, century
technical	"weir flooding"

introduced holding both date and time in a single field. In addition Oracle8 also knows about the Julian date. In the Unix environment a similar time structure is defined, which counts the number of seconds since 00:00:00 of January 1, 1970. It has functions to access the contents of the time structure as year, month, day, hour, minute, second and also contains a flag for the daylight saving time. Aside from this time structure Unix also keeps track of the current time zone which contains the difference between UTC and the local time, although this is done at the system level and not at the application level. The last two fields (daylight time, time zone) are seldom seen in standard applications.

It is clear from these few examples that there is a variety of implementations of time in today's computer systems. Most of them only cover a fraction of the information needed to reflect the full temporal information needed when handling geographical data spread all over the earth.

2.2.4 Rhythms Considered for Wildlife Data Analysis

Rhythms are sequences repeated at relatively regular intervals. In the biological and medical literature a lot of work has been done in the area of endogenous and exogenous rhythms (e.g., Borbely and Tobler, 1996; Camargo et al., 1999; Cederlund and Lemnell, 1980; Turner, 1980; Kurt, 1991). Most of this work has been done concerning the day/night rhythm. Wildlife animals are influenced by a wide range of natural and human influenced rhythms. Table 2.3 provides some examples which go far beyond a simple light/dark phase analysis.

Many of these rhythms are linked to quite complex spatial phenomena which require a lot of complex calculations to be analyzed, a task often beyond the capacity of a wildlife researcher. This is probably the main reason why such analysis is almost never performed in field research.

Rhythm is closely related to *time lags*. Analyzing rhythms and their origin often requires the inclusion of time lags. In some cases the amount of time used in an analysis to explore effects with a time lag might be known. In others it is unknown whether such an effect exists at all. In such cases good guessing or a tremendous amount of work might reveal such an effect. This makes the analysis task even more complex.

The title of this section might imply that all rhythms for the analysis of wildlife data could be determined in advance. This is of course not possible. Hence in a first stage of this research I want to include the most apparent rhythms in the framework, so that patterns caused by these can be easily discovered (Chapter 4). In a second stage, it will be the aim to build a generalized rhythm finder without predefined parameters (Chapter 5).

2.3 What is a Spatio-Temporal Pattern?

In our everyday language the term *pattern* is often used as a synonym to regularity. We use pattern in combination with time, sound, space and also to describe the structures behind our thinking [1]. There are two opposites of the term pattern. If we are not able to recognize any pattern, we may state that there is no pattern. The second way often used in the case of spatial data to describe the absence of a pattern is regarded as a separate class of pattern, the random pattern. Hudson and Fowler (1966) define pattern as '.. the zero-dimensional characteristic of a set of points which describes the location of these points in terms of the relative distances of one point to another'. This definition has strong resemblance to the hypothesis no. 3 in the introduction, that in this case space and spatial arrangements can be described by a single value.

We often use the term *spatial pattern*. What is meant by that expression? In the case of point objects this question is mostly reduced to the aspect of whether the objects are arranged randomly, clustered or evenly distributed. In the latter two cases inferences about the underlying mechanism are drawn from the data themselves as whether the distribution of the objects may be formed by a Poisson-, Cox-, Markov- or a similar kind of process. In the case of linear elements the literature is astonishingly quiet about statements of patterns in spite of their wide use as a representational form in Geographical Information Science. Common examples include elements such as streams, roads or border lines of certain habitat descriptives. In the case of polygonal data models a variety of aspects can be considered for analyzing pattern. Juxtaposition, interdispersion and measurements of size and shape are a few examples for describing certain aspects of arrangements and their regularity in polygon structures. Fields as another type of data model have their own set of descriptives for characterizing inherent patterns (e.g. variograms, Fourier series etc.). The idea of patterns

[1] e.g. Longman Dictionary: Pattern: A) A regularly repeated arrangement (especially of lines, shapes, etc. on a surface or of sounds, words, etc.). B) The way in which something happens or develops.

with this kind of data is often viewed as regularities as surface trends or periodical wave forms.

Most definitions of *pattern* state that pattern should be defined independently of scale. Aside from the fact that this is very difficult to achieve, it also contrasts the common use often found in the biological literature. As an example an animal using one hectare as its homerange is considered to show a different pattern from another one using 100 hectares just because of the areal difference. The distribution patterns often used in faunistics are another example where the requirement of scale independence is ignored. In linguistics it may be argued that these are two different terms hiding behind the same word. For scientific information transfer this is an awkward situation.

The terms *pattern* and *different pattern* are often applied in situations where a more concrete definition would sound cumbersome. As there are many aspects being considered in the process of analysis, the term pattern is often hiding a concrete finding. Using the expression '... it shows a different pattern' often pretends to have found a more general or globalised difference where in fact only one aspect was analyzed.

Because of the above reasoning I think the use of the term pattern should be limited to descriptions of theoretical ideas, and in the case of concrete study results more appropriate and concrete descriptions of the actual findings should be used.

Temporal patterns are probably best explained using the global language of music. On an abstract level it can be defined as a consecutive, organized arrangement of sounds. The melody consisting of tone level and rhythm can be considered as pattern homologous to the case of spatial phenomena. The speed at which a melody is played does not change the pattern. Interestingly the gamut used in singing a song does not alter its pattern, neither. Maybe we can find some causes of confusion in this context. Gamut and scale in the case of music can be used as synonyms. Transcribing a melody from C-major to F-major does not change the melody, but 'scaling' the tones by a factor for example 1.5 by altering the tone frequency from 400 Hz to 600 Hz, 460 Hz to 690 Hz and so on, will change the melody or even make it unrecognizable.

As it was illustrated above, there exists a variety of definitions, aspects and applications of the term pattern. In this thesis I will use the following definition: *Pattern is a general term for any recognizable regularity in the data.*

After looking at spatial and temporal aspects of patterns, we need to examine the often encountered term *spatio-temporal pattern*. As the term implies it describes a phenomenon which requires both spatial and temporal regularities to occur. A simple change in the spatial arrangement does not fulfill the requirement for a spatio-temporal pattern. A simple change in speed does not fulfill the requirements, neither. It is the combination of both space and time which needs to express regularities. Alterations in space which do not contain temporal regularities may be best named *spatial change*. If the regularities are changing, it may be expressed as a change in the specific spatial pattern.

An example for a spatio-temporal pattern can be found in car accidents with animals. They occur at different rates during the daytime, week and season and express 'preferences' for certain locations. Accidents with animals as red or roe deer for example mostly happen during dawn and dusk time. They vary according to the altitude of the sun changing during the year. These kinds of accidents occur more frequently during seasons where migrations take place. But then they often happen at locations different from the ones normally encountered during the year. In this case we have clear spatial and temporal regularities which in combination form a good example for a spatio-temporal pattern.

2.4 Questions Posed upon Point Objects

Whenever a wildlife study is being performed, the research goals can be classified into approximately seven questions. They reflect to some extent the evolution of the discipline.

- What is it? Due to a lack of binoculars, the first ornithologists were using a gun to be able to identify their 'observations', a common task in the high times of taxonomy.
- How many are there? This question is at the core of biologists. While most humans can simply enjoy watching animals, biologist can nothelp themselves start counting them.
- Where is it? After the two basic questions the focus is shifted to spatial aspects. The distribution of species became a large field of research.
- Why is it there? In the middle of the 20th century the whole complex of habitat analysis became popular.
- How does it influence others? Along with the developments of 'niche theory' and competition the whole ecological questions about how animals interact with their environment became a big research issue.
- Does it change? Due to the confrontation with progressing environmental disturbance and devastation the temporal aspects became more and more prominent in the questions of wildlife research.
- Where will it go?

Today's main focus in research and methodological developments lies in the last three questions. It is still very difficult to adequately characterize changes in wildlife data, and extrapolations into the future are even more difficult. The general topics mentioned in the above list need to be broken up into smaller questions to be operationalized. Table 2.4 provides a list of more detailed questions. It also includes a classification whether the questions include spatial and/or temporal aspects.

It is worth noting that almost all questions listed in table 2.4 include spatial aspects. About one third have temporal aspects. Of course the list is by far not complete, but the strong need for methods dealing with spatial and temporal aspects becomes evident.

In the following chapters I will try to give an overview of analysis methods that are available today. I will concentrate on methods dealing with point objects as the intention for this work stems from the work with wildlife data.

Table 2.4: List of example questions often posed in wildlife research when studying a certain species.

Spatial	Temporal	Question
		How many animals are there?
	t	Is there a change in the number of animals (increase, decrease)?
	t	How much does the number of animals fluctuate (σ, σ^2, stderr etc.)?
s		What is the density of animals?
s		Are there density differences between subareas?
s		Are there density clines (trends)?
s	t	Is there a change in density (measurements) over time?
s		Where are the animals?
s		How can I describe the distribution?
s		Does an animal always use the same paths?
s	t	Does it occur multiple times at the same place?
s		Which locations are important for an animal?
s		Which locations are used more intense than others?
s		Which locations are used less intense than others?
s		Where are corridors used by animals?
s	t	At which daytimes are certain areas used?
s	t	In which seasons are certain areas used?
s		Are the animals distributed in a randomly, aggregated (clustered) or evenly manner?
s	t	Is there a change in the distribution pattern over time?
s	t	Is there a positional shift over time?
s	t	Is there a dispersion change over time?
s	t	Is there a change in intensity of aggregation?
s		Are there minimal distances between points?
s		Is there a difference between 2 samples?
s		Is there a positional shift between the 2 (n) samples?
s		Is there a difference in dispersion between the 2 (n) samples?
s		Is there a difference in clustering between the 2 (n) samples?
s	t	Is there an influence of the moon or sun on the spatial distribution of the animals?
s		Is the distribution random according to other objects (points, lines, polygonal features)?
s	t	Are there distinct phases in which there is little change, which can be defined for further analysis?
s		What are the environmental parameters influencing (determining) the distribution of the animals?
s		What is the influence of a specific environmental parameter on the distribution of the animals?
s		Are there environmental parameter clines which could explain clines in the animal data?
s		What habitat is available to an animal?
s		How big is the area an animal is using (homerange)?
s	t	How does the size of a homerange change over time?
s		Are there barriers in the dispersion of animals?
s		What is the mean distance between objects?

3

Current Analysis Methods for Point Objects

3.1 Point Pattern Analysis in Biology

Compared to the literature on point pattern analysis in GIS (see below), a larger variety of methods is used in the biological and especially in the ecological literature. Several simple descriptive measurements have been used to describe aspects of point distributions. The easiest of them is probably the **mean center**, which is easily calculated as:

$$\bar{X} = \frac{\sum X_i}{N}, \bar{Y} = \frac{\sum Y_i}{N} \tag{3.1}$$

N = number of points
$\bar{X}, \bar{Y}$ = mean of X and Y

Since the mean center as calculated above is heavily influenced by outliers in the data, the median center is often considered to be a better method to describe a central locality.

The **median center** is described in two different ways. Cole and King (1968) and Hammond and McCullagh (1978) define the median center analogously to the one dimensional case as the intersection point for the separate medians in the x and y directions. This definition has the problem that the median center depends on the layout of the coordinate system. Rotating the coordinate system results in a different location for the median center. In contrast, Neft (1966), King (1962) and Smith (1975) define it as the point, where the sum of distances to all points is at a minimum (minimum aggregate travel). The point is calculated in an iterative process using successively finer grained grids.

Among the most commonly used easier measurements of spread in point data is the **standard distance**. The standard distance used in geostatistics is analogous to the standard deviation in simple, descriptive statistics. It can be calculated as:

$$SD = \sqrt{\frac{\sum(X_i - \bar{X})^2}{N} + \frac{\sum(Y_i - \bar{Y})^2}{N}} \tag{3.2}$$

which can be rewritten for faster computation as:

$$SD = \sqrt{(\frac{\sum(X_i{}^2)}{N} - \bar{X}^2) + (\frac{\sum(Y_i{}^2)}{N} - \bar{Y}^2)} \tag{3.3}$$

SD = standard distance
N = number of points
$\bar{X}, \bar{Y}$ = mean of X and Y

Since the standard distance uses squared distances, it exaggerates the importance of extreme points. In a more spatially explicit way a measurement of spread of points can be achieved by calculating the **standard deviational ellipse**.

Traditionally, quadrat-density analysis and nearest neighbor methods have been widely used (Legrendre and Fortin, 1989; Carpenter and Chaney, 1983) in biology to assess the deviation from complete spatial randomness.

There are two ways to check for random patterns by quadrat counts using the Poisson distribution as a reference:

- compare the actual quadrat counts to the expected distribution
- use the variance-mean ratio of the observed distribution

To compare the observed quadrat counts to a Poisson distribution expected under the hypothesis of complete spatial randomness, a χ^2 statistic can be calculated as follows:

$$\chi^2 = \sum \frac{(X_i - \bar{X})^2}{\bar{X}} \tag{3.4}$$

X_i = number of points in each quadrat
$\bar{X}$ = mean number of points per quadrat

with $N - 1$ degrees of freedom.

The Poisson distribution has a variance equal to the mean, thus departures from unity in the ratio reflect tendencies towards either clustering or regularity. The degree of departure from 1 can be converted to a z score after calculating the standard error of the difference (SE_x) from:

$$SE_x = \sqrt{\frac{2}{N-1}} \tag{3.5}$$

$$z = \frac{observed\ ratio - expected\ ratio}{SE_x} \tag{3.6}$$

N = total number of quadrats

Several indices for quadrat count data have been developed. An overview can be found in table 3.1.

These quadrat count methods have been enhanced for overcoming their dependency of scale by selecting a certain quadrat size. Greig-Smith (1952) basically proposed a resampling to successively coarser grid sizes using χ^2 statistics. Mead (1974) uses another approach using successively coarser blocks of 4 x 4 subblocks and testing the set of counts against a random selection from $(16!)/(4!)^5 = 2627265$ possibilities, as implied by complete spatial randomness.

Apart from the methods based on quadrat counts, a second family of analysis methods is widely used in the study of point patterns. One of the earliest methods available is the **nearest neighbor distance**.

Table 3.1: Indices for Quadrat Count Data (after Cressie (1993)).

Name	Index	Reference
Relative variance	I	Fisher et al. (1922)
David-Moore index	ICS	David and Moore (1954)
Index of Cluster Frequency	ICF	Douglas (1975)
Mean Crowding	$\dot{\mathrm{X}}$	Lloyd (1967)
Index of Patchiness	IP	Lloyd (1967)
Morisita's Index	I_δ	Morisita (1959)
(Xingping)	L_a	Xingping (1996)

Evans and Evans (1954) developed a measurement index and linked it to the Poisson probability distribution. The analysis compares an observed spacing of a point distribution to an expected random pattern. The average expected distances for a random pattern are calculated as:

$$\bar{r}_e = 0.5 * \sqrt{\frac{A}{N}} \tag{3.7}$$

$\bar{r}_e$ = average expected distance
A = area of study region
N = number of points

The above formula is intended for use with a study area without border effects. If the study area has border effects, the above formula underestimates the expected average nearest neighbor distance for a random pattern. As this is a common source of error, I will shortly illustrate this effect in the following paragraphs.

As an example the average nearest neighbor distance from 1000 random simulations is 158 meters for the badger setts data from Good (1997) for the Sihlwald area, whereas the above formula results in 144 meters. Considering only the 35 large setts (with more than 3 entrances) the corresponding values are 314 and 270 meters respectively.

There are two approaches to cope with these border effects. One can construct an outer edge of the study area where points lying inside this edge area are only used for distance calculations for points lying in the inner (kernel) area, but are not taken as observations. An other approach is to perform Monte Carlo simulations by generating random points within the study area and then calculating the nearest neighbor distances for each simulation. The advantage of this method is that one can use all observations in the calculations. The disadvantage is the large amount of computing time needed. Since biological data analyzed with nearest neighbor methods are often based on scarce samples, it is desirable to use all data available.

The effect of shape of the study area on the average nearest neighbor distance in a random distribution is illustrated in figure 3.1. The six shapes all have an area of 16 hectares. In each shape the mean nearest neighbor distance was approximated by simulations for 20, 50, 100 and 200 random points. Two results can be deduced from figure 3.1: (1) The mean nearest neighbor distance for random samples increases with an increase of the border effects. (2) This effect is stronger for smaller sample sizes.

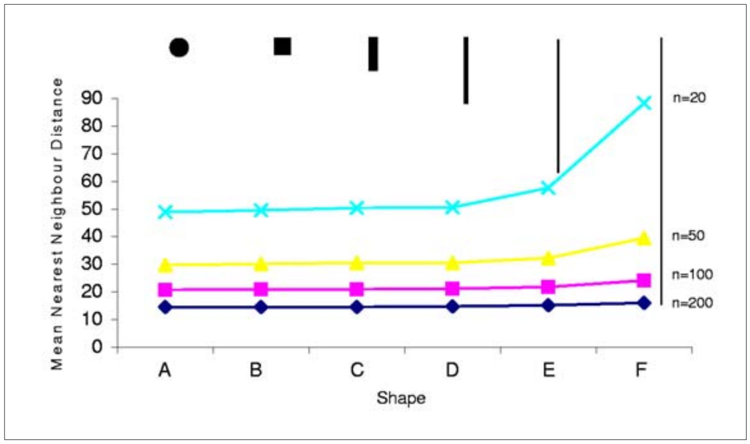


Figure 3.1: Effect of shape of the study area on the expected average nearest neighbor distance. Calculations made from simulations of random patterns.

The expected distance for a (maximally) dispersed pattern is given by:

$$\bar{r}_{dis} = \frac{\sqrt{2}}{3^{\frac{1}{4}}\sqrt{\frac{N}{A}}} = \frac{1.07453}{\sqrt{\frac{N}{A}}} \tag{3.8}$$

$\bar{r}_{dis}$ = expected distance
A = area of study region
N = number of points

To test the significance of patterns, Getis (1964) used the standard error of the expected average nearest neighbor distance to calculate a z value. The average expected distances for nth-order neighbor distances can be calculated as:

$$\bar{r}_{e_n} = \frac{1}{\sqrt{\frac{N}{A}\frac{(2n)!n}{(2^n n!)^2}}} \tag{3.9}$$

$\bar{r}_{e_n}$ = average expected distance to the nth neighbor
n = order of nearest neighbor
N = number of points
A = area of study region

These methods using distances have been enhanced for cases where measurements are taken from observation points to the events of interest (Diggle, 1983; Doguwa and Upton, 1989).

With the exception of the order neighbor distances, the above methods are also of limited use due to their inability to work over a range of scales. This difficulty has been addressed by several authors and resulted in the development of second-order functions (e.g., K-, L- and G-functions), which became more and more popular (Moutka and Penttinen, 1994; Getis and Franklin, 1987; Tomppo, 1986; Ripley, 1988; Getis, 1984). For these functions some methods for edge effect corrections have been published (e.g., Doguwa and Upton, 1989; Haase, 1995).

As another area of interest, different methods for density estimation are available. The earliest methods were simply counting the number of quadrats in which observations were found. Today a variety of methods are in use ranging from Dirichlet tessellations, least diagonal neighbor tessellations, weighted triangles and weighted polygons (Upton and B.Fingleton, 1985) to complex kernel estimations, harmonic mean methods and Bayesian smoothing (Worton, 1989; Dixon and Chapman, 1980). These methods are being widely used in wildlife research to gain more information from the observations than is possible using traditional homerange analysis methods such as the convex polygon.

Perry and Hewitt (1991) introduced a new class of tests for spatial patterns referred to as SADIE (Spatial Analysis by Distance Indices). They compare the spatial arrangement of the observed sample with other arrangements derived from it, such as those where the individuals are crowded or as regularly spaced as possible. Perry (1995) extends the method for two-dimensional patterns using Voronoi tessellations which are iteratively transformed into a regular pattern.

Area estimation for example in homerange analysis has also received special attention especially in wildlife research, and over a dozen methods have been used (e.g., Dixon and Chapman, 1980; Samuel and Garton, 1985; Worton, 1989; White and Garrot, 1990).

In recent years, surface pattern analysis methods such as Moran's I, Geary's c, correlograms, (semi-) variograms and two-dimensional spectral analysis appeared frequently in the biological literature (Renshaw and ED., 1984; Legrendre and Fortin, 1989). These methods are often used with aggregated data representing surfaces rather than point patterns.

Andersen (1992) used Ripley's K-function (Ripley, 1976) for interaction models using 'static' point pattern data. He expresses the need for statistical methods that explicitly allow for both spatial and temporal structure. Knox's method (Knox, 1964) for analyzing space-time interactions using 2-way tables is problematic because of the assumed independence of events (O'Kelly, 1994). O'Kelly (1994) presents a new approach using clustering techniques which allow for interdependence between the clusters.

Biondini et al. (1988) presented a new technique for analyzing multivariate patterns using permutation procedures. It has been applied to zoological point data (White and Garrot, 1990) with promising results for multivariate point patterns. In the zoological literature, habitat analysis is often done using Chi-squared analysis described in Neu et al. (1974) and its various modifications (see White and Garrot, 1990). Other statistical methods such as compositional analysis and log-linear models have also been reported (Aebischer and Robertson, 1993; Heisey, 1985). One of the major problems with these methods is the condition for independence of the observations (Swihart and Slade, 1985; Dunn and Gipson, 1977). Time aspects are only included by separately analyzing different time spans (e.g., Stoms et al., 1993).

3.2 Point Pattern Analysis in Commercial GIS

A common body of knowledge of point pattern analysis exists in the geographical, forestry and other sciences. Nevertheless, structural point pattern analysis is quite limited in geographic information systems. Present commercial GIS use elaborate techniques for spatial operations like buffering objects or overlaying different thematic layers. Concerning statistical spatial analysis of point objects (not to be confounded with point measurements), their capabilities are limited to simple descriptive measures such as minimum, maximum, mean and standard deviation. Some raster-based systems (e.g., IDRISI) offer more complex statistical analysis functions (e.g. measures of spatial autocorrelation), but they do not offer sophisticated algorithms for point pattern analysis.

Several authors (Openshaw, 1991; Goodchild et al., 1991) discuss possible ways to link GIS with statistical spatial analysis. Five strategies may be distinguished:

1. free standing spatial analysis systems
2. integration of basic GIS functionality to statistical software
3. 'loose coupling' of proprietary GIS to statistical software
4. 'close coupling' of GIS and statistical software
5. complete integration of statistical spatial analysis in GIS

The second strategy is applied in the experimental system SPLANCS by Rowlingson and Diggle (1993). They made some enhancements to the S-Plus system to produce a tool for display and analysis of spatial point pattern data. G-, F- and K-functions as well as a kernel smoothing procedure were implemented. The advantage of having the full statistical capabilities of S-Plus available is achieved at the expense of having no real GIS functionality. This approach, although criticized by Openshaw (1991), has also been adopted by Griffith (in Haining and Wise, 1991) and SAS Inc. producing a module SAS/GIS.

Several developments try to couple GIS with commercially available statistical packages such as GLIM, SAS, SPSS, generally using ASCII exports. They mostly concentrate on measures of spatial autocorrelation and association (Gatrell and Rowlingson, 1994). Others implemented methods for point pattern analysis concentrating on first- and second-order analysis (Gatrell and Rowlingson, 1994; MacLennan, 1991; Rowlingson and Diggle, 1993); they occasionally include some density estimation techniques (e.g., kernel density estimation). MacLennan (1991) implemented second-order analysis methods (G- and L-functions) into the GRASS system (GRASS, 1993). Analysis of time aspects has rarely been considered, partly because of the lack of infrastructure for temporal geographic information systems and the absence of established analysis methods.

Openshaw (1994) presents a new analysis approach by extending his earlier work to 'geographical analysis machines' GAM (Openshaw, 1987). He argues that traditional exploratory methods of pattern discovery are not feasible in a multivariate GIS environment with tri-space (geographical, temporal and attribute) information. He extends his GAM to 'space-time-attribute-machines' (STAM) and '-creatures' (STAC), using artificial life (see Beer, 1990; Langton, 1989) to search all three spaces (Openshaw and Perrée, 1996). Basically, a STAM is an automatic screening program which is searching all the locations in geography, time and attribute space for evidence of clustering.

Most of the above mentioned methods are intended for use with static data. This is not very astonishing when considering that their background is strongly connected to botanical data. Spatial movements are seldom of interest in this case. In the few examples where time is included explicitly in the methods

(Openshaw and Perrée, 1996; Cressie, 1993) the meaning of patterns to be searched for is reduced to comparing the clustering of points to complete spatio-temporal randomness.

The only approach available for handling time with spatial data from animal locations is the aggregation of observations into time slices [1], followed by the application of static analysis methods. Secondly almost all of these methods including the point process theory (for an introduction see Diggle (1983) or Cressie (1993)) are only of marginal interest here, as they are not dealing with mobile objects.

3.3 Analysis Methods for Point Objects Available in Current GIS

One thing must be pointed out here. The reader should not confound *point objects* with *point measurements*. Point objects are objects with no extent or of an extent that can be neglected. They may have attributes or not, but the main information is their location. Contrarily point measurements are measurements of fields at specific locations, so they are basically a representation of a field. The location itself is not so much of interest as the measurement itself. For this kind of data GIS provide a good collection of analytical methods. Variograms and kriging are two examples which interestingly were stimulating the idea for creating RDFs introduced in chapter 7.

At a first glance it is astonishing that only few analysis methods or operations are available in current commercial GIS for point objects. Until recently only three methods were available apart from the simple selection according to attribute data. The first methods is the intersection with polygons or fields. This enabled the user to identify in which polygon the point object is located or what value a continuous surface has at that location. The second method point objects can be used in is buffering. Point objects can be buffered to produce circular polygons around them. The third method available is the calculation of distances to other objects, whether they are also points, linear elements or polygon borders. In the last two years the situation was improved a little bit by implementations of simple density calculations of point objects (e.g. kernel density estimations).

To some extent this situation may be explained by the fact that one of the mainapplications of GIS today is in the area of administration and facility management. In these applications point objects exclusively appear as static, non-moving objects, for which the mapping itself is the main reason for its integration into a GIS. Although there is quite some activity in the scientific literature, the customer segment of researchers seems to be unattractively small for commercial GIS producers.

It may be argued that most GIS provide powerful tools for developing customized applications. To some extent this is true. But there are two difficulties with this approach. First and most important is that the dissemination of such 'user' developed methods is very difficult and often not of interest to the user due to financial or competitional reasons. The second difficulty in relying on

[1] Serial autocorrelation tests used in radiotelemetry are only used for data validation. Hence they are not considered here as analytical method.

'user' developed methods is the speed and integration level at which it can be done. Especially in research computing speed is often very important, stimulating isolated applications.

3.4 Database Aspects of the Analysis of Temporal Data

In the past few years the classical distinction in GIS between geometry - handled by specialized GIS modules - and attributes - treated in small integrated databases or exported to large external databases - has started to fade out. Several vendors of database products recently introduced spatial extensions for their systems, as for example the *Spatial Data Option* (SDO) from Oracle, the *Spatial Datablade* for Informix or the spatial data support in OpenIngres. They are mostly extensions to support spatial data types such as points, lines and polygons, but compared to geographic information systems, they provide little support for spatial analysis. A different approach was taken by ESRI, the producer of Arc/Info, by developing a spatial platform on top of large commercial database systems, called Spatial Database Engine (SDE), which allows to store both geometry and attributes in the database in a transparent way. Nevertheless both DBMS and GIS vendors focus on the 'snapshot oriented' systems concerning temporal aspects.

3.4.1 Temporal Databases

So far hundreds articles related to temporal databases have been published (Pissinou et al., 1994). The majority of work has been based on relational data models. Only very few approaches relate to different data models such as the object (e.g., Alfarmanesh et al., 1985) or entity-relationship data models (e.g., Studer, 1986). Langran (1992b) considered only relational databases as sufficiently mature for operational use. In database applications such as geographic information systems, Pissinou et al. (1994) considered the relational data model to be insufficient to represent the complex data structures needed for temporal extensions of these systems. In spite of the lively activity in temporal database research, no widely used temporal database management system is commercially available. In most prototypes (e.g., Ahn and Snodgrass, 1986) and implementations for specific applications, the interpretation of time was done by user application programs rather than by the database management system (DBMS) itself (Pissinou et al., 1994).

3.4.2 Temporal Query Languages

Peuquet (1994) mentions that the development of temporal query languages is a relatively new area that has not received much attention from research and development until recently. Nevertheless over 20 query languages have been proposed. Most developments are a result of extending existing query languages such as SQL (Structured Query Language) or Quel, a query language for the INGRES relational DBMS, to the spatio-temporal domain. Tansel et al. (1993) gives a survey of much of the work performed.

TOSQL developed by Ariav (1986) is an extension of SQL. It allows access to both current data and their previous versions, but does not include update semantics. Snodgrass (1987) extended Quel to TQuel with language constructs to retrieve facts that have been time stamped with a validity interval. HTQUEL (Homogeneous Temporal Quel) proposed by Gadia (1988) is based on the same query language as TQuel. Time is considered as discrete equidistant time intervals. Kim et al. (1990) developed ETQL as a front-end system to INGRES. ETQL supports abstract time (including relative time, e.g., last spring) to achieve easier specification of time in queries. In 1994 the specification of TSQL2 was published (Snodgrass et al., 1994) which later resulted in proposals for an extension to SQL3 called SQL3/temporal.

Standard SQL does not include time support except for user-defined time. Neither transaction time nor valid time is available (see Table 2.2). Date and time support in SQL-92 are similar to that in DB2 (Melton and Simon, 1993). The design for SQL3 only corrected some of the inconsistencies, but contains no additional temporal support over SQL-92 (Pissinou et al., 1994). However, until 1994 the SQL3 proposals included several constructs that can be useful for temporal extensions (e.g., interval data type). Since then several change proposals were submitted to the ANSI and ISO SQL3 standards committees for adding a new part termed SQL/Temporal (Snodgrass et al., 1996b,a; Snodgrass, 1997).

None of these temporal query languages are built for complex spatial queries. Oracle recently announced their support for the HHCODE standard for spatial attributes. Such efforts could induce further developments in the research of temporal query languages including spatial aspects.

Commercial database systems often support transaction time mechanism based on tuple time-stamping. Valid time is not supported as a built-in functionality.

3.4.3 Temporal GIS

The enormous advances in computer technology in recent years made it possible and necessary to give consideration to temporal geographic information systems. Several conceptual frameworks have been proposed, and a few partial implementations have been reported, but it will take a big step until off-the-shelf temporal GIS will be available. Despite these initial advances, the technical and conceptual difficulties still require a large amount of attention (Langran, 1993; Peuquet, 1994; Ramachandran et al., 1994).

Langran (1988) examined a concept called dimensional dominance, where access to the data is classified as either predominantly spatial or predominantly temporal to optimize data and algorithms. She examines four representational models for spatio-temporality based on existing non-temporal data models (Langran, 1992b): (1) Space-time cube, (2) Sequential snapshots, (3) Base state with amendments and (4) Space-time composite.

Except for the first representation, there are no implicit temporal relations between objects and states involved. 2) and 3) contain time as separate data layers, whereas 4) handles the temporal aspect separately in the non-spatial attribute database (Langran, 1992b; Kienast et al., 1991).

Peuquet (1994) proposes an integrated approach called triad representational framework, which is an extension of the dual representational framework described earlier (Peuquet, 1988). It integrates temporal, locational and object-related semantic aspects of the data using spatial learning and knowledge-based scene interpretation.

The efforts to build prototype temporal GIS were taken to satisfy specific needs and therefore were implemented only with partial support of temporal aspects. Beller et al. (1991) developed a prototype temporal GIS as a proof of concept to carry out global change research. The central concept is the temporal map set object (TMS), which is a collection of GIS maps representing the same area and theme at different times. The system is capable of performing interpolation between time slices and can incorporate events as separate binary TMSs. Langran (1993) describes implementation issues to consider when building temporal GIS, citing specific problems from automated hydrographic and aeronautical charting systems (Langran, 1990) and a forest GIS (Langran, 1992a). Main aspects seem to be representation, incremental updates, temporal generalization and longevity. Incremental updates produce artifacts and edge effects which can cause data to be inconsistent. The Environmental System Research Institute, Inc. (ESRI) has developed a temporal GIS application for a private forestry management firm (L. Montgomery, pers. comm.). The implementation is based on an area event system which maintains valid times and allows historical queries.

The application of GIScience in the domain of changing phenomena seems to have a promising future. Wachowicz and Healey (1994) stated that "by producing a lineage of data to track the historical information associated with real-world phenomena, temporal GIS will provide analytical tools for the recognition of patterns of change through time as well as the prediction of future changes, by implementing dynamic simulations". Today this is still more likely to be a wish than reality. In contrast to the numerous prototypes of temporal database systems most current GIS products remain snapshot-oriented systems capable only of static representations of data (Bagg and Ryan, 1997). In contrast to the database research the need for adequate representation of temporal data becomes an urgent need in GIS, because large databases are being built and updated for monitoring the continuously changing environment. In these databases spatial and temporal inconsistencies are not allowed, because today's analysis methodology is not able to cope with inconsistent data. Today's temporal GIS research is mostly concentrating on data representation and query (Güting et al., 1998; Worboys, 1994b; Peuquet and Wentz, 1994; Ramachandran et al., 1994; Yearsley and Worboys, 1995; Kemp and Kowalczyk, 1994), but the analysis part is not receiving much attention. Güting et al. (1998) provide a semantic foundation for handling time dependent geometries. They are the first authors considering moving objects (e.g. airplanes) in their work for founding their framework. This seems to be a major step (even though still on a theoretical level) to overcome the event centered view used by most other researchers. Most prototype applications are using time only as an additional selection criterion, but neglect the possibility of completely new analysis methods for temporal phenomena.

In the early 90ies, a frame free geographical representation was postulated as a research aim in GIScience (Tobler, 1989; Kemp, 1993). Although this has not been realized yet, it has become a prominent issue in handling temporal geodata.

II

Development of Methods

Conclusions from Chapters 1 to 3: Where is Research Needed?

This is a short but nevertheless important part. I shall recapitulate and integrate the findings in the previous chapters to point out in which directions the research that follows should be targeted.

The following conclusions can be drawn:

1. A system similar to the (spatial) coordinate transformation between different coordinate systems needs to be developed for the temporal domain (section 2.2.1).

2. No concepts for the representation of moving objects are available in (commercial) GIS.

3. Major efforts in GIS concerning temporal data are focussed on representational issues. The analysis part has not received any attention yet.

4. There are no effective ways to visualize, handle and analyze moving objects within GIS.

5. Apart from cartographic symbology and video-like animations no true analytical methodology exists for spatio-temporal data with the exception of the cluster search engines developed by Openshaw and Perrée (1996).

6. Temporal aspects are reduced to a minimum in today's analysis (at best to time slices).

7. There is a major disparity between wildlife researcher needs, GIS, and the analytical methods available.

8. Time slicing of wildlife data has mostly been done using some artificial human (or civilization) centered calendars (months, weeks, hours etc.). New methods need to be developed to identify and define biologically meaningful and homogeneous phases within data.

9. Up to now time has only been considered as a linear component. Other temporal aspects such as cyclic phenomena or rhythms have not been recognized as important aspects in temporal GIS research.

10. No or only rudimentary concepts are available on how to handle, integrate and analyze the surrounding environment of an observation in a spatially explicit way, neither in biology nor in GIScience.

11. The temporal sampling scheme should be changed from 'avoidance of temporal autocorrelation' to a scheme which describes movements as accurately as possible (section 2.1).

12. Since an analysis is much more complex when conducted in a spatially explicit manner, the analysis might need to be performed step by step for one aspect after another. As it is well known from cartography the accumulation of large amounts of information can produce maps that are hiding more than they are revealing.

With the exception of the first two items mentioned above, the next three chapters are going to provide new approaches to these important issues, equally important for both Biology and GIScience. Thus the main question is the following:

How can we include temporal aspects in the analysis of the spatial behavior of an animal within a GIS environment?

4

The Concept of Temporal Data Frames in GIS

4.1 Introduction to the Problem

In today's wildlife studies large investments are made in data collection. They require a lot of human and/or technical resources over long observation periods. Large volumes of data are gathered which are expressions of complex behaviors and environments. These datasets require powerful and fast mechanisms to be analyzed accurately in respect of their temporal components. Today's standard methods to analyze such data are static with very few exceptions. There is an enormous pressure in these studies to produce results in a very short time. This means that the data analysis is restricted by the available analytical methods, which can be easily applied to the data. Researchers apply these methods intended for use with static data, neglecting most temporal aspects inherent to the data. I would like to hypothesize that aside from an often incorrect application of such methods (e.g., James and McCulloch, 1990; Seaman and Jaeger, 1990) the neglect of time and its various aspects in such studies is a major source of error and hides much of the information and insights that could be available from the data.

A second reason why data is analyzed by static methods is that often the focus is set towards the complex habitat requirements of an animal. Even though a certain habitat or more precisely its definition may look very static and is thus considered to be constant, the environment an animal is living in is always changing. Analyzing dynamic animals with respect to a dynamic environment is still a very complex matter, normally too complex for ordinary studies. As I have shown above there is a lack of accurate methods for such data with distinct temporal properties.

In table 4.1 a classification of study types with respect to the perception of animals and habitat as either static or dynamic components in the analysis is shown. Most studies use animal observations as static, locationally fixed point objects. Habitat is seen as the static type of vegetation composition at a certain location. With these two views a study can be classified as type A study in table 4.1. It is clear that none of the two assumptions is correct. Animals are moving, and the habitat changes over time. Hence further research will be

Table 4.1: Classification scheme of habitat analysis studies with respect to the perception of animals and habitat. Both animals and habitat can be seen as static or dynamic data. See text.

		Habitat perception	
		static	*dynamic*
Animal perception	*static*	A	B
	dynamic	C	D

needed to achieve studies of class D, where animals are considered as moving and habitat is taken as a changing component. This will need substantial effort and will probably result in intermediate methods of type B and/or C, where one of either animal or habitat remains as a static component in the analysis.

Such methods for adequately analyzing temporal data in animal movements still need to be developed, defined and standardized. Standards are needed for acceptance and wider application of the methods.

4.2 Purpose

The aim of this chapter is to take the first of three steps towards a more natural and accurate handling of animal data in spatial analysis. It is a conceptual shift from statics to dynamics and will set the basis for the following chapters. It is a relatively technical but easy to read chapter which will introduce powerful instruments for the visualization and fast analysis of complex spatio-temporal patterns.

At a first glance temporal aspects might seem to be relatively simple to handle. But going a bit more into detail they become quite complex. In section 2.2 some of these problems were mentioned. Calculations of the beginning and end of dawn, for example, depend on the season, latitude and longitude and require a lot of geographical, astronomical and temporal computations. In the case of the tides calculations become even more complex. The complexity increases even more when combinations of such aspects are considered.

Shepherd (1995) provides a recent classification scheme for dynamic visualizations of geographical data. He distinguishes the following 6 classes according to the source of variation (table 4.2). I will introduce the *Temporal Data Frames* (TDF) concept as an extension to the first category of converting real time to display time. It is needed as a basic concept in the next chapter. It will provide fast methods for accessing large volumes of data.

In this chapter I shall try to elaborate a methodological framework for the following objectives:

- **Refinement of the temporal perception of biological data.** As I stated in the introduction temporal aspects of spatial phenomena have not received enough attention in the past. Time is handled in a very coarse way which probably hides a lot of information inherent to the data. Finer methods should be developed to analyze the data to their full extent.

Table 4.2: Classification scheme for dynamic visualizations from Shepherd (1995)

Class	Example/Explanation
Data	converting 'real time' to 'display time'
Representation	changing symbolism
Observer	moving observer
Agents	e.g. particles flowing
Entities	e.g. objects trigger symbol change
Designers	e.g. esthetic symbols

- **Integration of different time aspects.** Normally researchers in wildlife biology only include seasons as a temporal dimension in their analysis. As I have shown in section 2.2 there are several temporal aspects that are interlinked in complex manners. It is the aim to make these aspects accessible to the research community and provide the necessary methods for it.

- **Recognition of changes.** In the area of fast changes in the environment it is important to be able to detect such changes in a very efficient way. The following section shall develop mechanisms for the detection of changes in relation to different temporal aspects at a very fine level.

- **Recognition of biologically reasonable phases.** Preventing the application of 'unnatural' phase definitions in the analysis of biological data, methods should be developed to enable the researcher to perceive and define biologically reasonable phases in the data. It should provide the means to move away from the human-oriented definitions based on cultural calendars.

4.3 Introduction of TDF: the Basic Form

I shall use data on two badgers[1] in the Sihlwald area near Zurich (Switzerland) as an introductory example. It shall make the understanding easier.

Tracking an animal by means of direct observations or radio-telemetry gives at any instant a certain location of the animal. To summarize the whereabouts in a certain time period, for example within two hours, we can either try to keep it in our memory or writing all locations down on paper. Then we select the appropriate observations and draw them on a map. This is quite a time consuming procedure, but sometimes we are interested enough to repeat that procedure for another time period. This procedure is continued for at most a few dozens of times. Computer programs such as spreadsheets or databases help speeding it up a bit, but it still remains a lot of work. Although this may appear to be a very simple task, we have to acknowledge that this is exactly what wildlife researchers are doing with spatially referenced observations in 95 percent of the cases!

[1] Data courtesy of Karin Hindenlang

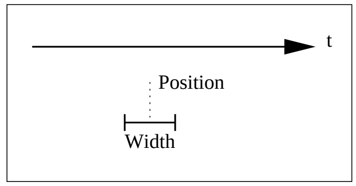


Figure 4.1: Basic form of a temporal data frame. Data are selected by a frame defined on the continuous time axis by specifying the position and width of the frame (period).

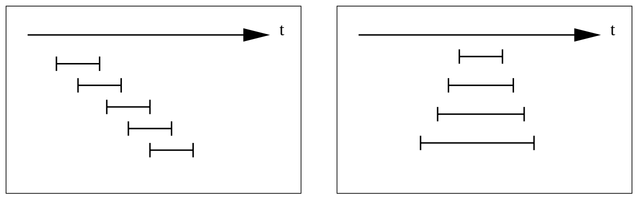


Figure 4.2: Methods of how to change a temporal data frame. Left: Change of position. Right: Change of width of the frame.

So what are they exactly doing? They define an interval (e.g. 2 hours) and select all observations at a certain instant within that interval. Then they move to another instant (normally shifted by the amount of time equal to the observation interval) and then use the resulting observations. A classic example are the seasonal homeranges often used in wildlife studies.

The procedures used hereby can be generalized to the concept of *temporal data frames*. It is a selection mechanism providing access to the data using the temporal domain. The general concept is illustrated in figure 4.1. The so called *temporal data frame* is defined by its position (e.g. 1.6.1998 16:00:00) and width (e.g. ± 1 hour). It can be used to select data within a defined period of time.

This now provides two parameters which can be changed dynamically: (1) the position, which means that one can drive through time selecting the data which are contained within the temporal data frame (figure 4.2 left). (2) The width of the frame. By changing the width of the data frame one can add observations which were just outside the edge of the frame (figure 4.2 right).

This is the easiest form of a temporal data frame (TDF) where the continuous time is used as the temporal aspect. Starting from here I will now continue with more complex temporal aspects in the next section.

4.4 Cyclic Aspects of Time

The alternation of day and night is important for most animals. It influences the behavior and use of their habitat. At first it looks straight forward to implement temporal data frames similar to the basic case above for this temporal aspect using a position and width parameter (figure 4.3). As an example it could be the aim to select all observations which occurred at dawn. At a second glance

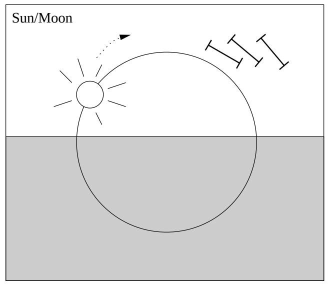


Figure 4.3: Illustration of temporal data frames. A temporal data frame can be used in a cyclic time aspect as for example the daily sun movements (geocentric) or moon movements. It is defined as in the basic form with a position and a width parameter. It selects all data e.g. within dusk, which basically means that multiple data frames are set up at intervals covering all dusk times over the whole observation period.

it is more complex than expected. There are three reasons for this:

1. The single temporal data frame is replaced by a large number of data frames. For every day a frame has to be constructed which selects the data within dawn at that day. This means a shift from a single TDF to multiple TDFs.

2. Sunset, sunrise and dawn times and all their related aspects are variable throughout the year and change with the position of the earth on its orbit. Figure 4.4 illustrates this fact. In the northern hemisphere the nights are longest in December and shortest in June. In the southern hemisphere this is reversed.

3. Dawn duration as one example depends on the latitude. This is illustrated in figure 4.5. There are two things worth noticing here. First the duration of dawn is longer the further we are from the equator. Second there are two (!) periods with long dawn durations in the year. In the northern hemisphere they are in December and June, while the shortest dawn durations occur in March and September.

This amount of complexity requires that the necessary calculations are automated and become an inherent part of a temporal data frame. These include coordinate system transformations (e.g. Swiss coordinate system to latitude/longitude), time transformations (UTC to JD) and calculations of the angle of the sun above or below the horizon.

The dawn was chosen as an easily understandable example. It is defined in three versions as the civil, nautical and astronomical twilight. They are defined as the time that starts when the sun is 9°, 12° and 18° below the horizon and ends at sunrise. It should be clear that the temporal data frames are not limited to these figures and can be applied at continuous ranges of values. The dawn is only one specific setting of values for a TDF.

I used the examples of sunrise, sunset and dawn above. There are other temporal aspects that are relevant in this context. Aside from solar there are

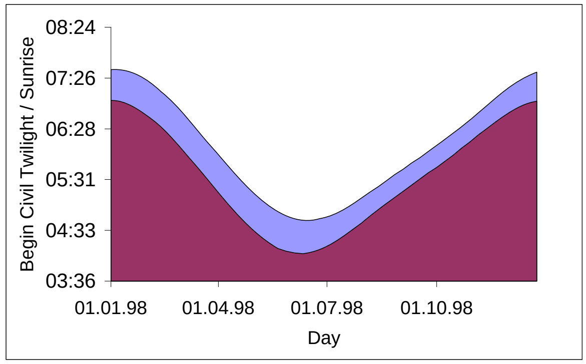


Figure 4.4: Changes of the sunrise (upper line) and the beginning of the civil twilight (lower line) during the year at a latitude of 40° north.

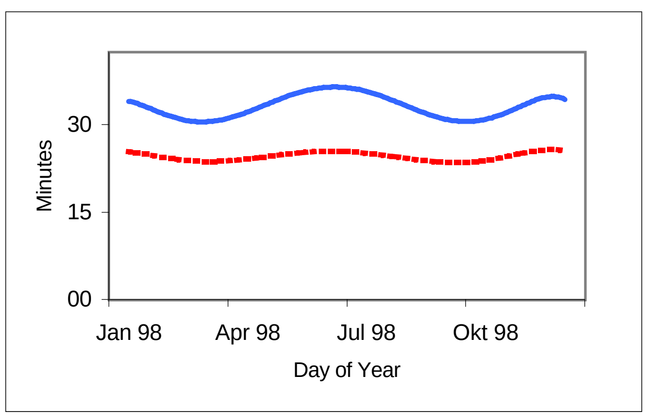


Figure 4.5: Changes in the duration of the twilight during the year. Upper line (blue): latitude: 40°. Lower line (red): latitude: 0°. Twilight in the northern hemisphere is shortest in spring and summer. The duration of the twilight is also shorter closer to the equator.

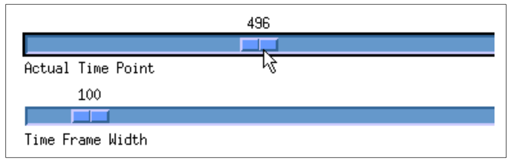


Figure 4.6: Implementation example of scrollbars as one method to define the parameters for a TDF.

lunar aspects that need to be considered. The lunar altitude, azimuth and its illumination are cyclic phenomena often standing on the wish list of a wildlife researcher for the analysis of his or her data. In most cases it remains a wish due to missing methods for handling these aspects. By introducing the concept of *temporal data frames* as a new method in the field of exploratory data analysis, it becomes possible to gain insights into animals' responses to such phenomena connected to solar and lunar and possibly other rhythms.

4.5 Forms of Implementation

It is always difficult to describe dynamic methods with static communication media like the paper this text is written on. I hope this limitation can be overcome to some extent by looking at the implementation details of the prototype application **TUPF** presented in Appendix A. The real effectiveness can only be seen and estimated by looking at full implementations on a computer system.

In the following I shall present some implementation topics for TDFs.

Without going into much detail I think it is useful to discuss some implementation issues at this point. Greater detail about the prototype implementation can be found in Appendix A.

As stated in the introduction fast methods are required to access the large data volumes commonly occurring in wildlife applications. Building a user interface (UI) for this purpose means that graphical methods are needed and used. Future UIs may be based on virtual reality applications or further developments, but for the time being they are not available to most researchers.

For the basic form of the TDFs two parameters need to be interactively changed: the position and the width. In a computer program two scrollbars can be used for this purpose, one that selects the position and one which defines the width (figure 4.6).

A more sophisticated solution would be the application of a diagram with two axes. The x-axis describes the position, the y-axis the frame width. The user can then use a pointer (e.g. a computer mouse) moving around in this space and simultaneously change both parameters as needed.

For cyclic temporal aspects the form of using two scrollbars can also be used, one for the selection of the time point, the other to determine the width of the time frame. In addition there can be cases where more sophisticated access methods are needed. One example for this is illustrated in figure 4.7. This way the problem of cyclic and circular parameters does not occur as in the application of scrollbars.

These two examples of implementation issues have shed some light on how

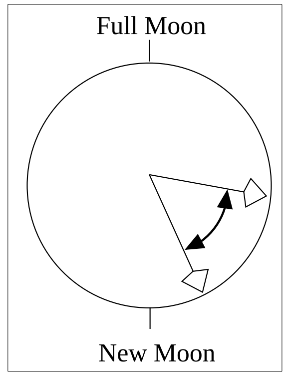


Figure 4.7: Implementation example of a graphical selection method for cyclic temporal aspects.

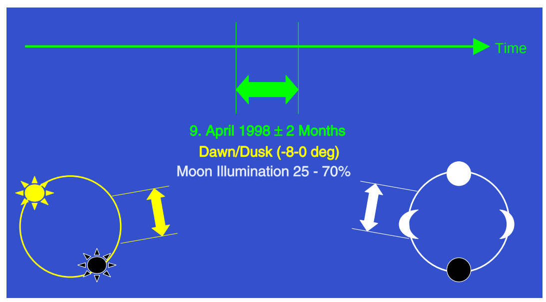


Figure 4.8: Multiple time aspects as an example in the application of TDFs. Data is selected within a certain time span (green), sun altitude (yellow) and lunar illumination (white).

TDFs can be used on a high level of computer abstraction in the field of user interface design. It is not the aim of this study to produce user interface design research, as there exists a large literature about this field itself. This look ahead to Appendix A is meant to give an impression of how TDFs can be used in a more concrete situation.

When multiple time aspects are considered simultaneously logical operators as *and, or, not* need to be defined and implemented to describe the relationship of the TDFs. Figure 4.8 illustrates a possible setup for analyzing data with the combination of three temporal data frames handling several temporal aspects. In this example three aspects are under investigation: The time as a linear component (selection of a two month period around 9.April 1998), the altitude of the sun (dawn and dusk times) and moon illumination (full moon phase).

Further developments of TDFs (e.g. window splitting, inverse selection functions, data shifting) are provided in the prototype application in Appendix A. It is easier to understand and introduce them by providing an example implementation in parallel. I will now provide an example application of the temporal data frames in the next section.

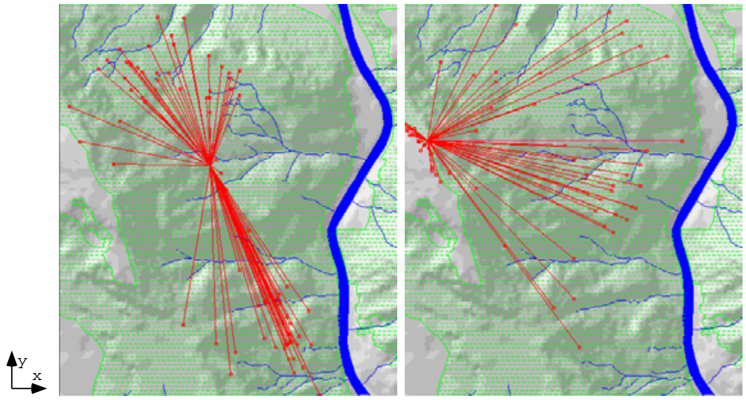


Figure 4.9: Radio locations of one badger in the Sihlwald in September 1996 (Data by K.Hindenlang). Left: full moon phase. Right: new moon phase. Red: star plot from the median center with the observations. A red line is drawn from the median center to each observation. Green = forest. Blue = Sihl river.

4.6 Example Application: Moon illumination and Badgers

Even though the dynamics is essential in the analysis of temporal data, the following paragraphs will try to present a simple analysis of this type on the static medium paper.

The animal species used in this example is the Badger *Meles meles*. They are nocturnal animals that can feed on a variety of food. They are known to feed in forested as well as agricultural areas. A study on the spatial behavior of badgers is conducted in the Sihlwald area by Karin Hindenlang (Zool. Instit., University of Zurich), using radio telemetry to locate the animals. As in many studies on wild animals there are questions about the influence of the moon on the spatial behavior of the badger. Due to a lack of efficient methods for analyzing such cyclic temporal aspects, lunar influences, however, have seldom been verified.

The sample data were analyzed for this purpose to see whether there is an influence in the spatial behavior in one of the badgers followed by telemetric means. As stated above I will only provide a static representation of the findings. In figure 4.9 the data were divided into two phases. In the left map all localizations of the badger during the full moon phases are plotted. The median center was calculated and a so called star or spider plot was produced. It shows clearly that all observations with one exception were located within the forest. In contrast to that the locations made during the new moon phase show a completely different picture (figure 4.9 right). The median center of the observations lies very close to the forest border and a majority of them was made outside of the forest in the meadows adjacent to the forest.

A more detailed analysis showed that the shift from the inner parts of the forest to the forest edge occurred in times from about one third (moon illumination 30 percent) before new moon to about one third after new moon, resulting in about one third of the time as meadow oriented and about two thirds of the time as forest oriented activity.

Using the prototype application TUPF, this kind of analysis took only a few minutes to be performed. Normally this would not even have been conducted. It would have taken a lot of work depending on the sample size and its spatial distribution. Neither the ideas and concepts nor the software for such analysis were available. This little example shows the power of the concept of temporal data frames.

The above data were used for demonstration purposes. The reader is reminded to be careful not to draw false conclusions from the illustration above. The sample dataset I received for this testing analysis was relatively small. To confirm whether there is really such an influence of the moon on this badger's spatial activity, a larger volume of data needs to be analyzed. This will be done as soon as the data are available according to the advances of the badger study.

In Appendix A movie examples for the temporal data frames are provided. They are video-shots of the prototype application TUPF. They shall serve as intermediate illustrations between the static (paper) and the fully interactive (TUPF) descriptions of the concept.

4.7 Outlook: Developments for more Complex Data Analysis

In the above sections the process of selection was illustrated. The application of TDFs is of course not limited to this process. All available statistical and other procedures can then be applied to the data in a very efficient way and changes in them can be discovered. The calculation of median center, dispersion indices or density matrices are only a few examples for this. Some interesting examples can be found in the prototype application described in Appendix A.

One of the major advantages of TDF is the speed at which data can be selected for analysis according to various temporal aspects. An application that could be available soon is closely linked to the technical developments of digital terrain models (DTM) where significant research efforts have been made in recent years. Creating DTMs with resolutions of one meter and below is now feasible at reasonable costs. Novel terrain data capture methods such as laser scanning even allow for the distinction between land surface and vegetation surface. With such data available highly refined analyses can be conducted on temporal phenomena. In wildlife research it was often speculated that shadows could influence the behavior of animals. Shadows from the sunlight or moonlight created at forest edges or other structures can be calculated with high resolution data. Such calculations increasingly important, because direct observations of animals, where such data can be recorded in the field, are replaced to a large extent by remote techniques as radio or satellite telemetry. Hence spatial databases are extremely important in these studies.

With changes in the spatial distribution in animals there are changes in environmental parameters taking place. The temporal data frame concept can be easily extended to a more general data frames concept where environmental parameters can serve as base criteria. Systems can implement such links between temporal and atemporal data and provide desired histograms or other plots instantly during the (exploratory) analysis phase.

Further developments are needed for spatially heterogenous temporal phenomena such as the tides and the much more complex weather. In wildlife studies the latter is very difficult to analyze as an influencing factor. This is due to the three facts that it is spatially heterogeneous, it changes all the time and there are time lag effects which by themselves are very hard to cope with.

There is a last topic to be mentioned here. In this chapter only phenomena with constant cycles are being considered. It is clear that some cyclic phenomena have some deviations from strictly constant periodicities, especially in biological systems. Methods need to be developed to cope with such data. The following chapter introduces a new method to analyze spatial periodicities in animal movements which are not strictly constant.

5

A new Family of Time-Plots

Perception of multidimensional data is a very tricky thing. Usually two dimensional data (e.g. a scatter plot of body height versus body weight) can be easily inspected by eye, whereas tree dimensions are often difficult to analyze. The focus of interest in analysis of such data lies on the interactions between the variables. The complexity of interactions between the variables does not increase in a linear way, but in the order of n!. This problem often urges researchers to abandon visual inspection and interpretation of their data and to start the analysis by statistical means, applying models offered by their statistical packages rather than the researcher's knowledge about the data.

Geographical data which change over time have by definition at least four to five and often more dimensions: two or three dimensions representing space, one dimension for time, and one or several dimensions for the attributes of interest.

In the past several methods were developed to deal with the problem of higher dimensional data. One common approach is to use color (hue, saturation, value), texture and symbols additionally to the two dimensions available on a sheet of paper or a computer monitor to "squeeze" all the dimensions into one graphical representation (Bertin, 1977; Winfree, 1980; Healey, 1997). Basically this effort assumes that by depicting all information into one graphic, one should be able to recognize the patterns hidden in the data. An example of such a plot is given in figure 5.1 for an ant walking on a petri dish.

This method works fine for presenting results at a final stage as it is often done in cartographic representations. For the exploratory and analytical steps it disregards the fact that the human perception is seldom capable of making use of multiple aspects for pattern detection at once. Most of the time overlaying several attributes onto the same graphic is hiding more than it is revealing. Receiving a complex picture we often try to extract partial information out of it to identify the information needed as we do for example in the tests for red/green blindness.

In this section I will present a novel approach for analyzing the patterns in concern here, the movements of animals. The basic idea is to reduce the dimensionality of the data to a level of complexity that does not over-burden our perception. This is performed by reducing the spatial aspects to one dimension (e.g. distance, angle, parallelity) and introducing one or two time axes. Using this technique new analytical plots are created to help identify regularities in the data. By introducing time axes into these plots, an overview of the spatio-

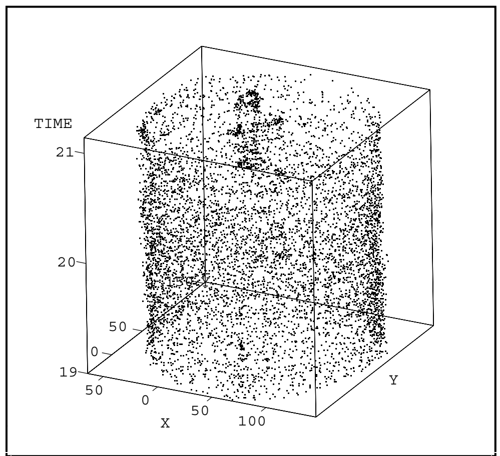
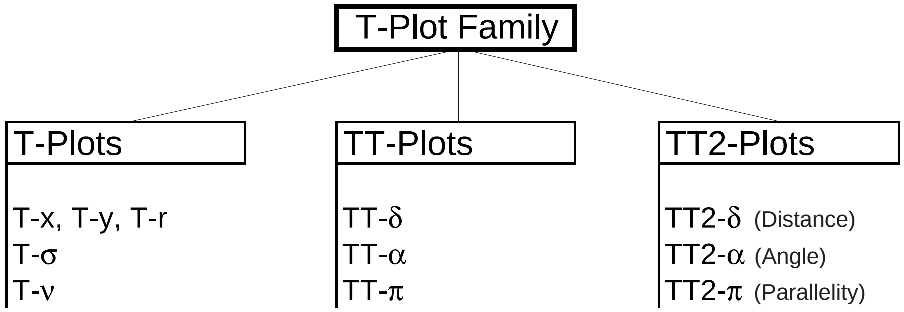


Figure 5.1: Example 3-D plot with two spatial (x, y) and one temporal axis of an ant walking on a petri dish. It illustrates how difficult it is to interpret higher dimensional representations.



Properties		
1 time axis	2 time axes	2 time axes
1 spatial aspect	1 spatial aspect	1 spatial aspect
n datasets	1 dataset	2 datasets
temporal data frames		

Figure 5.2: Classification of the Time-Plot Family.

temporal changes in the movement data can be obtained.

In the following sections the so-called *Family of Time-Plots* (figure 5.2) is presented. The name comes from the *T*ime axis used. The following discussion is divided into three parts. The first part introduces the T-plots (sections 5.1-5.4). These are relatively simple plots with one time axis, where the *Concept of Temporal Data Frames* (Chapter 4) is applied. They visualize the change of one spatial aspect. The next section introduces the TT-Plots, which contain two time axes (sections 5.5- 5.9). They are used to analyze intra-dataset parameters, i.e. characteristics originating from a single dataset. Then the TT2-Plots will finish the section (section 5.10). They also contain two time axes, but they are intended for use with 2 datasets to describe inter-dataset characteristics).

If someone tries to identify geometrical shapes in snow patterns, he or she needs to be able to name a set of shapes such as squares, circles, lines and ellipses or more concrete ones such as a house, a bird, or a sheep as it is done in the Japanese yukigata (snow pattern descriptions, Yamada (1996)). The same is the case when someone tries to identify patterns in point arrangements or movement patterns. Hence the plots are first illustrated using artificial data, starting with simple patterns followed by more complex data (figure 5.3). Then a first approach to an identification catalogue is presented. In chapter 6 applications of these analysis methods to real biological data are presented.

The artificial movement patterns used for demonstration purposes are shown in figure 5.3. Beginning with the most simple form, the first pattern represents a point walking forth and back on a diagonal line (5.3a). The second movement is an object walking on a circle, whereas the third describes a 8-shaped movement. The last and most complicated artificial movement pattern introduced here is a star shape, where the object passes several times through a center in the middle of its activity (5.3d).

These four movement patterns will serve as introductory examples in the next sections.

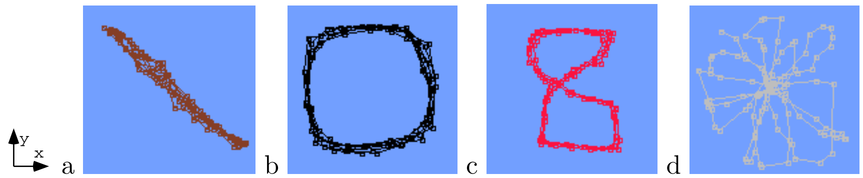


Figure 5.3: Artificial movement patterns used for demonstration purposes (from left to right): a: line, b: circle, c: 8-shaped, d: star.

5.1 Creation of T-Plots

I shall start introducing the Time-Plot family with the easiest type of plots, the T-plots, which only consist of one time axis and one spatial aspect on the second axis. There are three types of T-plots. (1) The first type (T-x, T-y, T-r) considers besides the temporal aspect only one spatial aspect and ignores all others. (2) The second type (T-ν) works with derivatives of the whole spatial information. (3) The third type (T-σ) compresses the spatial information to one measurement as the standard deviation by applying the *temporal data frames concept* described in chapter 4. The first use of these plots is for revealing homogeneous phases within datasets. A second application can be the detection of trends. These plots are easy to understand and some of them may have been used earlier, but they have never been integrated in a comprehensive way. They will give a good introduction to the 'thinking in time' that will be needed when presenting the TT-plots.

5.2 T-x, T-y and T-r Plot

In this section I will introduce the T-x, T-y and its generalized version, the T-r plot. The idea is to neglect one of the coordinates defining the point in space and replacing it with a time axis. In the case of the T-x plot, the x axis is retained and the y axis is replaced by a time axis, whereas in the T-y plot, the x axis is replaced by a time axis. Figure 5.4 shows T-x and T-y plots for the object walking along a line. In this artificial example a sinus-like graph is seen in both plots. In figure 5.4a we can follow the (computer-)animal starting from the bottom left of the graph. The horizontal axis represents the x coordinate, whereas the vertical axis indicates time. The animal then starts moving east until it reaches the right edge, turning sharply around and walking in direction west, restarting again at the left edge. In figure 5.4b the graph is rotated by 90 degrees, now representing the y coordinate on the vertical axis and time on the horizontal axis. This is done for illustration purposes. In this example a similar pattern is produced indicating a movement starting north and then periodically going from south to north and reverse.

Both T-x and T-y plots are bound to the coordinate system used. Using a different coordinate system could result in different results. This limitation is overcome with the T-r(otation) plot, which allows rotating the coordinate system dynamically. This is roughly illustrated in figure 5.5 in which three rotation angles for the coordinate system are used. The little triangle in the upper right corner indicates the rotation angle. In figure 5.5a the coordinate

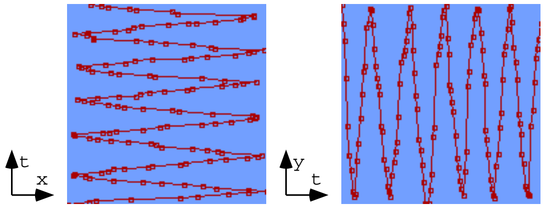


Figure 5.4: Left: T-x plot Right: T-y plot for the artificial movement pattern walking along a line.

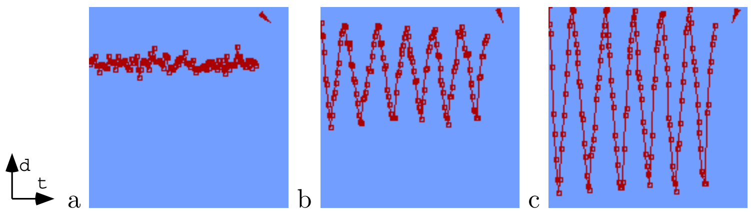


Figure 5.5: T-r(rotation) plot for the artificial movement pattern walking along a line. The arrow in the top-right corner indicates the angle used as axis for the specific plot (from left to right): 132, 161 and 202 degrees.

system is rotated to be more or less parallel to the main movement direction of the (computer-)animal resulting in a narrow horizontal band. In figure 5.5b the coordinate system is not parallel to the main movement direction anymore revealing an oscillating movement, which has its highest amplitude at a rotation angle of about 200 degrees in figure 5.5c.

5.3 T-ν Plot

As an example of a second type of T-plots, the T-ν plot will be introduced here. One of the variables of interest in analyzing movement patterns is the speed an object is moving. This aspect can be easily converted into a time plot where the horizontal axis represents time and the vertical axis speed. Such a plot is illustrated in figure 5.6. One advantage of this kind of plot often found in biological data is remarkable: outliers are conspicuous, often identified as errors in the data. This T-ν plot has been used before, but it is nevertheless presented here as a representative of an own kind of T-plots.

5.4 T-σ Plot

The last category of T-plots I will bring in is represented by the T-σ plot. It is used to describe the change in spread of the point object. As a contrast to the T-r plot, which describes one spatial aspect, and the T-ν plot, which describes a derivative of the spatial information using 2 observations, the T-σ works with a measurement derived from several observations (in the case here with the standard deviation).

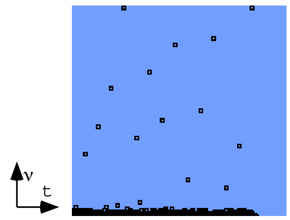


Figure 5.6: T-ν plot for the artificial movement pattern walking along a line.

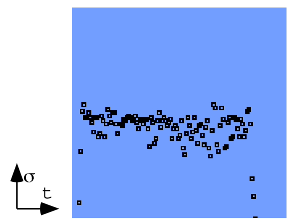


Figure 5.7: T-σ plot for the artificial movement pattern walking along a line.

The approach used here is using the *temporal data frames* introduced in chapter 4. The plot uses time for the x-axis and as one example of a spread index the standard deviation on the y-axis. For calculating the standard deviation at time-point x, all observations within the time frame $x \pm \frac{1}{2} timeframe$ are considered. This is done repeatedly for all time-points. Such a plot is shown in figure 5.7, indicating the change in spread during the observation period. One may assume that a critical and also essential point is how to determine the width of the time frame. This is not the case here, because as with all exploratory data analysis (EDA) techniques (**?**) it is important to be able to change the perspectives and views quickly to search for patterns within the data instead of imposing a rigid pseudo correct corset. As the scientific saying goes, *one man's data is another man's noise*. There is no 'correct' value, instead you need the ability to change and compare different values.

The approach used here - an equal period of time for the selection of objects considered for the calculation of the statistics - could also be replaced by a fixed number of observations resulting in varying time frame widths. This would result in more homogeneous results with scarce data and more inhomogeneous values in dense data.

5.5 Creation of TT-Plots

The construction of a TT-plot is outlined with the example of the TT-δ plot. The TT-plots are a new way to transform the three dimensional data consisting of two spatial and one temporal axis to a two dimensional representation by reducing the spatial component to an inter-event distance matrix and introducing a second time axis. This is performed by creating a matrix of spatial distances for every time-point to all other time-points. The plot can be constructed as follows: The x and y axis are both t (time), while the z axis at the point t_1, t_2

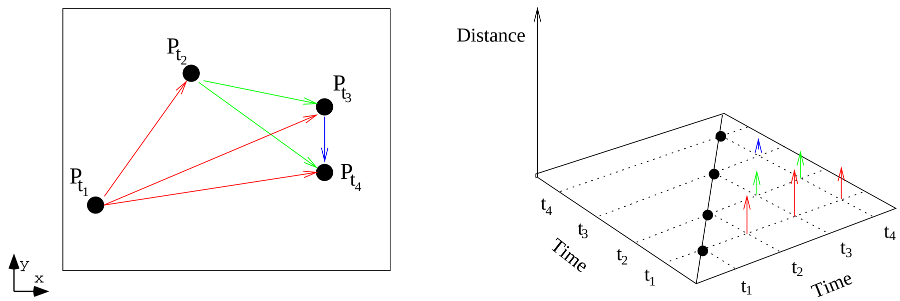


Figure 5.8: Example TT-δ plot construction. Left: Planar view of an animal's path with four observed locations $P_{t_1} - P_{t_4}$. Right: 3-D view of a TT-δ plot construction. See text.

represents the geographical distance δ between the two locations at P_{t_1} and P_{t_2}. Let us have a look at figure 5.8 where four observations are plotted ($P_{t_1} - P_{t_4}$).

For x = t_1, y = t_1 the calculated z is 0 ($P_{t_1} = P_{t_1}$). For x = t_2, y = t_1, z would be the geographical distance between P_{t_1} and P_{t_2}, for x = t_3, y = t_1, z would be the distance between P_{t_1} and P_{t_3} and so forth. Now the calculated distances are drawn as z values (arrows) at the coordinates t_1, t_1 and t_2, t_1 and t_3, t_1 (figure 5.8). This results in a xyz scatter plot that can be interpolated to a surface for easier pattern recognition. The z values lying above the indicated diagonal line from the lower left to the upper right in figure 5.8 can be viewed as mirrored from the lower triangle in TT-plots, because the distance $P_{t_1} - P_{t_2}$ is equal to $P_{t_2} - P_{t_1}$. This duplication of information is done for easier recognition of patterns, especially for the beginning (lower left) and the end (upper right) of the observation period. A second reason for this mirroring is to keep the plots consistent with the TT2-plots introduced in Section 5.10 which need the full quadrangle to represent all the information in the data. Constructing the TT-δ plot in this manner, the distance from a location at time t_1 to the location at time t_n can be 'seen'. In a general mathematical form, the z value is a function of two locations at two time-points of an object:

$$Z_{TT} = f(P_{t_1}, P_{t_2}) \tag{5.1}$$

Z_{TT} = Calculated value at plot location t_1, t_2
P_{t_x} = Object's location at time-point x

Figure 5.9 shows a three dimensional plot with the standard color scheme used throughout this thesis. Short distances are indicated in blue, medium distances in green and large distances are colored in red (table 5.1). This figure serves as a link picture between the illustration for the plot construction (figure 5.8) and the remaining illustrations of TT-, and TT2-plots, which will only contain the color shading without the 3D effect.

Figure 5.10 illustrates the basic features found in a TT-plot that will be used in the following sections: the base diagonal line (5.10:A), a line parallel to it (5.10:B), another line at a 90° angle to it (5.10:C), and the lower right triangle (5.10:T).

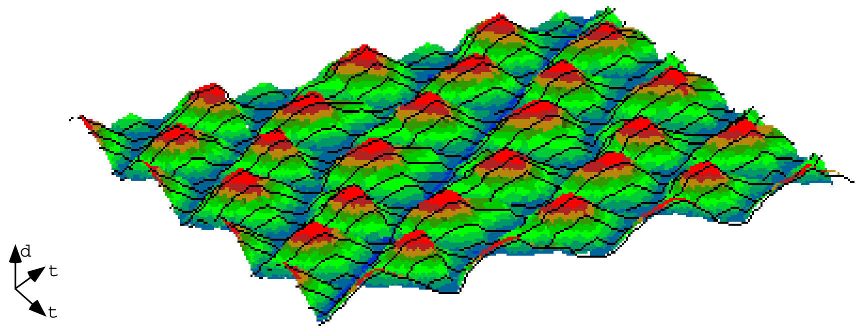


Figure 5.9: Example TT-δ plot (3-d perspective) from the artificial data 'walking on a line'. Short distances are indicated in blue, medium ones in green and large distances have a yellow to red color.

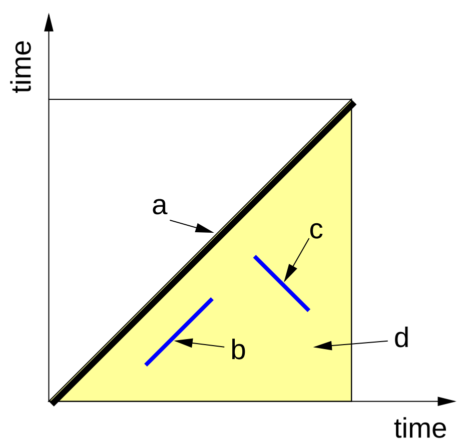


Figure 5.10: Basic features found in a TT-plot: A = base diagonal line, B = parallel line to a, C = line at a 90° angle to a, T = lower right triangle.

Table 5.1: Color scheme for TT-δ plots

Color	Distance (space)
blue	short
green	medium
red	large

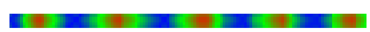


Figure 5.11: Transect for the example TT-δ plot from figure 5.12a.

5.6 TT-δ Plot (Spatial Distance)

The first type of TT-plot introduced here is the TT-δ plot. It describes intra-dataset distances, i.e. the distance between two locations of an animal at two different times. The general procedure of how to create this plot is outlined above. The Z value is calculated as the Euclidian distance between two locations, where the formula for the TT-δ plot can be written as:

$$Z_{TT-\delta} = \sqrt{(x_{P_{t_1}} - x_{P_{t_2}})^2 + (y_{P_{t_1}} - y_{P_{t_2}})^2} \tag{5.2}$$

$Z_{TT-\delta}$ = Distance between locations at time-points t_1 and t_2
$x_{P_{t_x}}$ = x coordinate of object at time-point t_x
$y_{P_{t_x}}$ = y coordinate of object at time-point t_x

5.6.1 TT-δ Plots Generated from Artificial Point Data

To become used to reading TT-δ plots, the following section illustrates several plots resulting from the artificially produced point patterns (figure 5.3). This will help us sharpen our ability to recognize patterns in temporal point data.

The first pattern represents an animal walking forth and back along a line (figure 5.3a). The resulting TT-δ plot is shown in figure 5.12a. The color scheme used in the following figures indicates the distance according to table 5.1.

Now we can start the first steps reading a TT-δ plot. Let us take a look at a thin horizontal transect through such a pattern provided in figure 5.11.

Starting at the left side of the line the animal starts moving. As time passes by when we shift the view in direction to the right, the animal moves away from the starting point. This results in a color change to green and later to red, because the distance represented by the colors on this horizontal transect is measured in relation to the starting point. Now the animal turns around and moves back in direction to the starting point - the color changes again back to green - and finally reaches it, as it is indicated by the recurrence of the blue color. As the original movement path of the animal is 'walking stereotypically to and fro on a line' (as often seen in animals in captivity), the pattern is restarted from the beginning and continues until the end of the transect.

With this knowledge we can now start the interpretation of a whole TT-δ plot.

The first thing (A) we can note in figure 5.12a is the blue diagonal line from the lower left to the upper right corner. On this line, all the distances calculated

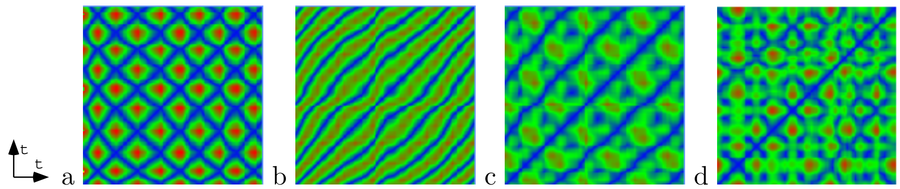


Figure 5.12: Example TT-δ plots from artificial data (from left to right): a) line, b) circle, c) 8-shaped, d) star movement patterns. Short distances are indicated in blue, medium ones in green and large distances have a yellow to red color.

are from location at t_x (horizontal axis) to the location at t_x (vertical axis), i.e. to the location itself and is therefore zero, as indicated by the blue color. This line will hereafter be referenced as the *base diagonal line*. Remember that in TT-plots the lower right triangle delimited by this line is mirrored to the upper left triangle.

As a second pattern (B) we can recognize the blue lines parallel to the before mentioned diagonal. They indicate that the animal walked the same path in the same direction as it did some time ago.

The third features (C) we can see are the blue diagonal lines from the top left to the bottom right at an angle of 90° to the diagonals described above. They show that the animal walked the same path as before, but in the opposite direction.

As the last feature (D) I would like to mention the equally spaced red spots. They indicate that the animal is far away at regular intervals, not only for certain points, but for all points visited.

The second TT-δ plot (figure 5.12b) is an artificial movement from a (computer-)mouse running around in a circle (as shown in figure 5.3b). Looking at a horizontal transect through the figure we have a similar pattern as in the horizontal transect from figure 5.12a, starting close (blue), then going away (to green to red) and then returning to the original point. The patterns (A) and (B) above can also be found in this figure. But as an obvious difference there are no diagonal lines from the upper left to the lower right, indicating that the animal never walks the same way in the *opposite direction*.

The third example is a bit more complex. The movement describes an 8-shape. We can see the parallel blue lines from the lower left to the upper right as in the previous examples, showing that the animal walks the same pathway in the same direction at a later time. In addition to these lines we can see blue spots at regular intervals in between the parallel lines. They result from the intersection point in the middle of the 8-shape.

The last artificial dataset used here is the movement path shown in figure 5.3d which is describing a star shape. Here we meet again several patterns we know from the previous examples, sometimes only as fragments. There are two new characteristics I would like to point out. First, there is a relatively regular blue dot pattern, sometimes overlapping with other structures. The second pattern are the X-shapes in the lower right part of the plot. Such a shape indicates that the animal not only used the same path as some time ago, it also walked back on the same path as it did the first time.

This leads to another characteristic that can be seen within these plots.

When an animal is using the same path twice, one can see the speed it used the second time relative to the speed it had the first time. More concretely, when a T-δ plot shows a blue line in a 45° direction, the animal used the same path in the same speed as the first time. A distortion towards a larger angle shows a faster speed, one towards a smaller angle a slower speed than the first time. The same is true when the angle is around 135°, but the animal is then using the opposite direction on this path.

These four examples shall serve as a short introduction on how to read TT-δ plots. In the next section I will try to compile a first catalogue of features and patterns that can be discovered in this type of plots.

5.6.2 Interpretation Catalogue for TT-δ Plots

From the observations above I shall try to provide a first interpretation catalogue for TT-δ plots. It will start with patterns that are easy to recognize. I shall then provide pattern interpretations which are intended for a more refined analysis.

In tables 5.2ff the base diagonal line from the lower left to the upper right corner is drawn in black. The objects of interest are drawn in blue. Objects mentioned in a former entry are also drawn in black.

The first catalogue entry is a square shape around the base diagonal line (table 5.2a). It indicates that the object resides at the 'same' place for a certain amount of time. In practice the square shape will not be sharply bounded at first. If needed the data can be easily filtered to display only areas that represent intra-event distances up to a defined distance. The second catalogue entry (5.2b) is a line at an angle of 90° to the base diagonal to which it is connected. It shows that the object uses the same travel path in the opposite direction without leaving the original path. When the animal uses a path and after some time reuses a travel path used some time ago, the pattern in table 5.2c occurs. The disjoint lines show that the animal did not turn around and walk back, but made some kind of a loop before rejoining to the original path. In table 5.2d the line of interest is parallel to the base diagonal line. For the observation time in question it means that the animal uses the same travel path in the same direction as before. The amount of time passed between the two lines is equivalent to the horizontal distance between the two lines. Smaller distances mean a shorter interval and vice versa. The last simple catalogue entry is shown in table 5.2e. Whenever a location is visited a second time, but the animal uses a different path to and a different path from the location than the first time, a single (blue) dot occurs in the TT-δ plot. The next two catalogue entries are more complex. 5.2f is a repeated pattern of entry b. The same path is used repeatedly in both directions, describing twice a forward and backward movement. The last pattern comes from an animal running four times on a loop. It does not have to be a regular loop as a circle or a rectangle, but can be any shape which does not contain intersections.

Table 5.3 lists four extensions to the first catalogue entry (5.2a). In 5.3a1, after staying at a location for some time, the animal leaves the area using the same path as it came from. When an animal is staying for some amount of time at the same location it did some time ago, the resulting pattern (5.3a2) in a TT-δ plot is a rectangle at the same y-coordinate (t_2) but displaced to the right (t_1). How long the animal stayed in an area is proportional to the size (horizontal diameter) of the rectangle. In the special case shown in table 5.3a3

Table 5.2: Basic interpretation catalogue for TT-δ plots. The shapes of interest are drawn in blue. Shapes that already occurred in previous catalogue entries are drawn in black. The black diagonal line (lower left to upper right) indicates the base diagonal line (see figure 5.10).

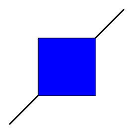
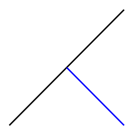
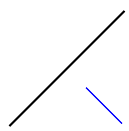
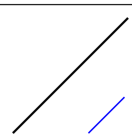
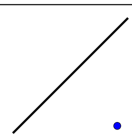
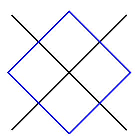
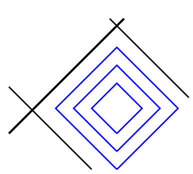
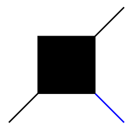
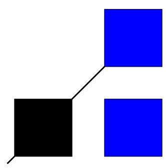
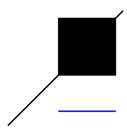
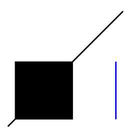
	a. The animal stays for a period of time at the same place. (see also table 5.3)
	b. The animal walks on the same path as it came from, but in the opposite direction. (see also table 5.4)
	c. The animal walks for some time on a path and then returns to the path used earlier, continuing its way on the same path but in the opposite direction. (see also table 5.5)
	d. The animal walks the same path some time later in the same direction as the first time.
	e. The animal passes by a location it already passed before, coming from and going in a different direction than the first time.
	f. The animal is using a path forth and back and a second time forth and back.
	g. The animal is using the same path four times in one direction (e.g. walking on a circle), then reverts its direction, walking the same path four times in the opposite direction, but still on the same path, and then again uses the path four times in the same direction as the first time.

Table 5.3: Extended interpretation catalogue for TT-δ plots of variant a.

	a1. The animal stays for some time at the same location and leaves it via the same path as it came from.
	a2. The animal stays at the same place it did some time ago.
	a3. The animal visited a place for a short time where it stayed for a longer time afterwards ('recognition tour').
	a4. The animal visits the same place again for a short time where it stayed for a longer time before.

the animal stays for a longer time in an area it visited only for a short time before. This could be interpreted as some kind of searching for good places and then returning to the 'best' one. The opposite pattern is shown in table 5.3a4, where the animal passes the same location it stayed for a longer period only for a short time.

A finer analysis of the catalogue entries b and c (table 5.2) can reveal relative travel speed differences (tables 5.4 and 5.5). If the angle to the horizontal becomes smaller than 45°, it reveals that the animal was walking at a slower speed than the first time (b1 and c1) and vice versa (b2 and c2). Relative speed changes are indicated by changing angles (b3, b4 and c3, c4).

5.7 TT-π Plot (Parallelity)

The section above used the distance as one spatial aspect. This section introduces a second type of the TT-Plots, called TT-π plots, which considers the intra-dataset parallelity aspects in a time plot. One example application for this type of plot is the analysis of foraging movements of an animal which is 'scanning' an area. In figure 5.13 the creation of a TT-π plot is illustrated.

First the point P1 is considered as basis for the calculations. The differences in the direction of the walking path from point P1 to all subsequent points (P2-P4) are then calculated. This results in values from 0-π (0-180 degrees). Then the same procedure is applied starting from the next point (P2) as basis. This results in a plot similar to figure 5.8b) with the difference that the length of the arrow now represents the angular deviation from the base point. This matrix is then transformed into the same color scheme used in the previous plots

Table 5.4: Extended interpretation catalogue for TT-δ plots of variant b.

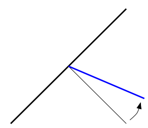
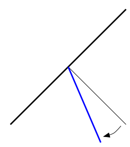
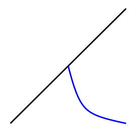
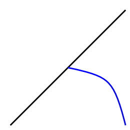
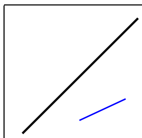
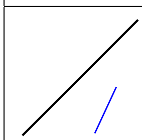
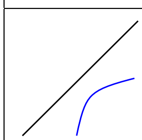
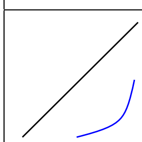
	b1. If the angle from the horizontal becomes smaller than 45°, it means that the animal is walking slower than the first time.
	b2. If the angle from the horizontal becomes larger than 45°, it means that the animal is walking faster than the first time.
	b3. Speed changes on the same travel path are indicated by changing angles from the horizontal. In this example the animal starts walking back on the same path very fast, then becoming slower.
	b4. In this example the animal starts walking back on the same path very slowly, then becoming faster.

Table 5.5: Extended interpretation catalogue for TT-δ plots of variant c.

	c1. The smaller the angle from the horizontal, the slower the animal is walking along the same path than it did before.
	c2. The larger the angle from the horizontal, the faster the animal is walking along the same path than it did before.
	c3. Speed changes on the same travel path are indicated by changing angles from the horizontal. In this example the animal starts walking on the same path very fast, then becoming slower.
	c4. In this example the animal starts walking on the same path very slow, then becoming faster.

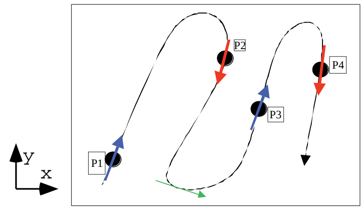


Figure 5.13: Creation of a TT-π plot. The colors indicate the parallelity relative to P1.

Table 5.6: Color scheme for TT-π plots

Color	Parallelity
blue	parallel
green	90°
red	anti-parallel (180°)

(table 5.6) to maintain a consistent appearance. Blue colors indicate a parallel walking direction, green colors show a walking direction of about 90° and red colors represent an opposite walking direction.

5.7.1 TT-π Plots Generated from Artificial Point Data

As it was done for the TT-δ plots, examples from artificially produced point movement patterns are presented here. The original data can be seen in figures 5.3a-d.

The first pattern results from the animal walking forth and back along a line (figure 5.3a). The resulting TT-π plot is presented in figure 5.14a. There are basically three characteristics that I would like to point out in this plot. First (A) a base diagonal line from the lower left to the upper right can also be found in this type of plot. Here it represents the difference in direction from a point to itself, which is always zero, i.e. they are exactly parallel. The second feature (B) is a blue/red pattern similar to a checker board with hardly any other colors in between. It shows that the animal almost exclusively uses two directions which are at a 180° angle. The third feature (C) are the thin green lines which indicate that there are short periods at which the animal uses another direction, in this case probably turn-points as the one indicated with a green arrow in figure 5.13. From the size of the blue squares it can be easily determined (visually or by calculations from the matrix) how long the animal keeps on going into the same or similar direction.

The second pattern illustrated in figure 5.14b comes from an animal running around in a circle. The resulting TT-π plot is very different from the previous one. Here the width of the 'blue phases' is much smaller than in the previous example indicating a smaller tendency to go on walking into the same direction. The checker board pattern is replaced by a cluster of diagonal lines going repeatedly from parallel (blue) to a 90° angle (green) to anti-parallel (red) and then back to parallel (blue). It is important to know that this pattern is independent

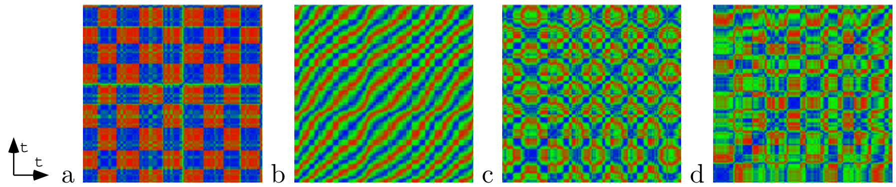


Figure 5.14: Example TT-π plots from artificial data (from left to right): a) line, b) circle, c) 8-shaped, d) star movement patterns. Blue indicates parallel, green intermediate, red anti-parallel.

from the place where this circular path is located. When the animal is shifting to a direction, but maintains the circular path, the same TT-π plot results.

The third example (figure 5.14c) includes a similar structure as the one above, the blue lines being parallel to the base diagonal line. Additionally we find blue parallel lines at an angle of 90° to the base diagonal line. They indicate a fact I shall illustrate with a simple situation. It is a movement which is, from the point of view of the direction, mirrored at a line. For example if the animal is making a half circle turning to the left and then performing a half-circle turning to the right, the resulting pattern in a TT-π plot would be a blue x-shape. The two half circles can be transformed into each other by mirroring one of them at a line. At what angle from the horizontal the upper left/lower right diagonal line is appearing depends on the relative speed used in the two parts of the movement. The third structure that can be recognized are the more or less distorted red circles which are arranged in two alternating rows and columns. They indicate a regular movement which is divided into two parts as the one mentioned above, but in an anti-parallel direction. According to the size of the blue areas the tendency to keep on going into the same direction is in between the first and second example above.

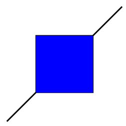
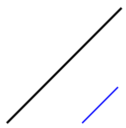
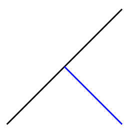
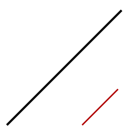
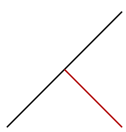
The TT-π plot for the last artificial movement presented here is shown in figure 5.14d. It is based on the star-shaped movement of the (computer-)mouse. First we can see at several places in the plot an alternation of parallel (blue) and anti-parallel (red) movements (e.g. lower left corner, second and last quarters of the observation period). There seems to be a tendency towards a pendular movement with an angle of approximately 180°. The exact values would have to be extracted from the plot. The blue squares are not so regular as in the previous examples. From the statements above it can be deduced that the animal in concern has a higher variability in how long it keeps on going in the same direction.

5.7.2 Interpretation Catalogue for TT-π Plots

As we are more used to think in terms of distances than in terms of parallelity patterns, the following start of an interpretation catalogue is limited to the basic forms.

The first catalogue entry (table 5.7a) looks like the first entry in the TT-δ plot section, although the meaning is quite different. Here the animal is walking continuously along a straight line. When an animal is walking on a path parallel to one used previously the pattern in table 5.7b occurs in the TT-

Table 5.7: Basic interpretation catalogue for TT-π plots. Blue indicates parallel, red anti-parallel.

	a. The animal is walking a straight line.
	b. The animal is walking parallel to a path used earlier.
	c. The animal walks along a path which is (in terms of parallelity) mirrored at a line.
	d. The animal walks anti-parallel to a path used earlier.
	e. The animal walks anti-parallel to a path used earlier in the opposite direction.

π plot. A blue line at an angle of 90° to the base diagonal line can indicate a more complicated fact. The animal first uses a certain path. Then it uses a path which describes the mirrored shape of the original path used the first time, where the mirroring was done across a line. This is the general interpretation of this pattern. A special case is the inversion of the direction. A simple example would be a circle shape which is used at a different location a later time as the travel path, but this time in the opposite direction (e.g. clockwise and then counterclockwise). The pattern shown in table 5.7d represents an animal which is using the opposite direction used before. (Note: This does not mean that it is walking on a straight line). The last pattern shown in table 5.7e is comparable to the situation in table 5.7c, where the following is true, but now for an anti-parallel walking direction: the further the animal goes, the longer the time has passed since the animal was using the opposite direction to the current one. This pattern can also occur when the direction is inverted.

5.8 Further Analysis of TT-Plot Data

In the previous sections I often used a visual language using colors to make communication easier. It must not be forgotten that the color schemes used for presenting the TT-plots are only a representational feature. They always represent measurements extracted from the original data in the form of distances or angles which finally build up the matrices called TT-plots.

The aim of the Time-Plot family is to provide fast methods for recognizing and pinpointing special features in large datasets. After a first inspection the analysis can be refined and extended in several ways:

- zooming in by using a smaller time frame
- reselection of values within a TT-plot (defining values of interest)
- reclassing the values into categories
- performing transformations on the values (e.g. logarithm)
- building histograms of values
- calculating statistical values from the matrix (mean, variance etc.)
- Fourier analysis of the TT-matrix

This list is not complete but indicates in what directions further analysis of these plots may lead. Examples of this kind are given in chapter 6 together with the presentation of applications to biological data.

The above sections introduced two kinds of TT-plots, the TT-δ plot and the TT-π plot which describe distance measurements and parallelity aspects, respectively. These are only two aspects that can be used for a temporal analysis in TT-plots. The creation of other TT-plots can be more or less performed in a similar manner as above. For example the following characteristics could be used for other TT-plots:

- altitudinal difference
- differences in attribute data angle difference from a specific direction (e.g. north (implemented in the prototype), or valley main direction)
- relative speed
- derivatives from the above (speed change, direction change etc.)

The list is not intended to be comprehensive but instead to give some ideas for further developments of this technique.

5.9 Combined and Superimposed TT-Plots

In this short section I will provide two representational extensions to the TT-plots. Both of them are combinations of different spatial aspects into one comprehensive plot.

In the TT-plots introduced by now the information was always mirrored from the lower left triangle to the upper right triangle of the square plot area. This was done for two reasons. The first one is often applied in pattern detection. It is the replication of information, so that patterns at the border of a plot can also be seen, because often only half of the pattern is visible due to border effects. The second reason is to keep the TT-plots consistent with the plots presented later (section 5.10), which make use of the whole area in the plot to depict all

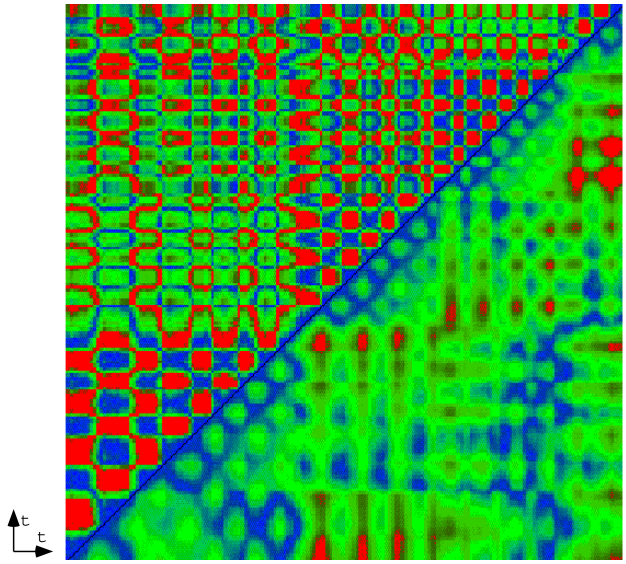


Figure 5.15: Example for a superimposed TT-plot from artificial data. It is a combination of a TT-δ plot (lower right triangle, short distances are indicated in blue, medium ones in green and large distances have a yellow to red color) with a TT-π plot (upper left triangle, blue indicates parallel, green intermediate, red anti-parallel).

the information. Now I will break this restriction in spite of the general rules mentioned in chapter II.

Another way of using the space normally used by mirroring the data at the base diagonal line in a single TT-plot is to combine two different TT-plots into a single graph. This is illustrated in figure 5.15.

The lower right triangle is representing a TT-δ plot describing distance characteristics whereas the upper left triangle of the graph is used for a TT-π plot visualizing the parallelity aspects. Here the main purpose is the graphical representation of these aspects and no further analytical computations can be performed upon this data in contrast to the previous TT-plots (cf. section 5.8). This combination of two TT-plots into one graph has the advantage that they can be directly compared with time as reference system.

A second way of combining two TT-plots into one plot is to use different representational techniques for each of them. In figure 5.16 a TT-δ plot is plotted in the standard manner with the color scheme. But in addition the graph contains an overlayed TT-π plot represented by the small black arrows indicating the parallelity. An arrow with a northern direction represents a parallel direction, south represents anti-parallel. Compared to figure 5.15 it is easier to look at both aspects in areas with a large time difference, whereas in the previous graph it is easier to interpret the data lying relatively close to the base diagonal line.

It is not my intention to go into much detail about visualization techniques. Instead I would only like to encourage the readers, software developers and users to have a look at the specialized literature in visualization (e.g., Shepherd, 1995; Lippert-Stephan, 1996; **?**; **?**).

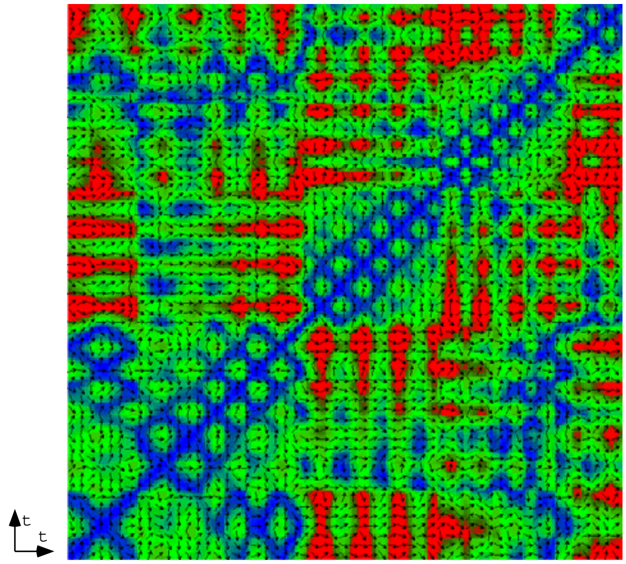


Figure 5.16: Example for a superimposed TT-plot from artificial data. The TT-δ plot is overlayed with the TT-π plot (arrows). Short distances are indicated in blue, medium ones in green and large distances have a yellow to red color. Arrows are indicating the parallelity: north = parallel, south = antiparallel.

5.10 Extensions for Two Objects: TT2-Plots

In the previous sections about the TT-plots intra-dataset features were considered as the point of interest. Now I would like to extend this concept to the analysis of inter-dataset characteristics, i.e. characteristics originating from two datasets. The movements of an animal in its environment is also influenced by the movements of its neighbors. This can be accomplished by simultaneously observing the animals, but it is very difficult to find time lag effects with the traditional methods as for example Markov chain analysis in behavior research. Spatial effects are hardly ever considered. The extension of TT-plots to enable the analysis of two animals' movements can provide insights into the spatial components of the behavior of two animals.

TT2-plots can be constructed in a similar way as TT-plots. The difference is that the measurement of interest does not refer to one single animal, but from one animal to the other. In a TT2-δ plot, which describes spatial distances, the distance from one animal to the other is measured for all time lags (figure 5.17).

This results in a distance matrix which can be visualized in the same way as the TT-plots, where the x and y axis are used for time, and the z axis represents the distance between the two animals. Also the same types of TT2-plots can be constructed as the ones mentioned in the TT-plot section (e.g. distance, parallelity etc.).

There are two main differences between TT- and TT2-plots which need to be outlined. First the TT2-plots do not contain a symmetry axis from the lower left to the upper right of the graph anymore. The distance from animal A at time point t_1 to the animal B at time point t_2 is not equal to the distance A(t_2)-

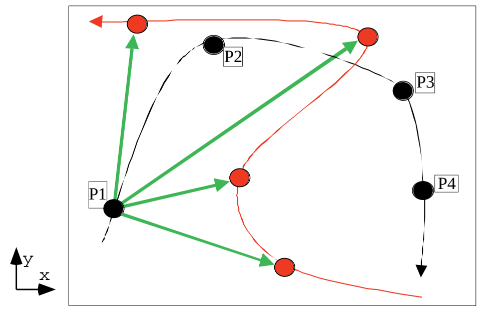


Figure 5.17: Creation of a TT2-δ plot.

B(t_1). Hence the full plot area is needed to represent the whole information. Second there is no base diagonal line visible because of the same reason.

The interpretation of TT2-plots can be easily deduced from the interpretation of TT-plots. To avoid repeating the same statements made earlier I will concentrate in the remainder of this section on the illustration of a single artificial dataset illustrated in figure 5.18a-c.

There are two (artificial) animals walking around simultaneously. The first one is the same one used in the previous examples walking on a straight line. The second one is walking around in a meandering movement (figure 5.18a). The corresponding TT2-δ plot is shown in figure 5.18b. Having a look at the (formerly called) base diagonal line we can see that the animals have six occasions where they come relatively close to each other, indicated by the blue phases lying on the diagonal line. The first animal (black) often passes again locations it crossed before, and which are additionally lying on the other animal's path. The second animal on the contrary does seldomly pass a location twice where it encountered the other animal's track. This can be seen when building vertical and horizontal transects through the TT2-δ plot. Possibly two phases can be distinguished by the blue '<-shapes' and '>-shapes'. In the TT2-π plot (figure 5.18c) a distinct pattern can be seen at the beginning of the observation period. The two animals walk parallel to each other, then anti-parallel, and afterwards the first animal keeps on performing the same parallel-anti-parallel walking pattern while the second changes its direction mainly to a 90° direction from the first animal. In the second half of the observation period it changes its direction to an angle of approximately 45° from the first animals direction until in the last phase it approaches again similar directions (parallel and anti-parallel) as the first animal.

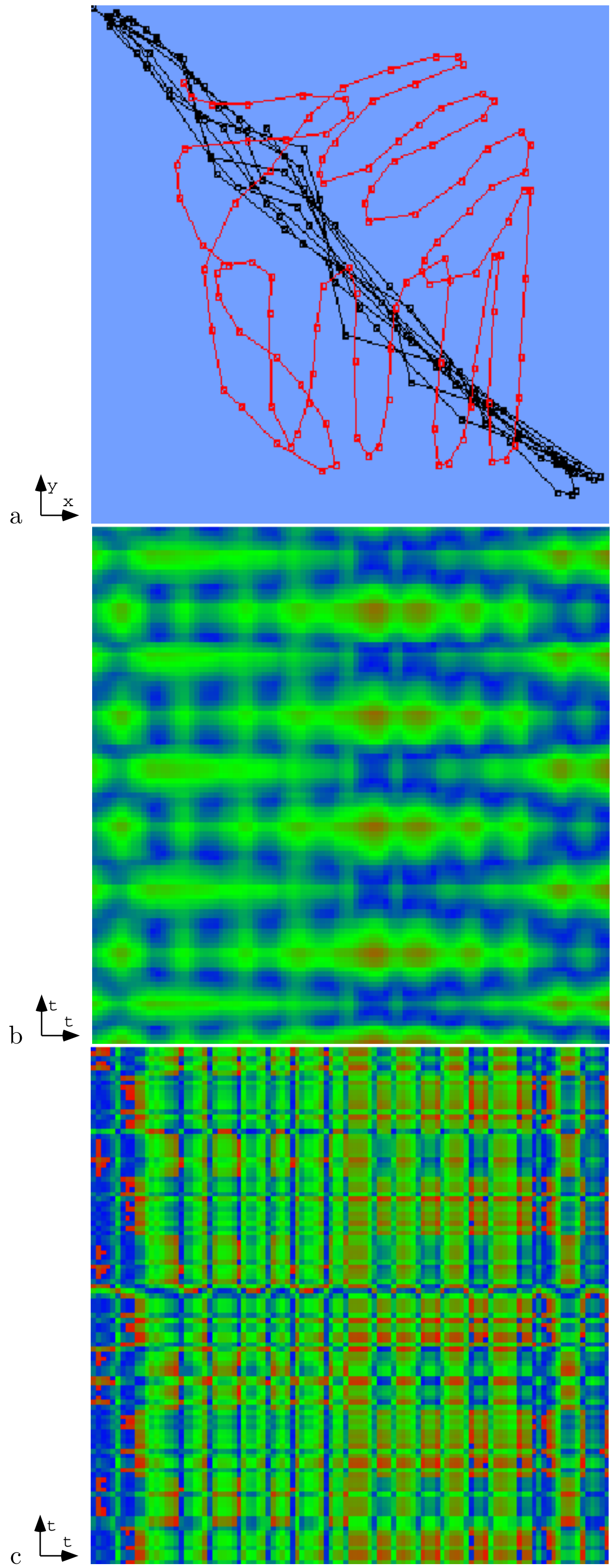


Figure 5.18: Example TT2-δ and TT2-π plots from artificial data: a: original data. b: TT2-δ plot. Short distances are indicated in blue, medium ones in green and large distances have a yellow to red color. c: TT2-π plot. Blue indicates parallel, green intermediate, red anti-parallel.

6

Applications of the Time-Plot family to real biological data

The previous chapter provided the theoretical and technical basis for the Time-Plot family. Now I would like to apply it to real biological data. This will illustrate and verify the methods introduced and will provide a first test whether they can function as useful instruments to a wildlife researcher.

I will use data from five species of animals in the following sections: four woodstorks *Mycteria americana* , an ant *Formicidae*, four lynx *Lynx lynx* and a greater mouse-eard bat *Myotis myotis*. The first data set was collected by a research group using satellite telemetry. The last two data sets were collected by researchers using traditional radiotelemetry installations, one using hand-held antennas, the others using fixed towers for collecting the bat's locations.

In my experience the TT-plots need some time to get familiarized with. I will therefore put the main focus in this chapter on the application of TT-plots, mainly TT-δ plots.

6.1 Defining Biological Phases with T-Plots: Woodstorks

The data used in this section is from four woodstorks *Mycteria americana* monitored by the Woodstork Project in the area of Georgia and Florida in the US (http://allison.clark.net/pub/wcsweb/stork/). The data were collected from 27.8.1996 to 20.5.1997. An overview of the data is given in figure 6.1.

In the T-x plot illustrated in figure 6.2b there seem to be no major changes within the observation period. When examining the figure in greater detail, a difference can be seen that in the first half of the observation period there is a higher tendency for excursions in the x-direction compared to the second part of the observation period. The picture is quite different when looking at the T-y plot shown in figure 6.2b. Here we can distinguish several phases in the data. In the first phase, the birds stay relatively constant at the same location (latitude in this case). After that, quite an abrupt and simultaneous change

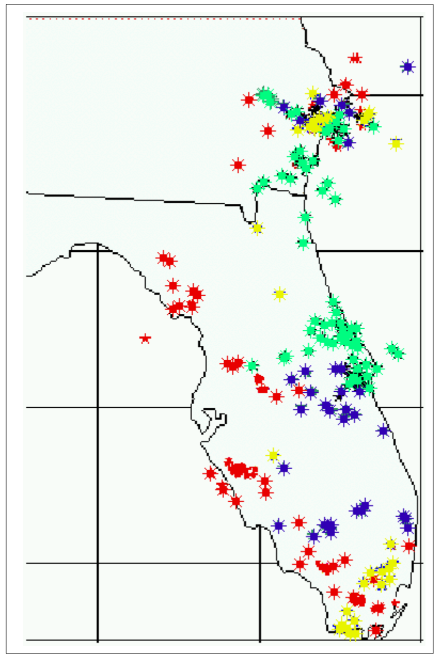


Figure 6.1: Overview of the Woodstork data from Florida and Georgia.

can be observed when all four birds leave the area. Now we can distinguish two types of movements. The first type is shown by the woodstork represented with yellow markers in figure 6.2b. It moves directly without any longer resting periods to the southernmost area where it will stay. The same is true for the woodstork with the green color in figure 6.2b, with the difference that it only migrates about half the distance compared to the former bird. The second type of movements can be seen in the two birds marked with red and blue colors. They also leave the area at about the same time as the other ones, but they make a longer stay at the same latitude as the 'green' bird. After that they go on migrating further south until they reach approximately the same latitude as the 'yellow' woodstork which directly migrated down to that area.

In the third period all birds[1] seem to have stayed at the same locations without larger excursions. Interestingly there is a change in location in two birds (red and blue) moving further north, and when examining the plot for the bird indicated in green, the same is true yet to a much smaller amount. This may well indicate that this should be considered as a fourth period in the data. The fifth period of time starts with the spring migration back to the north. It is interesting that the woodstork which already made a longer stay during the migration southwards again stays for some time at a latitude in between the winter residence and its breeding site. Another thing that is noteworthy is that it looks as if the spring migration takes a longer time than the migration in fall. In European regions the inverse is normally observed for birds. The spring migration is shorter than the fall migration.

[1]Unfortunately only data until 17.12.1996 were available for the bird indicated in the plots in yellow.

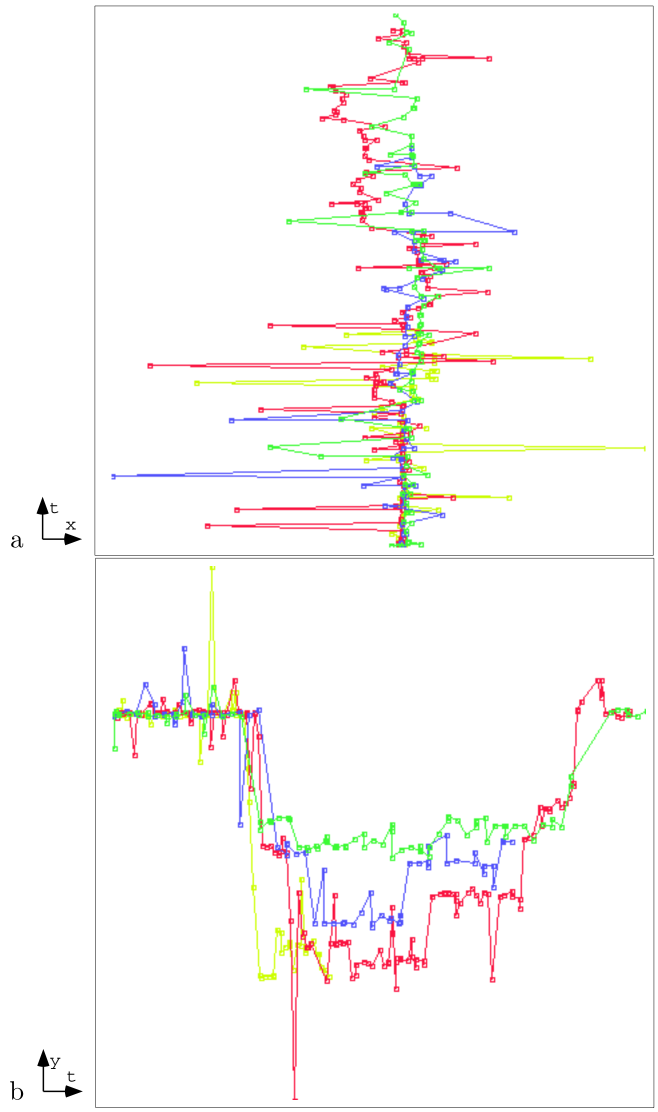


Figure 6.2: a) T-x and b) T-y plots for the woodstork data from 27.8.1996 to 20.5.1997.

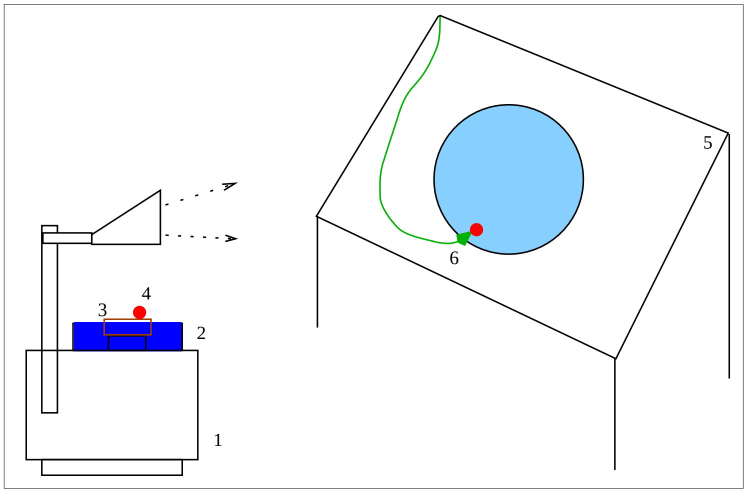


Figure 6.3: Instrumental setup for the data collection of the ant. 1. overhead projector, 2. tank filled with water to keep the heat from the animal and to prevent the ant from escaping, 3. petri dish, 4. ant, 5. digitizer tablet, 6. digitizer mouse.

These two T-plots illustrated are easy to create, but as became evident above, they provide a splendid overview of the spatial movements of these birds.

6.2 TT-δ Plots from Biological Point Data

6.2.1 Ant in a Petri Dish

In a laboratory experiment I recorded the spatial movements of an ant. The recording needed to be very effective with the smallest amount of technology possible. The following setup was then created (figure 6.3. The ant was put on the backside of a petri dish. To prevent the ant from escaping the petri dish was put in the middle of a tank filled with water, so that the surface of the petri dish and the water were at the same level. The water also prevented the heat from the overhead projector used to affect the ant. The projector made it possible to digitize the ant's movements directly on a digitizer tablet, to which the installation was projected.

With this setup, four runs of five minutes each with pauses of five minutes in between were recorded. Locations were recorded every second. In the first two runs, only the animal was situated on the petri dish. After another pause of five minutes, a small piece of a rubber cannula was put onto the petri dish, and immediately after that the third run lasting another five minutes was recorded. After the third pause, a piece of sugar was offered to the ant and left on the petri dish, immediately followed by the recording of the last run. After this procedure the ant was immediately returned to the place of its capture. The observations after its release gave no indications of abnormal behavior.

In figure 6.4 a T-y plot is shown for the first run. Three things can be noted. The first is that the animal was almost always moving. The second thing is

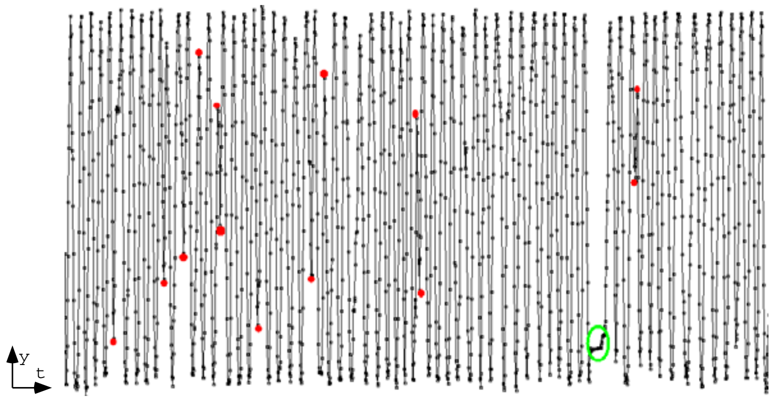


Figure 6.4: T-y plot for the ant in phase 1. Turn-points are marked with red bullets, the other event mentioned in the text is marked green.

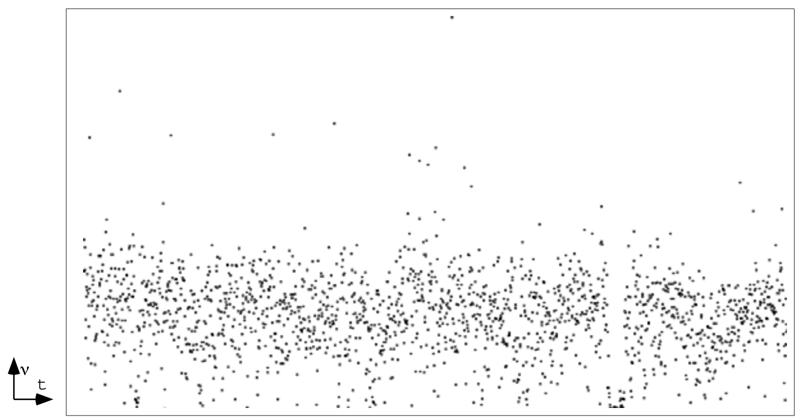


Figure 6.5: T-ν plot for the ant in phase 1.

that the animal made a longer rest after about three thirds of the time, marked with a green circle in figure 6.4. The third and probably most interesting thing noticeable are the turn-points of the ant, i.e. the time-points at which the animal was changing its direction by a sharp turn. They are indicated by red bullets in the figure. More precisely it can be seen that in the first half of the run there was a tendency to show more turn-points than in the second half.

In figure 6.5 a T-ν plot indicates the speed at which the animal was running in phase 1. The speed of the ant is decreasing during time (Regression values corrected for temporal autocorrelation (data when the ant was standing still excluded): r^2=0.23, p=0.041).

As a last T-plot with a single time axis a T-σ plot is illustrated in figure 6.6.

In this plot the standard deviation for the locations within a moving temporal data frame of 1 minute and 15 seconds is displayed. The period in which the ant stayed for some time at one location can be easily depicted in the last quarter of the plot. Aside from that several times there are periods with a minimum

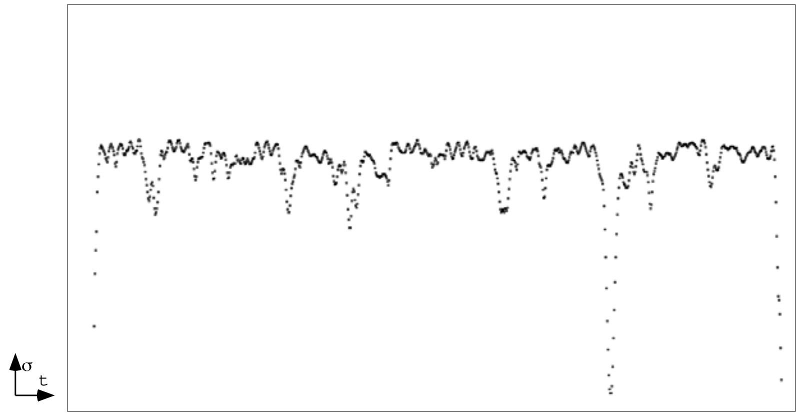


Figure 6.6: T-σ plot for the ant in in the middle of the observation period in phase 1. Time-frame width is set to 1 minute and 15 seconds.

in the standard deviation. They coincide with the turn-points discovered in figure 6.4 with few exceptions. These exceptions are probably additional turn-points which have not been identified before due to fact that the T-x plot was used as a static plot. When used in the dynamic form as a T-r plot, in which the direction of the spatial axis for which the plot is drawn can be changed interactively, they would have been easily recognised. This is again a restriction based on the static characteristics that printed graphics are confronted with. It is not an ideal way to illustrate dynamic interactive processes.

After these relatively simple plots I will now provide the TT-δ and TT-π plots for the four runs. The plots are illustrated in figures 6.7-6.14.

Let us first have a look at the TT-δ plots for the first two 'undisturbed' runs (figures 6.7 and 6.8). The first characteristic we can see is that almost all of the plot is covered with a more or less regular pattern of parallel and antiparallel blue lines in respect to the base diagonal line. They indicate that the ant is continuously running on the same path (in this case along the edge of the petri dish), sometimes changing the direction in which the path is run. This is a very stereotype movement pattern. The second thing noted is the blue square in the upper right quarter of the plot. This shows a rest taken by the animal. The animal has passed this location several times before, and passed it several times afterwards, as it is indicated by the numerous blue horizontal lines to the left and right (and of course below and above) the blue square. In the second run (figure 6.8) these regularities persist, but there are at least two things that are different from the first run. The first is that in the middle of the run the animal continued running in the same direction for a long time. After that a second difference can be noted when the animal changed its behavior. It then ran for five sequences in a row three times in one direction, and then three times in the other direction (see table 5.2g).

When a rubber cannula was put onto the petri dish at the beginning of run three (figure 6.9) this 'pattern' changed a little. It can now be seen that the ant stayed at several occasions at a location. By examining the TT-δ plot in more detail, it can be found that there were six places where the animal was

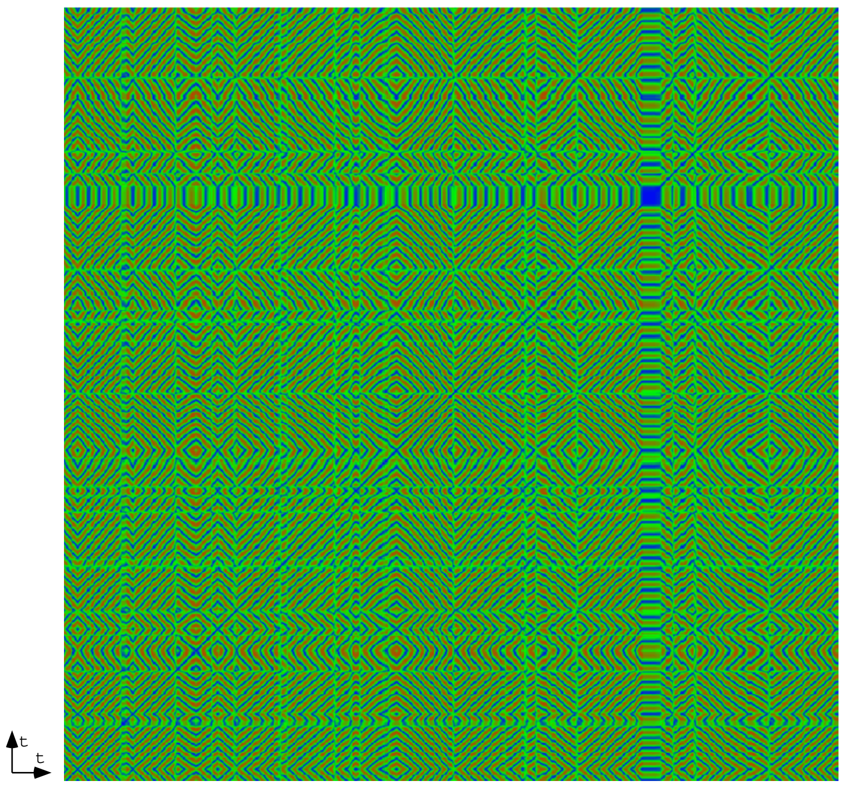


Figure 6.7: TT-δ plot for the ant in the petri dish phase 1 (undisturbed). Short distances are indicated in blue, medium ones in green and large distances have a yellow to red color.

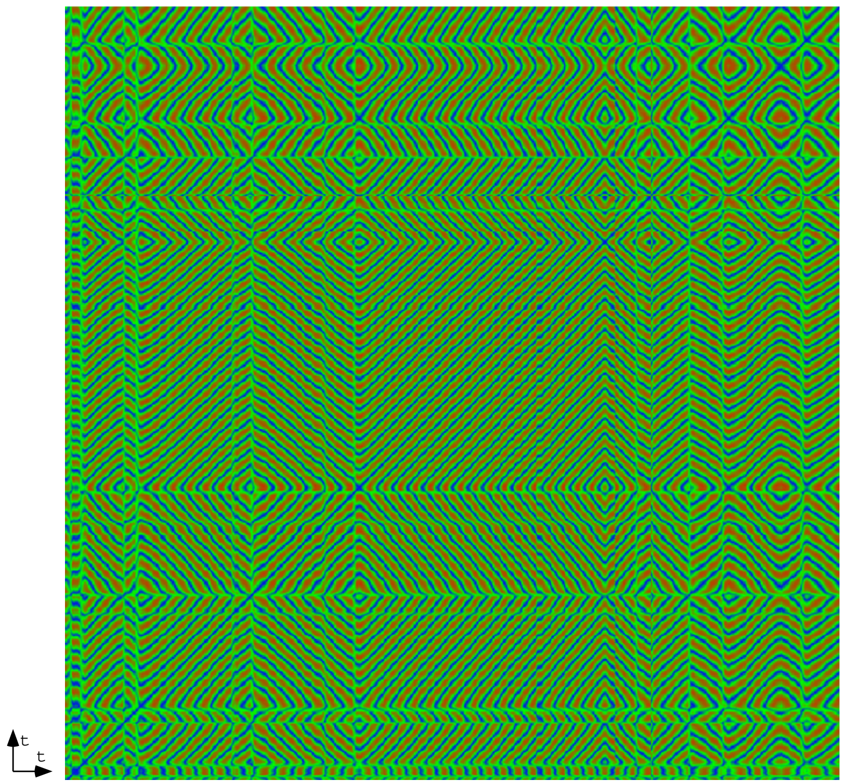


Figure 6.8: TT-δ plot for the ant in the petri dish phase 2 (undisturbed).

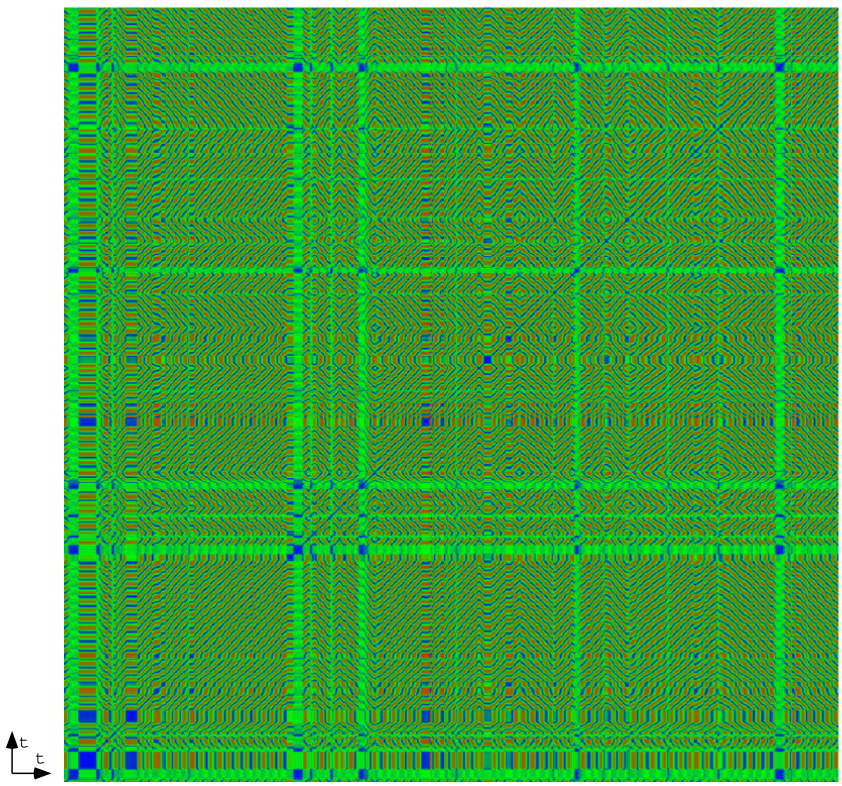


Figure 6.9: TT-δ plot for the ant in the petri dish phase 3 (rubber cannula).

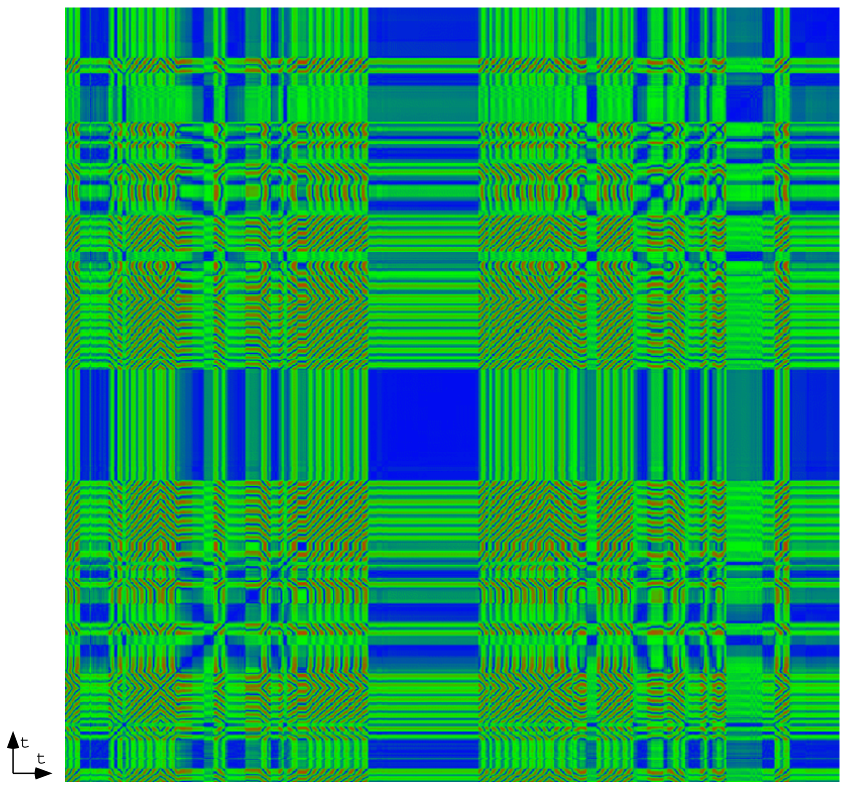


Figure 6.10: TT-δ plot for the ant in the petri dish phase 4 (sugar).

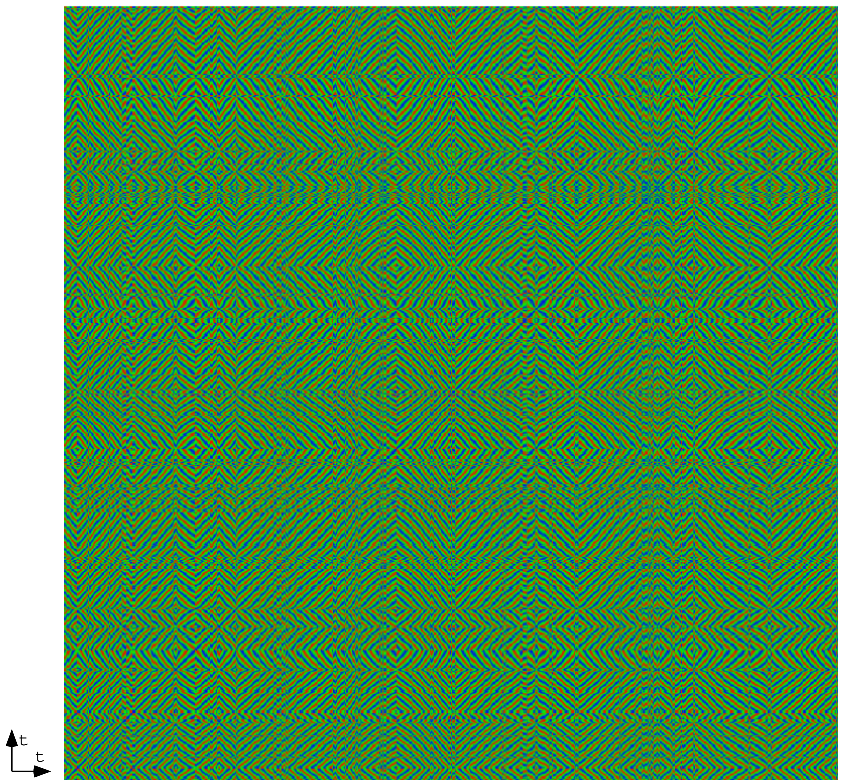


Figure 6.11: TT-π plot for the ant in the petri dish phase 1 (undisturbed). Blue indicates parallel, green intermediate, red anti-parallel.

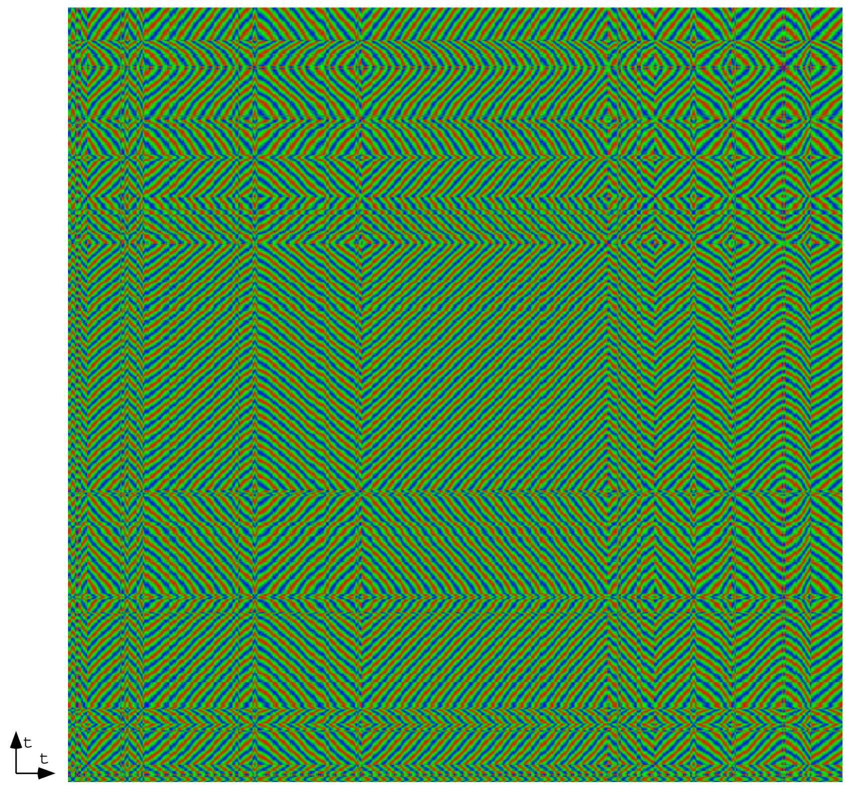


Figure 6.12: TT-π plot for the ant in the petri dish phase 2 (undisturbed).

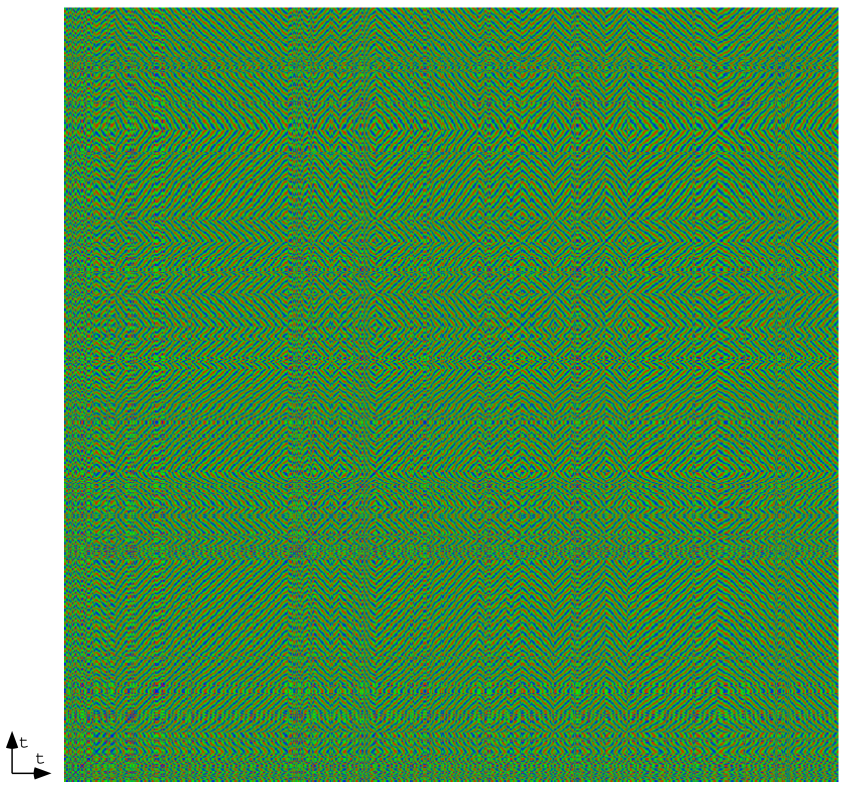


Figure 6.13: TT-π plot for the ant in the petri dish phase 3 (rubber cannula).

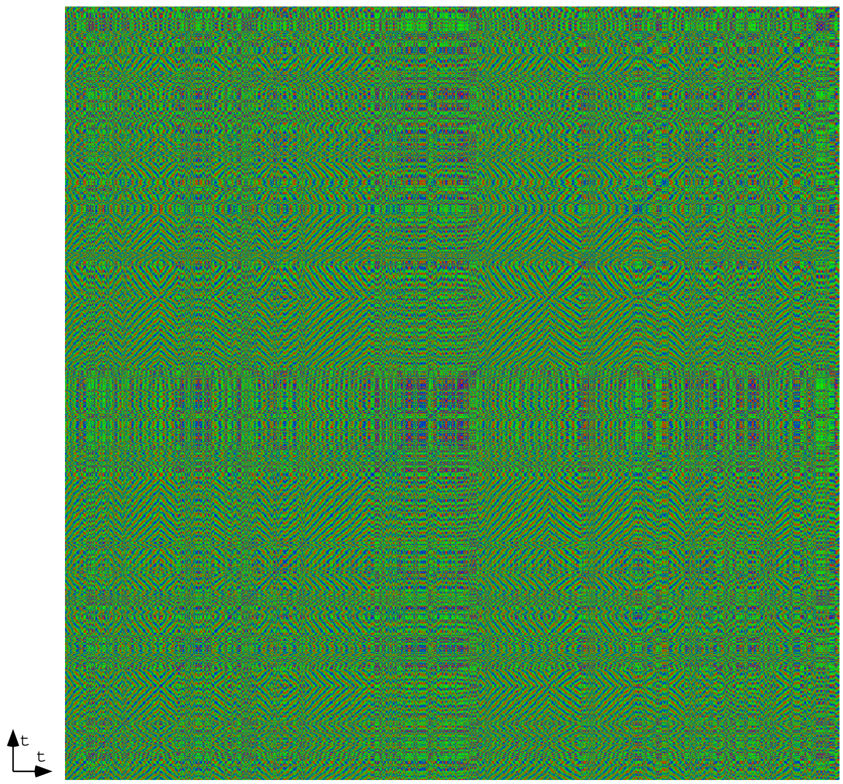


Figure 6.14: TT-π plot for the ant in the petri dish phase 4 (sugar).

spending a longer time. It was expected that the animal stayed for longer times at the location where the rubber cannula was placed. This became true, but additionaly the ant also stayed at other locations more frequently than in the previous runs. As the rubber cannula was placed in the center of the petri dish and the 'long runs' of the ant were taken place at the edge of the petri dish, it can easily be determined, which rests were located at the cannula place. They are the blue squares that have no blue bands to the left or right in the plot, as they were not passed by the animal while running around the edges of the petri dish. Astonishingly the longest two resting periods have not taken place at the cannula, but somewhere on the edge of the petri dish, both of them at the same location.

In the last run, a piece of sugar was put onto the petri dish, at some distance to the rubber cannula and the edge of the petri dish. As can be seen in figure 6.10, the movements of the ant changed dramatically. Although the regularities known from the previous runs are still visible in parts of the TT-plot, the animal used a significant time spending on one location, the piece of sugar and at other locations. The most prominent feature in this plot is the large blue square located in the center. I would like to examine that period to greater detail using the data of the TT-π plot shown in figure 6.14. First I generated the histogram of the values representing the parallelity of the movements (figure 6.15 for the whole TT-π plot. These values seem to be distributed evenly over the whole range from 0 to π. In a second step I created a histogram for the values in the middle of the TT-π plot (the large blue square in figure 6.10). This is shown in figure 6.16. Now a very distinct regularity appears. Five angles of parallelity show high values. These are: 0, $\pi/4$, $\pi/2$, $3/4\pi$ and π. Three of these values can be easily explained by the edges of the piece of sugar, which had the approximate shape of a square. The other two indicate either that the animal could see the edge lying on the diagonal of the sugar, or that it had constructed a mental map of the place knowing that it needed to run in that specific direction to get to the other corner. This could be further investigated by further subdividing this period in smaller time-frames. It is the intent of this thesis to develop new methods and not to produce biological results on animal behavior. Therefore, I will now leave this example of an ant's (techniquely limited) spatial movements.

6.2.2 Lynx

In figures 6.17 to 6.23 TT-δ plots are shown for four different lynx (Lynx lynx) using the same color scheme for distance coding (maximal distance of 70km is coded as red). The first lynx expressed a relatively simple movement (figure 6.17). It basically moved from one location (small blue square at the bottom left) to a different one (large blue square at the top right). The time in between these two phases shows an interesting movement indicated by the x-shape occurring before the second large blue square (figure 6.18). The animal was travelling, then turned around and walked back on the same path and then stopped at the location it stayed for the rest of the observation period. The lynx had passed that location once before. This can be recognized by the horizontal blue band underneath the large blue square, as it was explained in the interpretation catalogue (table 5.7).

The second lynx shows a more complicated movement (figure 6.19). It seems

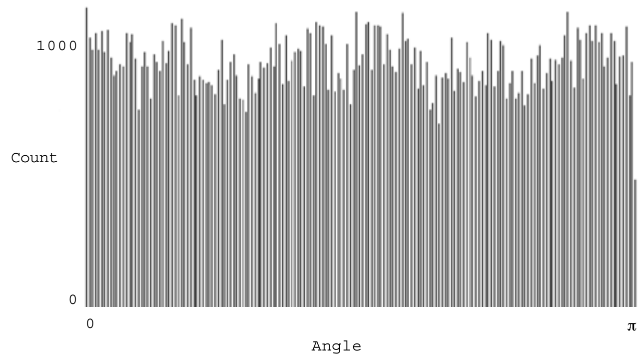


Figure 6.15: Histogram of parallelity values for the TT-π plot for the whole phase 4.

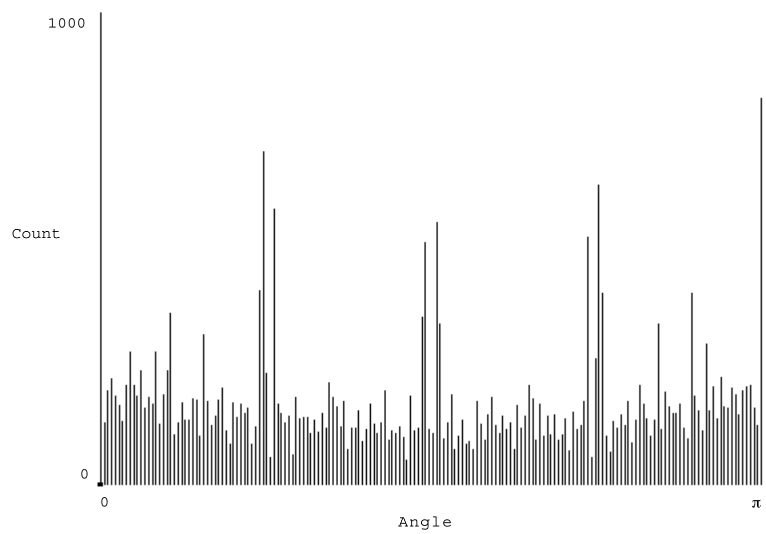


Figure 6.16: Histogram of parallelity values for the TT-π plot for the time with little movement in the middle of phase 4.

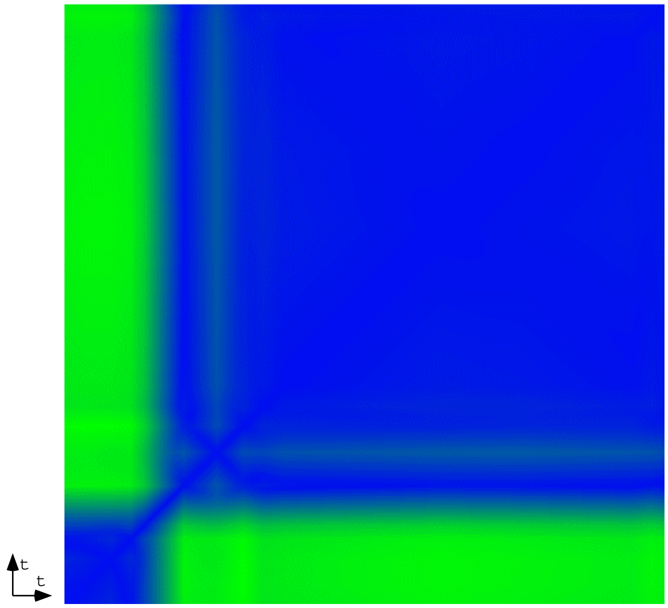


Figure 6.17: TT-δ plot of lynx 1 for a three months period in spring. Short distances are indicated in blue, medium ones in green and large distances have a yellow to red color. The same color scheme is used in the following figures.

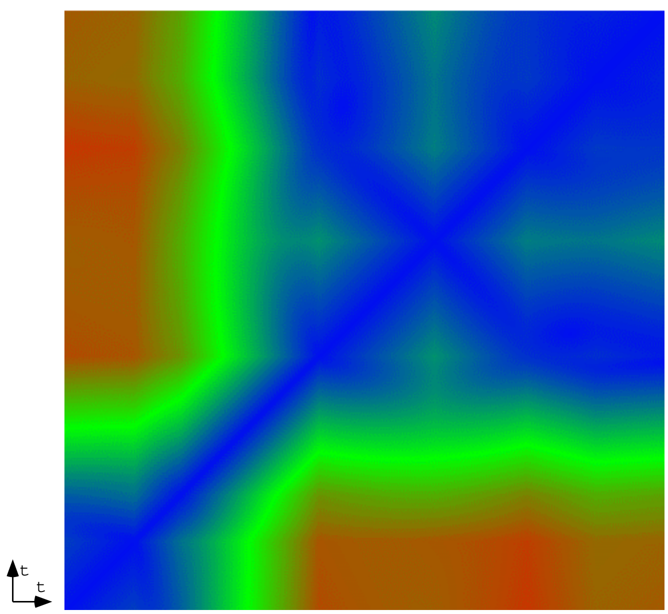


Figure 6.18: Detail of the TT-δ plot for lynx 1. The color scheme was rescaled to account for the maximum distance within this part of the observation period.

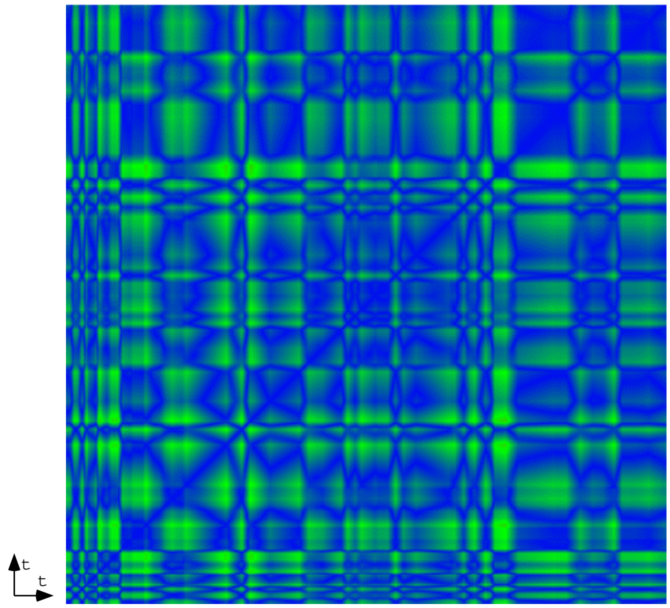


Figure 6.19: TT-δ plot of lynx 2 for a one year period.

to be exchanging between two places at a high rate, sometimes staying in these areas for some time. It often uses similar travel paths to change between the areas, which is indicated by the many blue angular lines all over the TT-δ plot. A special pattern found in different variations is shown in detail in figure 6.20. It is an x-shape surrounded by a circle. It takes some time to understand the underlying movement. It shows that the animal was walking along a single path forward, then walked in the opposite direction, back, and then repeated this movement once more on the same path. It is basically the same movement as the one illustrated in figure 5.12a. The difference is that the animal used different speeds at different times. The first and the last time the animal used the path it walked relatively fast. In between the lynx was walking at a lower pace.

In the third lynx (figure 6.21) three larger blocks can be detected. In the first part (lower left) the animal showed a similar movement as the previous animal. It is changing between two locations at relatively constant intervals. After a longer stay in one of the two areas it leaves for another region. Now the same movement pattern can be seen as in the first lynx. After walking for some time, it goes back on the same path to a place it encountered before and then stays for a longer time in that area. This behavior is repeated again in the last third of the observation period.

The last TT-δ plot with data of a lynx is shown in figure 6.22. This animal also changes between areas twice. The TT-δ plot shows again the pattern from figure 6.20, but the image is more blurred, indicating that the animal does not adhere to the pathways as strongly as the other animals did.

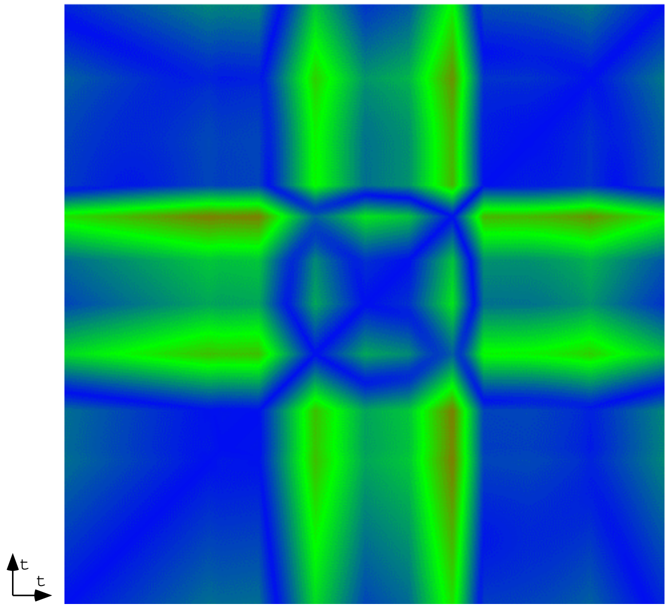


Figure 6.20: Detail of the TT-δ plot for lynx 2. The color scheme was rescaled to account for the maximum distance within this part of the observation period.

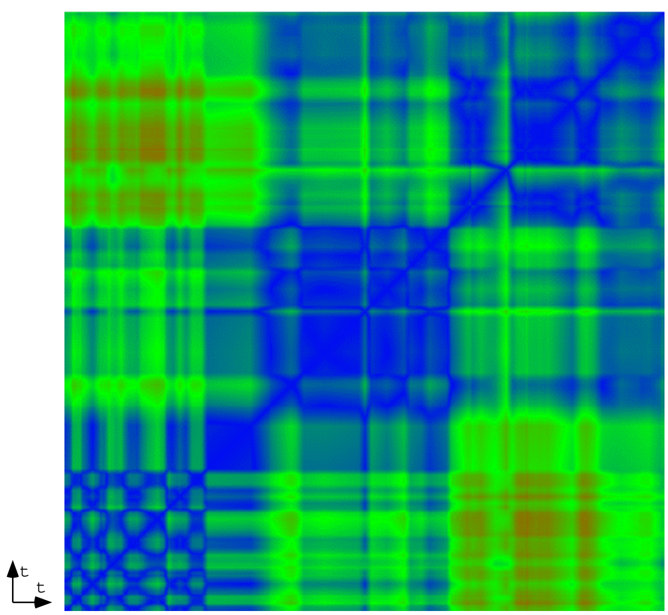


Figure 6.21: TT-δ plot of lynx 3 for a one year period.

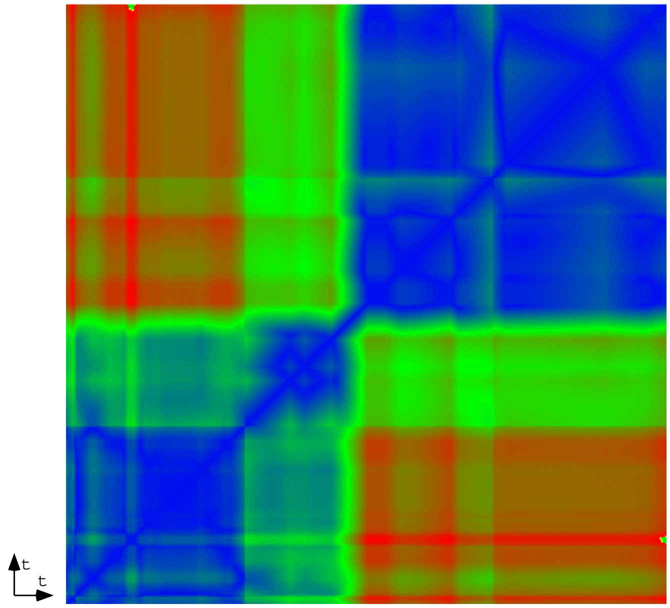


Figure 6.22: TT-δ plot of lynx 4 for a one year period.

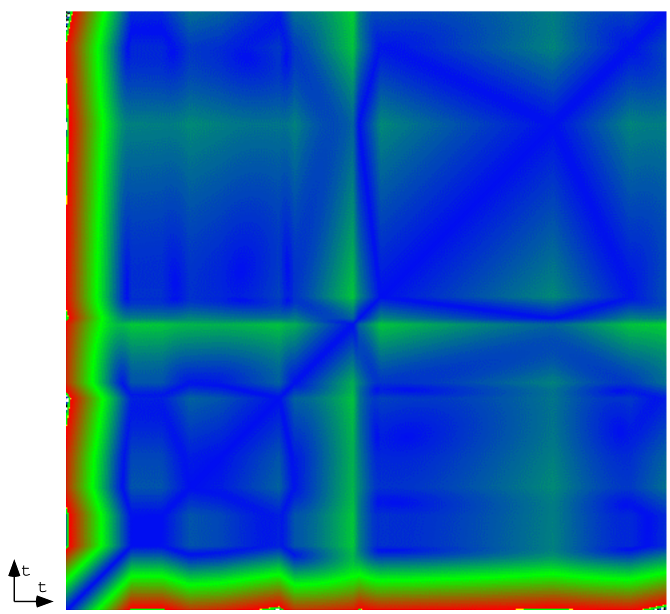


Figure 6.23: Detail of the TT-δ plot for lynx 4. The color scheme was rescaled to account for the maximum distance within this part of the observation period.

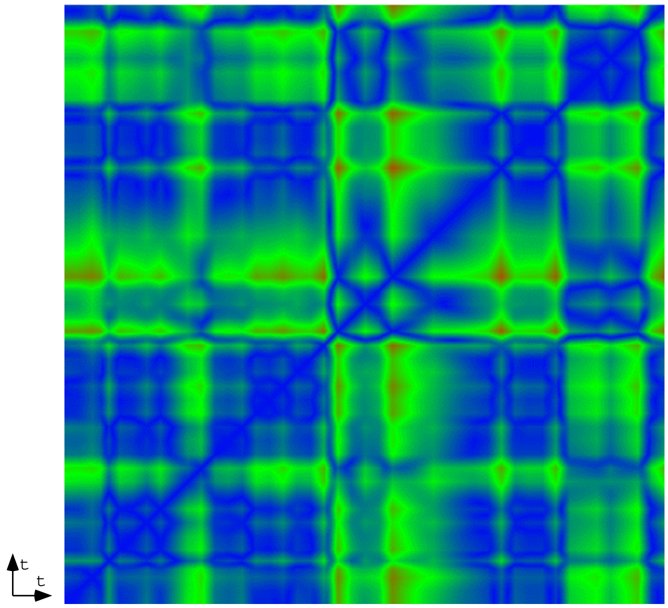


Figure 6.24: TT-δ plot for a Greater Mouse-eard bat *Myotis myotis* from Portugal.

6.2.3 Myotis myotis from Portugal

The last example of an application of TT-δ plots is from a Greater Mouse-eard bat *Myotis myotis* from Portugal (data courtesy of Ana Rainho, ICN/DHE Portugal). It shows a different pattern (figure 6.24) than the previous ones originating from lynx. The bat utilises two areas dividing the observation period in five phases of equal duration. It first flies in one area, and after a short excursion, it returns to that area again. After that in the third phase it moves to the second area showing a similar movement as the one illustrated in figure 6.20. It then returns back to the first area and finally in the last phase it goes back to the second area. The TT-δ plot appears to suggest that the bat uses known corridors as it uses similar pathways in different phases within an area.

7

Radial Distance Functions (RDF)

In the previous chapter the main focus was set on the analysis of the pure locational component ignoring any environmental attributes. In this chapter I will work on methods on how to link observations of an animal to its environment using the spatial domain.

Professor Hans Kummer at the University of Zurich once tried to explain the concept of motivation in ethology to a group of students. A plant as an immobile organism needs all resources at the same location. If one resource is missing, the plant cannot exist at that spot. Animals overcame this restriction by 'inventing' locomotion. This enabled the organisms to use resources dispersed over space and made areas accessible to them which do not contain everything at the same spot. But locomotion required another system which urges the animal to move to another location: this is called motivation. An inner mechanism to provoke the animal to do something.

By putting animal observations into one of today's GIS, the ability of locomotion is removed from the animal in most cases. In a subsequent habitat analysis this fact often vanishes in the minds of the researchers and gets replaced by an acribic application of sophisticated statistical tools. Traditionally three methods for describing the environment at an animal's location have been used. These are:

- value of parameter x (e.g. vegetation type) at this location
- distance to first object of type y (e.g. distance to nearest road)
- quantity of parameter z within a certain area around the animal (e.g. percentage of vegetation cover within 2 meters of the object)

Today these are still the standard methods. GIS technology made them very efficient, but up to now it did not generate new or enhanced methods of how to measure the environment around an animal.

There are several unsolved problems when using these description methods. In the case of measured distances often only the closest object is considered and further ones are ignored in today's habitat analysis methods. In the traditional

Figure 7.1: A cow in the middle of a road in Zurich (Switzerland). This picture illustrates three widespread problems in performing calculations with animal locations in GIS. 1. The accuracy of the determined location is often less precise than the GIS data. 2. The resources an animal needs are distributed throughout space and can hardly be determined by point in polygon tests. 3. One of the most important abilities of animals is mostly ignored in GIS: locomotion (note the concrete underneath the cow).

point pattern analysis such approaches exist for describing the clustering of a point pattern (e.g., MacLennan, 1991), but they have not been adopted for habitat analysis in animals. When describing parameters such as the amount of one vegetation type, a reasonable or 'correct' areal extent (e.g. a hectare) has to be chosen. It is widely accepted that animals interact with the environment at different scales. This makes guessing of a 'correct' diameter for such measurements hard for any researcher[1].

It is important to point out the difference between the European and American definition of the term *habitat*. In the American literature the term habitat is used as a type of surface cover class (e.g. woodland, pasture etc.). In Europe most of the time habitat refers to the sum of factors influencing a location (e.g. shrub density, altitude, soil humidity etc.). This makes a habitat analysis conducted in Europe a very tricky task involving a lot of decisions on how to measure the habitat parameters. On the American continent a habitat analysis can be performed very easily by using a point in polygon test to see in which habitat class the animal was located[2].

Nevertheless measurement of habitat parameters within a defined area around an observation remains the standard technique, sometimes enhanced by picking two different diameters. Moving to a more general level, the *configuration* of the habitat is mostly ignored, even papers calling their models 'spatially explicit' use only a fraction of the information on the spatial configuration. 'Spatially explicit' if often reduced to an inclusion of the nearest neighbor fields in calculations of gridded data (e.g., Augustin et al., 1996). Another problem that often arises in studies of wild animals but is almost never considered in their analysis is the problem of spatial autocorrelation in the environmental data. Sampling data at distances where spatial autocorrelation in the environmental parameters is still present reduces the validity of the analysis. Nevertheless this is almost never considered (e.g., Warrik and Cypher, 1998). This may also be a reason for the findings by McClean et al. (1998) that different habitat analysis methods produce contradictory results.

The human (and most probably animal) perception is absolutely brilliant in recognizing a habitat configuration by looking at it in the field. But for a scientific analysis we need to break that information apart into different pieces. This process is very hard for complex environments and cannot be done by simply looking in the field because of visibility limitations and large perspective distortions.

It is often desirable to integrate the environment around an observation for further analysis of the requirements of an animal. Minimum (spatial) resource requirements are very difficult to determine. It requires that the spatial configuration of such resources are included in such an analysis, but how?

There are different scientific fields that work on spatial structures. In landscape ecology the concepts of interdispersion, juxtaposition and fractal dimensions (Olsen et al., 1994) are used among others to describe landscape structures. They are methods to reduce the spatial information to single values which are

[1]In the analysis of populations, regression statistics is often applied using density indices, where a regular grid size with a defined cell size is used for habitat description. Here the same problems of multiple scales apply, but are ignored mostly because of a lack of datasets at different scales and the lack of knowledge about animal perception of different scales.

[2]This difference in definition is not so strong anymore, because both sides of the Atlantic start using both definitions in different situations.

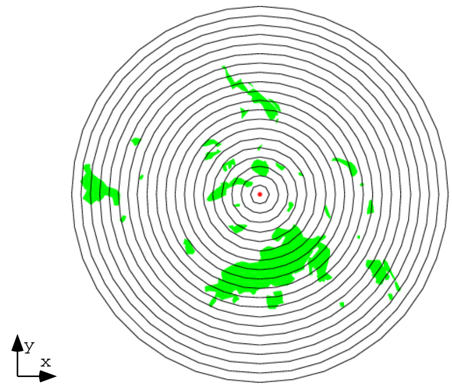


Figure 7.2: Example RDF configuration map.

often scale dependent (see chapter 1). Using these values in habitat analysis of animal requirements again raises the problem of point measurements at the location of the observation.

The aim of this chapter is to enhance the description of environmental factors. We should overcome the restriction of point measurements and start including the *spatial arrangement* of habitat features in the analysis. This may lead to a different way of analyzing the requirements of animals of their environment. It will be a lengthy and difficult process to develop scale independent, spatially explicit analysis methods. The aim of this work here is to provide a first step and a conceptual framework for further research in this direction.

In this chapter I will provide a new method for approaching these problems. As in the previous chapter the method and its extensions is explained first. In the second part two sets of biological data will illustrate their use.

7.1 Creation of RDF-Functions

The basic idea stemmed from an inspection of variograms used in kriging (section 3.3). Variograms and more often semi-variograms are used to describe spatial autocorrelation with respect to distance for estimating interpolation parameters (Cressie, 1993; Bucher, 1998). A habitat parameter description in relation to distance is exactly what is needed and used here.

I will begin with a simple observation point in a woodland area as illustrated in figure 7.2. The amount of woodland around this point can be estimated by calculating its areal extent within a certain radius around the observation.

This is done for different radii (hence the name radial distance function) until a suitable resolution and extent is achieved. Further increase in resolution would finally result in a measure of circle length at a certain radius. In a second step a graph is created using these calculated values. In figure 7.3 the *RDF-area* plot is shown. The horizontal axis represents the distance from the observation point, whereas the vertical axis shows the area (m^2) of woodland. This is the simplest form of a RDF plot which can be guessed approximately by visual inspection of figure 7.2. In the close vicinity of the observation (center) only little woodland is available. At distances around 1000m the absolute amount of

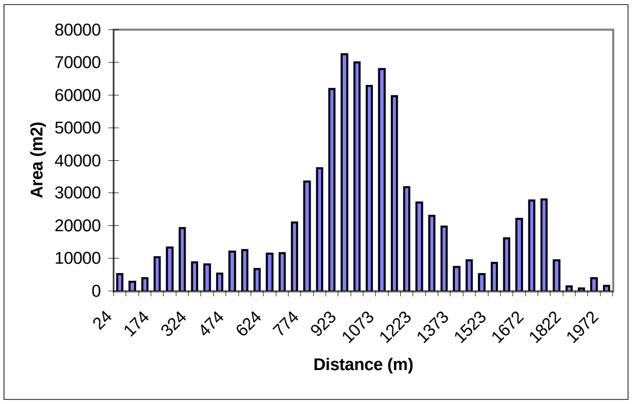


Figure 7.3: Example RDF area plot.

woodland is highest. Below and above that value only a small area of wood is present with two secondary peaks around 275 and 1750m, whereas at distances around 400-600m and 1500m local minima are present.

This procedure can be applied to a variety of measurements of the surrounding habitat in different ways. The scope of the following section is to provide an overview of the variety of aspects that can be used in RDFs and the different types of RDF-plots.

7.2 Types of RDF-Functions

The basic way in which a RDF is constructed was illustrated above. The diversity of functions used comes from the different ways in which parameters can be measured. The absolute amount of area used in the previous example is only one way to look at the surroundings. Sometimes it is more interesting to use the percentage of a parameter instead (e.g. percentage of woodland). The example from figure 7.2 is treated as such in figure 7.4, showing an *RDF-percent-area* plot. Here the picture looks quite different from the absolute RDF-area plot shown in figure 7.3. Now there are two equally high modes around 250m and 1000m reaching about 10 percent of the area at the specific distance. The third peak that occurred in the RDF-area plot has almost vanished. In this plot it can also be easily seen that the observation lies in an area with about 30 percent of wood in the closest vicinity.

For estimations of resources an animal needs the cumulative amount with respect to the distance is more useful. This is done in figure 7.5.

The cumulative area plot gives an overview of how much of a resource is available within a certain distance. In figure 7.5 it can be seen that up to a distance of about 700m only little woodland is available. From 700m to about 1200m distance a steep increase is shown and at distances over 1200m the graph again shows a relatively flat slope indicating that travel distances above 1200m provide only a small amount of additional forest areas. In a case where the observation is a nest or sett and the parameter a resource needed by the animal one would expect a large amount of travel distances between 700m and 1200m.

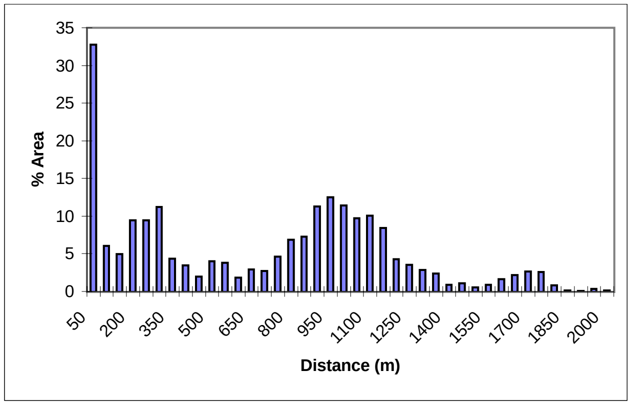


Figure 7.4: Example RDF percent area plot.

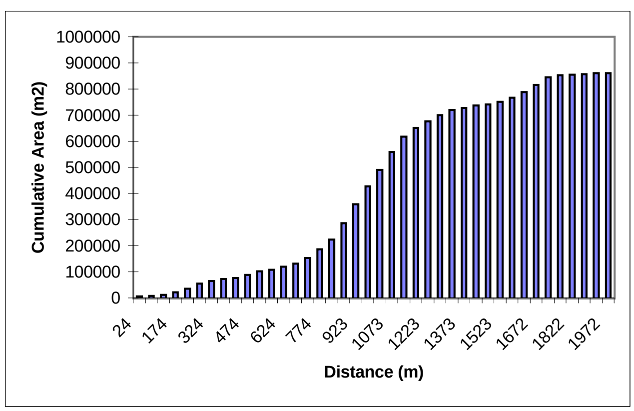


Figure 7.5: Example RDF cumulative area plot.

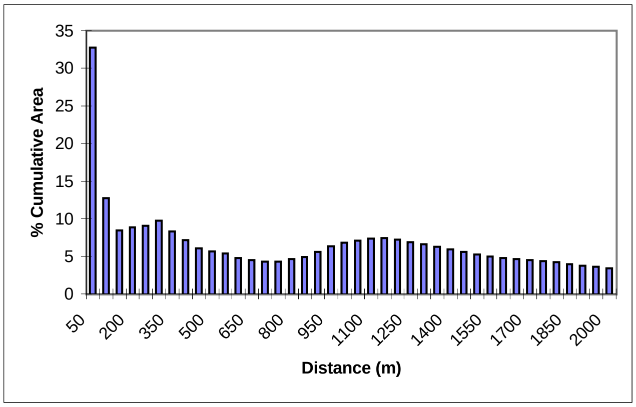


Figure 7.6: Example RDF cumulative percent area plot.

As it was the case above with the RDF-area plot, the cumulative versions can also be used in a form using percentage calculations. Figure 7.6 provides an example for the same example area used above. In the previous plots it was becoming obvious that these facts cannot be analyzed and inspected by simply eyeballing the map shown in figure 7.2.

Radial Distance Functions are not limited to polygonal structures. They can also be applied to other elements such as points and linear objects. To provide an overview of possible applications of the concept figure 7.7 shows 23 basic types of RDFs. The figure is divided into three subareas. The upper part contains aspects concerning polygonal structures, the lower right are applications to linear elements and the lower left point objects. The two classes of RDFs mentioned before are drawn in different colors. Black is used for the standard type whereas red is used for the cumulative versions. Green is used for proportional versions. Note that in principle, almost any geometric or thematic variable of relevance for the study of a particular phenomenon could be represented in RDFs.

From these basic forms of RDFs extensions and variants can easily be derived and used in the same way. I will provide more information on this in sections 7.3 and 7.6.

7.3 Example Application of RDF-Functions: Badger Setts

In this section I will present two applications of RDF-Functions. They shall illustrate their potential use.

The examples presented here were done using the locations of six main badger setts in the area of the Knonaueramt (Switzerland) from a study conducted by Emmanuel Do Linh San (San, 1997). In order to illustrate the use of RDFs the meadows were taken as object of concern. As shown by different authors (e.g., Kruuk, 1987), badgers often feed on earthworms found in meadows. Hence a RDF-area plot (figure 7.8) and the corresponding cumulative RDF-area plot

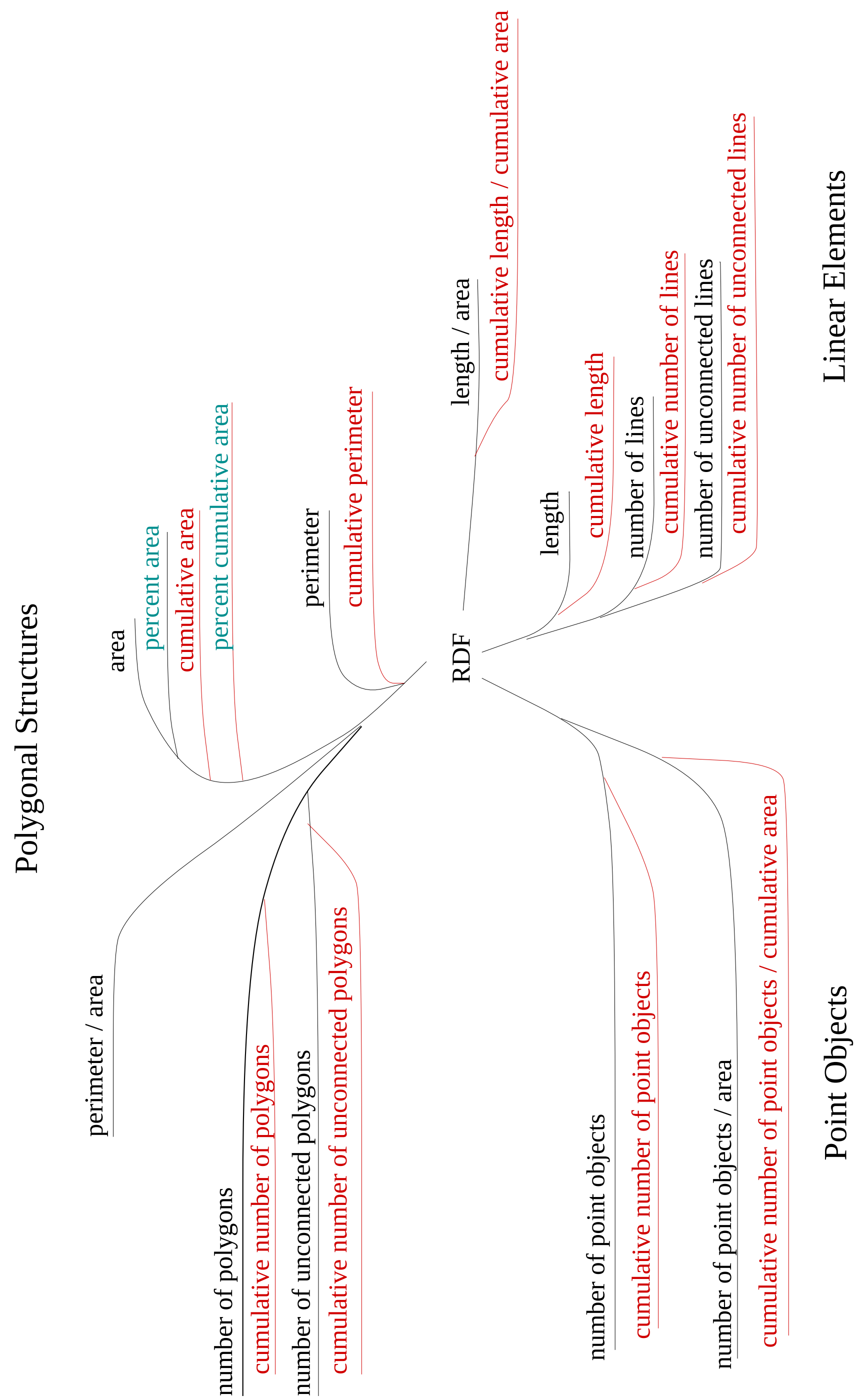


Figure 7.7: The RDF family. Sample types of RDF plots for polygonal structures, linear elements and point objects. Cumulative functions in red, percentages in green.

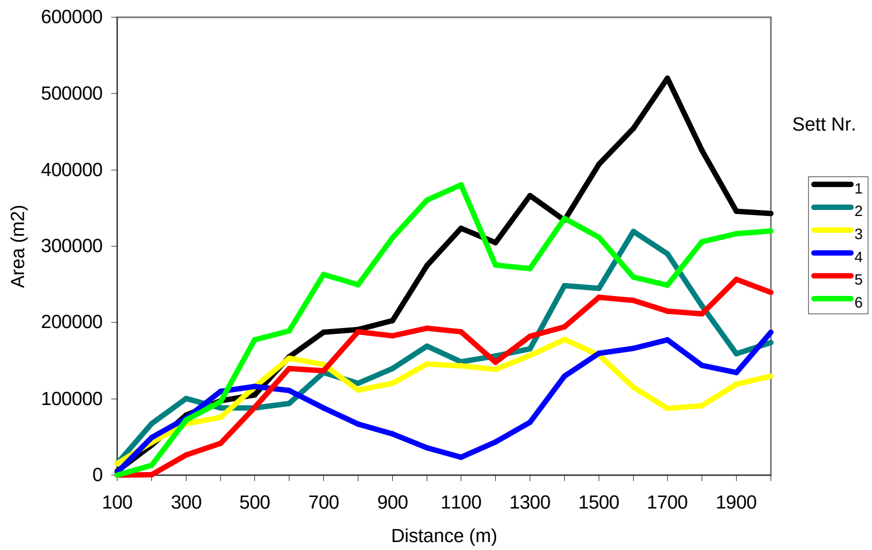


Figure 7.8: RDF for the main badger setts in the Knonaueramt area (Switzerland: amount of meadows (area in m^2) with respect to the distance from the sett.

(figure 7.9) were created for the six main setts in the area.

In figure 7.8 one can see that up to a distance of about 400m around the sett all of them except sett number 5 have comparable amounts of meadows in their vicinity. Sett number 5 has less areas with meadows within these distances, but at larger distances there is no difference compared to the other setts. A different picture is shown for sett number 4. It has a clear local minimum at distances around 1100 meters from the sett. The variation among the setts in general becomes greater at larger distances.

In this RDF plot it can be easily understood why choosing a reasonable diameter for evaluating habitat characteristics is so difficult. For each of the distances that might be used (e.g. 200m, 400m, 600m, 800m), the amount of meadows around each of the setts changes so much (200 percent and more) that completely different results can be expected in further statistical analysis steps. This situation is prevalent in many situations, but is normally neglected.

The total amount (cumulative RDF area plot) of meadows around each badger sett is shown in figure 7.9. Here we see that sett no.4 has comparable total amounts of meadows up to a distance of about 700m. After that distance only a small increase can be seen until 1300m from the sett. The total amount at a distance of 2000 meters is less than half of what is available to setts no.1 and 6.

Such findings can be used to either make predictions about the expected behavior of the badgers inhabiting these setts or helping to explain behavioral differences found in a study.

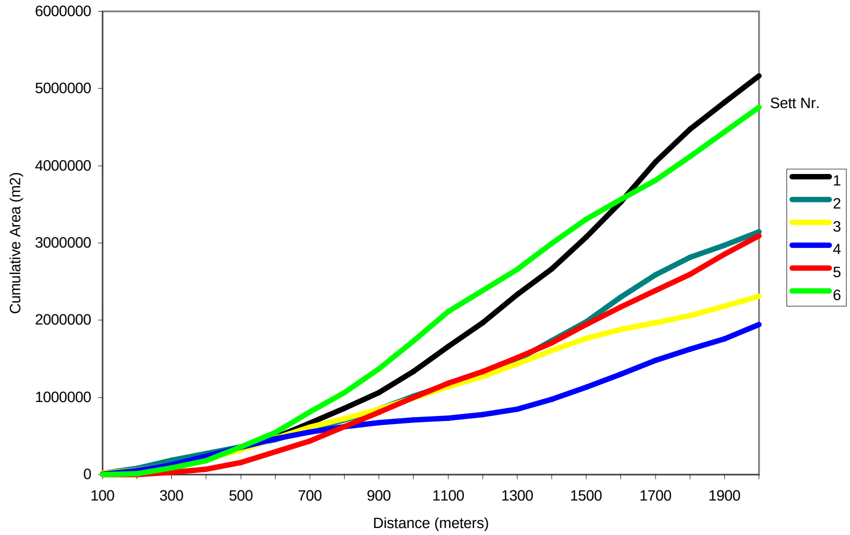


Figure 7.9: Cumulative RDF for the main badger setts in the Knonaueramt area: cumulative area of meadows in respect to the distance from the sett.

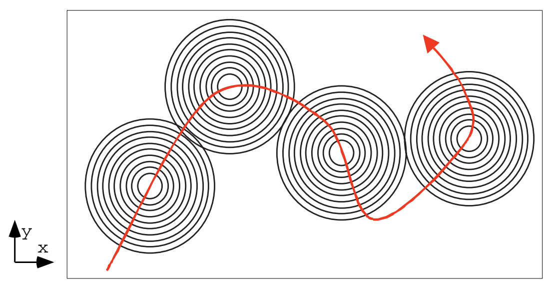


Figure 7.10: The creation of a temporal RDF-plot.

7.4 Temporal RDF-Functions

In the previous sections the radial distance functions were always used in a static context. As stated earlier it is important to include temporal aspects in the analysis when analyzing animal observations. Now I would like to extend the RDF-plots in a way that allows for the recognition of changes in the environment of an animal's surrounding. This extension is done with techniques used in the T-plots introduced earlier.

The creation is done by calculating RDFs for each location passed by the animal (see figure 7.10). In the actual implementation the path is sampled at a very fine interval and at each sampling location a standard RDF is created.

After that step all the calculations are combined into a single graph in the following way. Two axes used in the RDF plot (distance and value of parameter of interest) will be retained in the graph, although the value of the parameter will be coded in a color scheme on a z-axis instead of using the y-axis. A third

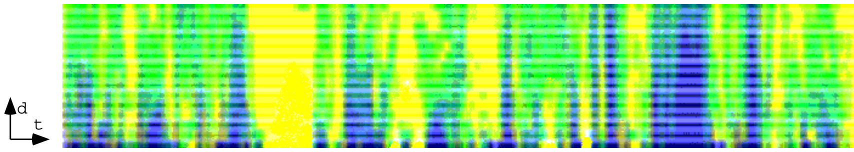


Figure 7.11: Example of a temporal RDF-plot. The amount of woodland (percent) is plotted in this example for a lynx during the observation period in 1996. Dark blue indicates high percentage of forest, green medium and yellow small percentage of forested areas around the location.

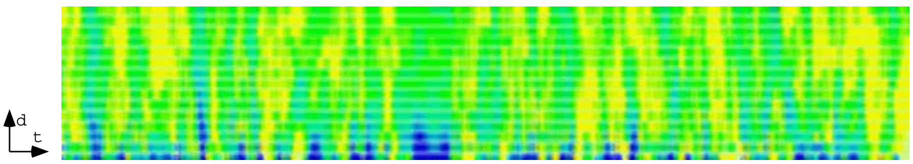


Figure 7.12: Second example of a temporal RDF-plot. The amount of woodland (percent) is plotted in this example for another lynx during the observation period in 1996. Same color scheme as in figure 7.11.

axis will contain the distance values from the sampling locations. This results in a graph having the time on the x-axis, the distance from the object in the y-axis and the value of interest (e.g. amount of woodland) coded in color on the z-axis.

An example is given in figure 7.11 for a lynx. Three distinct features can be easily seen in that figure. Firstly in the close vicinity of the animal there is most of the time a relatively high percentage of forest. At larger distances the amount of forests is usually medium to small. Secondly there were three periods where the animal stayed for a longer period of time in relatively open areas as indicated by the three larger yellow (vertical) bars in the second quarter of the figure (from left to right). The third feature can be seen in the two smaller and the large blue vertical bars in the second half of the figure. They indicate sojourns in large forested areas.

The main focus in this kind of analytical plot rests upon the temporal dynamics. It is important to see at what times and at which intervals or cycles the animal is using certain configurations in its surroundings.

In figure 7.12 a second example is illustrated. It shows another lynx living in the same region but using it in a different way than the first one. Periods of using widely open or widely forested areas are missing. In the very close vicinity it uses areas with high amounts of forests, whereas the further vicinity consists of medium forested, relatively homogeneous areas.

7.5 Directional RDF-Functions

The last extension to RDFs discussed in greater detail are the *direction dependent* RDFs. I shall explain them by directly using an example with real

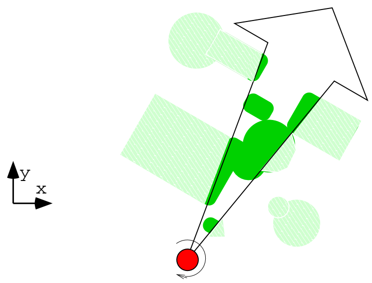


Figure 7.13: Directional RDF configurations. Left: angular configuration (slice). Right: transect configuration (transect).

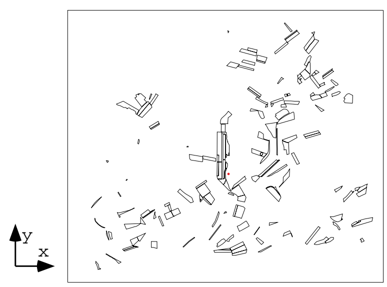


Figure 7.14: Spatial configuration of the meadows around one of the badger setts in the Knonaueramt. This data was used to calculate the following directional RDFs. The red dot indicates the sett.

data.

In the previous RDFs a circular configuration was used for calculations. This basically uses the assumption of a circular perception and use of the environment without any preferences for specific directions. Sometimes animals do not use their environment equally in all directions. Güttinger (1997) showed for example that Greater Mouse-eard bats *Myotis myotis* of one colony used the habitat around their roost preferentially in the two directions NNW and SSE.

Such findings made it clear that in some cases RDFs need to be extended by a directional component. This should allow for a (spatially) finer analysis of the environment, but it also introduces greater complexity.

There are several ways how this can be achieved. Two of them are illustrated in figure 7.13. In the case of a static RDF the calculations can be extended by dividing the surrounding area into different slices or transects for which the calculations are performed separately (figure 7.13). To be able to interpret the results, again the data have to be edited and compiled into an appropriate graphical representation due to the complexity of the data. Such a procedure was performed for one of the badger setts mentioned earlier (San, 1997) showing the relative amount of meadows in its surrounding areas.

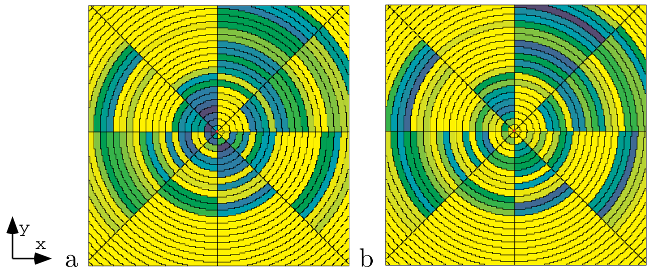


Figure 7.15: Directional RDFs of the meadows around one of the badger setts in the Knonaueramt. Left: the absolute amount of meadows within a zone. Right: the percentage of meadows within a zone. Yellow = small amount, green = medium amount, blue = large amount of meadows. The length of a side of the square is 2000m. The number of sectors used to divide the area (8 in this example) can be chosen as needed.

In figure 7.14 the meadows around one of the before mentioned badger setts are shown. The red dot in the middle indicates the sett.

The main focus here is to overcome the vague qualitative interpretations of the surrounding area. Normally we would only be able to make statements like *there seems to be more in the upper left part* or *the upper left part has a relatively small amount of meadows*. The other approach that was possible up to now is to draw a circle around the sett and then state that there are e.g. 30 percent meadows around it within 800 meters. To overcome these restrictions a much finer and much more quantitative approach needs to be applied.

From the configuration in figure 7.14 four directional RDFs have been created. In figure 7.15 the results are shown for the RDF-area (a) and RDF-percent-area (b) plots. The RDF-area plot gives a very detailed picture of the total amount of meadows around the badger sett. As it is normally the case in RDF-area plots, higher amounts are found at larger distances[3]. But here there are two exceptions to the general behavior. At medium distances in the sectors 1 and 4 (counted clockwise from the north direction) there is a zone with high amounts of meadows. The RDF-percent-area plot (shown in figure 7.15b) provides a completely different picture. The highest concentrations of meadows occur in the closer vicinity in the sectors 4 and 8.

Figure 7.16 shows the cumulative versions of the RDF-area plots. The cumulative RDF-area plot (figure 7.16: left) shows a characteristic which might result in difficulties for an animal. The directions with larger amounts of meadows in closer vicinities (4 and 8) do not or only partially correspond to the directions where large amounts of meadows are available at larger distances (sectors 1, 2 and 6). As a hypothesis this may result in larger distances which have to be traveled by the animal when compared to a configuration where closer and further areas are in the same directions. This may result in pathways describing loops or hook shapes with start and end points at the sett. This plot might be very helpful when evaluating minimum resource requirements of animals.

In figure 7.16 (right) the cumulative percentage of meadows around the bad-

[3]This problem can be avoided in directional RDFs by using the transect configuration to calculate dRDFs as illustrated in figure 7.13

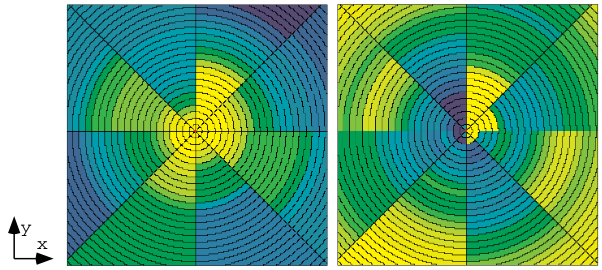


Figure 7.16: Directional RDFs of the meadows around one of the badger setts in the Knonaueramt. Left: the cumulative absolute amount of meadows within zones and sectors. Right: the cumulative percentage of meadows within a zone. Yellow = small amount, green = medium amount, blue = large amount of meadows.

ger sett is shown. The areas with high values are clearly in the directions of sectors 4 and 8 and partly sector 6. This was partly recognizable before in the RDF-percent-area plot (figure 7.15). In the example used here (meadows as feeding places for badgers) the direction of an animal is probably much influenced by the chance to find another suitable feeding patch. A second factor influencing this direction is the amount of patches already found in this direction. If the interval of finding patches declines the animal will change its direction, whereas when the interval is becoming shorter it will keep going in that direction[4]. For this aspect the cumulative RDF-percentage-area plot may be very useful. It shows the cumulative relative amount of meadows around the badger sett.

This would suggest that if travel path length is limited to this scale level directions in sectors 4, 8 and to some lesser extent in sectors 5-7 should preferably be used. This contrasts completely the intuitive rating when looking at figure 7.15 that the preferred directions would be towards the upper right and lower left part of the area.

Maybe the reader is still a little bit sceptic about the value of RDF plots. I have now introduced RDF-plots in different applications and with some examples. So I assume the reader got accustomed to reading and interpreting them in a relatively easy way. To illustrate their power I would like to ask the reader to find out which badger sett shown in figure 7.8 is corresponding to the meadows configuration presented in figure 7.14.

With these examples I would like to finish the introduction of the basic forms of radial distance functions. They were all illustrated using the areal extent as the base measurement. As it was shown in figure 7.7, the application of RDFs can be used in a variety of other measurements. Since the basic principle is well explained above and the applications with other measurements is straight forward, I would like to go on to discuss some further extensions of RDFs in the next section.

[4]There are several other factors influencing the directions used by an animal. This should only serve as an example of where the considerations could lead to.

7.6 Extensions to RDF-Functions

The applications of RDFs above were limited in two respects. First the areal extent of some type of habitat class was considered. Second the shape used in the calculations was strictly circular. These two limitations are not always necessary and extensions for both of these aspects need to be developed and applied.

The areal extent might be replaced by various forms. Many of them are indicated in figure 7.7: perimeter (edge length), number of objects or number of different habitat types are only a few examples. In a biological context of resource allocation the biomass of food is an important factor which needs to be considered.

The circular structures used above are only a first, still relatively rough attempt to quantify the environment around a location. In the context of geographic information systems refined approaches are possible and desirable. The first and relatively easy extension is the inclusion of barriers in the calculations. Not all areas are accessible to an animal. This means that the (spatially) available biomass will produce more realistic indications for understanding the way an animal uses its habitat. Going even further with the refinements the euclidian distances used in the calculations so far can be replaced by the real travel distance or even with the energy needed to access these locations. Calculations of travel distances over surfaces including barriers and the like are techniques available in good geographic information systems. Further research could lead to RDF functions which can handle multiple aspects.

In studies where the third order habitat selection of an animal, i.e. the use of habitat within homeranges (Johnson, 1980), is being investigated, the shapes used in RDF calculations can be adjusted to the shape of the homerange.

I think it now becomes clear that RDFs are a versatile framework of how to gain closer and more detailed insights into the habitat configuration around objects, even when the objects are moving as animals do.

8

Discussion and Epilogue

> *'Analyses of space-time processes require not only effective tools for multidimensional data, but also detailed spatial and temporal information'* (Miller, 1997).

Eric Miller's statement is a crucial observation for any analysis in the field of temporal phenomena. Today we are confronted with huge masses of spatial data which are in one or another way temporally referenced.

Miller implicitly mentions three parts: the data, the handling of them and their analysis. The first aspect of data acquisition is probably further developed than the remaining two. Airborne and satellite remote sensing techniques produce large datasets with a variable temporal resolution often used as input sources for GI-databases. In wildlife research large collections of data are being collected by means of radiotelemetry and especially satellite telemetry. Also in the field of animal behavior research spatial movements of animals are being recorded automatically at intervals of seconds or at even shorter intervals (e.g., Buma et al., 1998).

Some of the methods developed in this work require a high sampling intensity. If the sampling intensity is too low, spurious results may be produced. TT-plots and temporal RDFs require data sets that do not have larger gaps without any information. The present implementation within the TUPF prototype uses interpolation techniques between two observations. This has the benefit of relatively homogeneous plots. If larger gaps occurred in the data, they should be indicated as missing information in a commercial software or other production system, an effect not taken into account in the prototype application developed here.

The advances made in storing and querying such data are quite large, but still require major efforts until they become available in productive off-the-shelf systems (see section 3.4.3). What has not been recognized yet is the fact that by extending GIS with a temporal domain, new analytical methods need to be developed. Today's extensions in this context are mainly concentrating on queries (e.g., Bagg and Ryan, 1997). The discrepancy between the very high level questions asked about spatial change e.g. in global change issues and the available methods for answering such questions is still enormous.

One of the goals set up at the beginning of this work was to develop methods which allow for a better recognition and definition of biologically meaningful phases compared to the methods currently being used. All approaches established in chapters 4-7 are achieving this goal. Especially when used in combination, they provide efficient means to define such phases which need to be analyzed separately.

The conceptional change from a space centered view of the data to a time centered view is completely new in both the biological and the GIS literature body. The TT-plots are a first method performing this shift and thus require some time to become familiar with it. Once familiarized with them, they provide powerful instruments to analyse temporal changes in moving point objects by means of EDA. As this is a completely new approach, it is difficult to compare them with other methods. The only comparable methods are the ones provided by Openshaw and Perrée (1996) discussed earlier and the analytical plots used in sleep research (Borbely et al., 1981; of Psychopharmacology and Sleep Research, 1998), although the latter is used strictly with non-spatial data. In these plots the activity (active, not active) of a monitored person or animal is plotted with the x-axis representing the daytime and the y-axis representing the calendar date. The activity is indicated in black or white at the specific location to recognize temporal patterns within the data.

One of the biggest advantages of TT-plots over traditional approaches is the speed at which changes in the data can be located. The plots are computed within seconds, and make use of our most powerful way of pattern recognition, the visual perception. As an example the definition of homogeneous phases, which normally takes one or two days, can be performed within minutes (pers.comm. Prof. Dr. U. Reyer, University of Zurich).

Even though the form chosen in this work for representing the TT-plots was based on a color scheme, the reader needs to keep in mind that TT-plots consist of a matrix of calculated values such as distances or angle, which can be used for further analysis. They are not to be confounded with visual techniques such as a DTM hill-shading or the like. The calculated values are directly usable and available for further analytical steps or the calculation of derivatives. It may well be the case that other techniques are developed to visualize the data or make their interpretation easier.

The basic concept behind TT-plots is not limited to the examples presented in chapters 5 and 6. It opens up a large new field for extensions and adoptions. A wide variety of parameters and measurements can be used in the calculations of TT-plots. It is also possible that the concept could be adapted to other non-spatial problems where regularities in datasets are of interest.

The interpretation of TT-plots needs some training just like any other analysis method. The idea of creating automatic interpretation mechanisms and programs may be an approach in the future. The design of TT-plots was an open one to depict various kinds of regularities in the data. Using fixed search algorithms in computer programs for this task of finding regularities might contrast this basic idea. Maybe the application of a modified version of Openshaw's STACs (Openshaw, 1991) would allow for an appropriate machine learning process for new occurring regularities.

In chapter 4 a switch from manual, text-oriented selection of data to a graphical selection was made. In combination with effective methods to select for different temporal aspects such as the ones related to the sun or moon allows for a very efficient handling of such aspects in biological analysis. The neglect of these important factors in most previous studies concerned with wildlife animals expresses the need that such instruments for calculating and handling biologically important influences on the behaviors of animals should be available within the most capable analytical framework for such data, the GIS. Today such calculations can only be performed if professional help from astronomers or highly skilled computer technicians with knowledge in Fortran programming. This situation needs to be improved. As it becomes more and more clear that GIS will become the standard platform for analyzing data from wildlife research studies, it seems to be reasonable to stimulate and concentrate efforts for such extensions in GIS.

Bagg and Ryan (1997) recently reported on an application on historic land ownership, creating a temporal model within the Illustra database management system [1] to store and query the data. They use a similar approach as used in the TDF concept to retrieve data. Due to the application in historic changes of land ownership, it was constrained to a single linear time aspect. They also implemented different temporal rules of temporal topology (e.g. before, during, overlap, etc.), but did not include mechanisms for dealing with relative time. Time was only considered as a selection criteria and not as a separate domain for analytical purposes as in the present work.

In section 2.2 the need for 'time systems' comparable to coordinate systems in the case of space was expressed. Whenever time is recorded, the information about its reference is needed, although it is almost always forgotten to document it. This includes aspects as for example the system used (UTC, time zone, Julian Date, etc.) the actual time zone, the use of daylight saving time and others. It is an often encountered source of error that researchers forget to adjust their data for the daylight saving time in calculations. These 'time systems' are by far not as complex as the coordinate systems, but complex enough to prevent the casual GIS user to transform between them for calculations like sunset or moon rise. Especially in cases where data are exchanged it is necessary to know exactly what the time values mean.

TDF and Time-Plots are very efficient with respect to computing resources. Temporal RDFs require a lot of geometrical calculations and thus need several minutes to a few hours to be calculated, depending on the temporal resolution and the length of the observation period. In the present implementation of the RDFs all the calculations are performed using vector operations. Implementing the RDF in a grid-based environment would presumably enhance the calculating speed to a large extent. Nevertheless the computing time for temporal RDFs (and of course for the basic RDFs, too) depends strongly on the geometrical complexity of the data as well as the temporal and spatial resolutions required for the analysis in question. Considering the development of computing power over the next five or ten years, these issues probably become neglectable.

A problem that is present throughout the analysis of spatial data in wildlife research and spatial information in general is the problem of scale. Problems of generalization show up when implementing TT-plots or TRDFs. The approach used here to address these issues was a pragmatic one. Today most data collected in wildlife research are collected independently from GIS. The scale and accuracy of the data are determined by the field researcher. In the case of visualization data such as TT-plots, the possibility to zoom into the data allows the user

[1] The Illustra company and its technology was acquired by Informix. Some of the functionality was incorporated in the new Informix Universal Server software.

to browse over different scales of the data, provided the user does stay within the limits imposed by the data themselves. The TT-plots are intended for interactive usa, where this is possible. For the creation of visualizations the limits imposed by the hardware (screen resolution) are being used.

In the introduction to this thesis I provided some of the general objectives pursued in this work. I hope I was able to provide a better understanding for time and temporal aspects. The goal of developing improved and more accurate analysis methods for temporal aspects in wildlife research data within geographical information systems was achieved, even though it is clear that this can only be the start of a large field of research. I think it became evident that in the context of *temporal geographical information systems* not only sophisticated data models to represent temporal aspects are needed, but also new and improved analytical methods need to be developed to adequately analyze the data. This work concentrated on the single data type of point objects. Having seen the interesting topics involved in exploring time in such data, I assume that further developments for other data types such as lines, polygons or fields can be even more exciting, but also more complex. I think the aim to develop analytical instruments to perceive and define biologically 'meaningful' phases in a wildlife researcher's data was fully achieved.

As stated in the introduction, even though most examples and statements were made primarily with a focus on wildlife research, the methodology is of course applicable to any context concerned about the spatial movements of point objects.

A lot of research is being performed with the aim of identifying a very specific thing. The goal in this work was to open up new ways of thinking and analyzing, as my colleage Emmanuel Schmitt once stated:

> Questions should be answered in a way that more possibilities emerge.
> (E.Schmitt, personal communication)

After all, it might be a good advice to keep in mind that not everything we see necessarily needs to be true. As Couclelis (1996) states in her article about 'Geographic Illusion Systems' (GIS?)

> ... that GIS through its images, and what can be done with them, creates beliefs (mythologies, some would say) and molds habits of mind in thinking about the world unlike any that would exist without it.

Appendix A

The Prototype Application TUPF

Some of the developed methods require large amounts of calculations to be performed, thus a prototype application was programmed. It was crucial to have a high computational performance, so the following technical environment was used:

- Operating System: Unix (Sun Solaris)
- Hardware: Sun Sparc
- Programming Language: C
- GUI: X-Windows with Xt- and Motif-Toolkits
- GIS-Integration: ODE from ESRI

For portability and reasons of distribution the choice of higher level languages such as Java or Avenue were considered, but because it was necessary to achieve maximum performance the above listed environment was chosen. The hardware chosen was determined by the availability at the institute and the need for integration with a GIS (Arc/Info), which was only available on the Sun Sparc platform.

In the following I will explain the basic elements of the user interface, followed by several examples showing the capabilities of the prototype application TUPF. As it is difficult to describe dynamic methods on a static printed paper, the CD-Version of this work additionally includes several animations of the application.

The application consists of a main window and several auxiliary windows. The main window itself contains four elements (figure A.1):

- the menu bar (A)
- the map window (B)
- the selection bars to determine the standard temporal data frame to be displayed in the map window (C)
- and an area for visualisation and statistical tools (D)

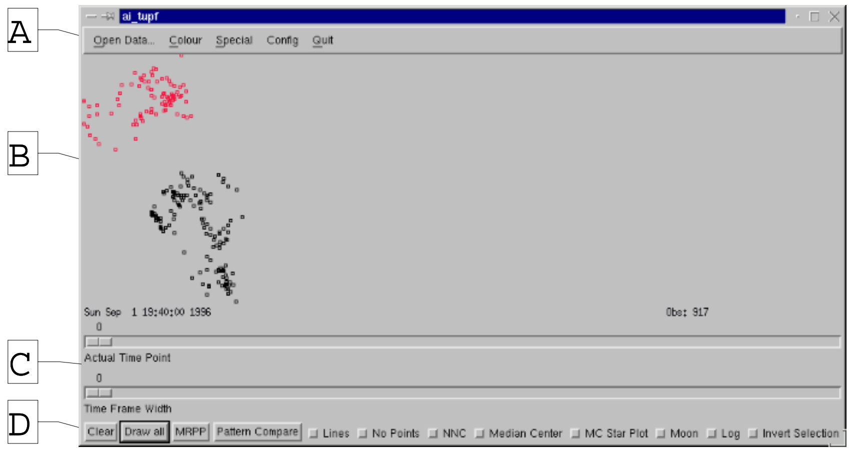


Figure A.1: The prototype application TUPF: the elements of the main window. A: menu bar, B: map window, C: selection bars to determine the standard temporal data frame to be displayed in the map window, D: visualisation and statistical tools.

As can be seen in figure A.1 a special test was rewritten and integrated into the application for a powerful comparison of two or more point distributions. It is Biondini's Multi Response Permutation Procedure which tests whether the groups originate from the same or a different distribution by exchanging observations between the groups (Biondini et al., 1988).

In figures A.2 and A.3 examples are shown for the basic selection of data. They can be refined by applying several other temporal data frames. Four sets of cyclic aspects were implemented in the prototype:

- Solar elevation
- Solar azimuth
- Lunar elevation
- Lunar illumination

As the main map window is used for georeferenced data representation, an auxiliary window is used for the calculation of the various types of plots from the *Time Plot Family*. TT-Plots and TT2-Plots can be easily explained in a static illustration. One of the Time Plots that can only be understood when seeing a dynamic example is the T-r plot. Such a plot is therefor illustrated in figure A.5.

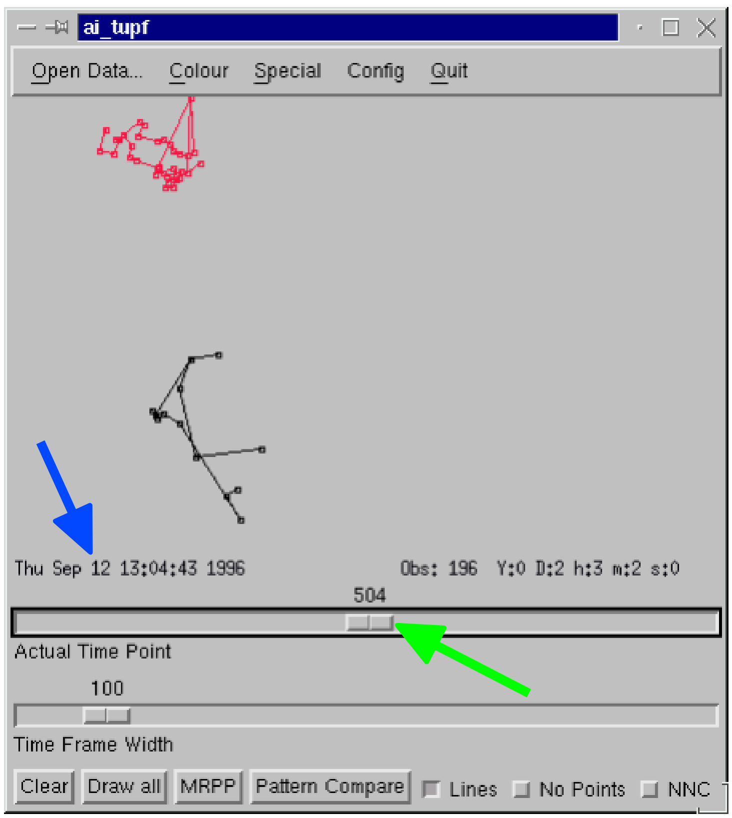


Figure A.2: The prototype application TUPF: With the scalebar indicated with the green arrow the time point of the temporal data frame can be defined and changed. The blue arrow points to the place where the actual time point (middle point of the TDF) is indicated.

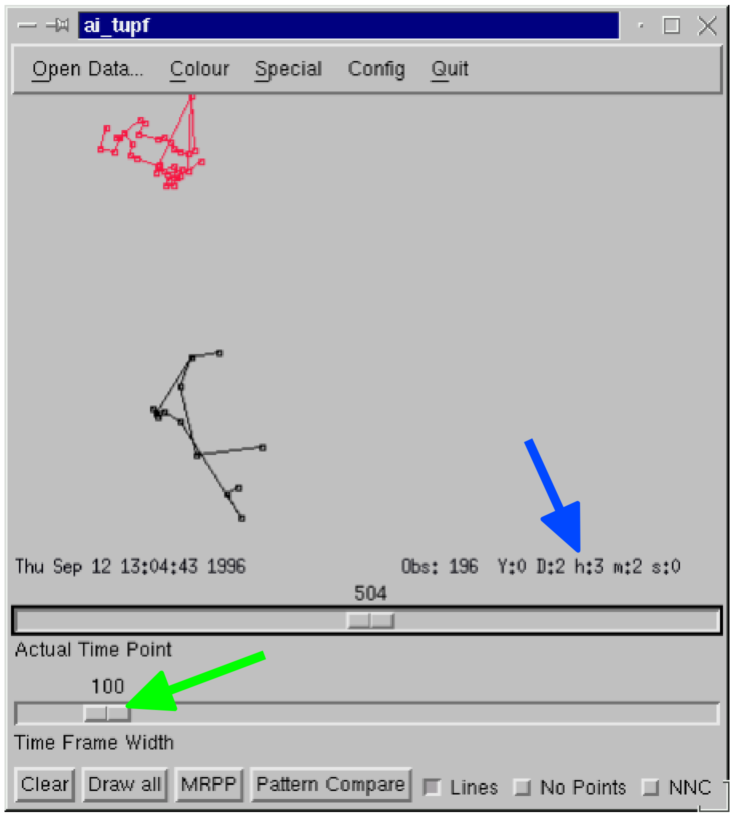


Figure A.3: The prototype application TUPF: With the scalebar indicated with the green arrow the width of the temporal data frame can be defined and changed. The blue arrow points to the place where the data frame width (in days, hours, minutes and seconds) is indicated.

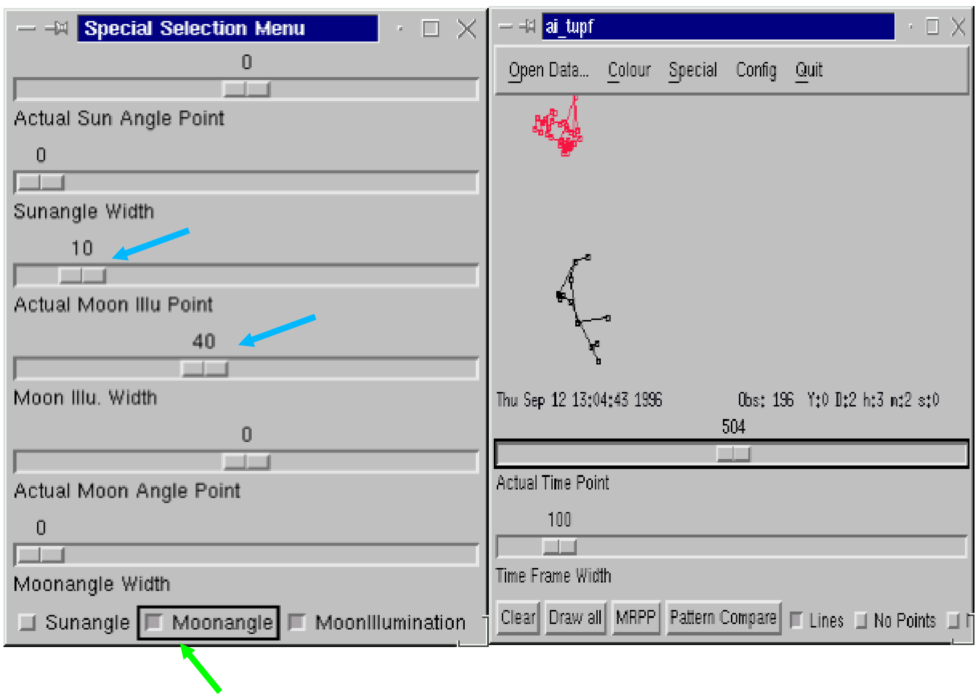


Figure A.4: The prototype application TUPF: with the auxiliary window on the left special selections of the data can be performed. The example here uses the amount of moon illumination as a selection criteria. First, the mechanism needs to be activated by pushing the button pointed to by the green arrow. Then the width of the data frame (lower blue arrow) and the actual time point (percent illumination) can be defined and changed. The window is live linked to the map window shown on the right side.

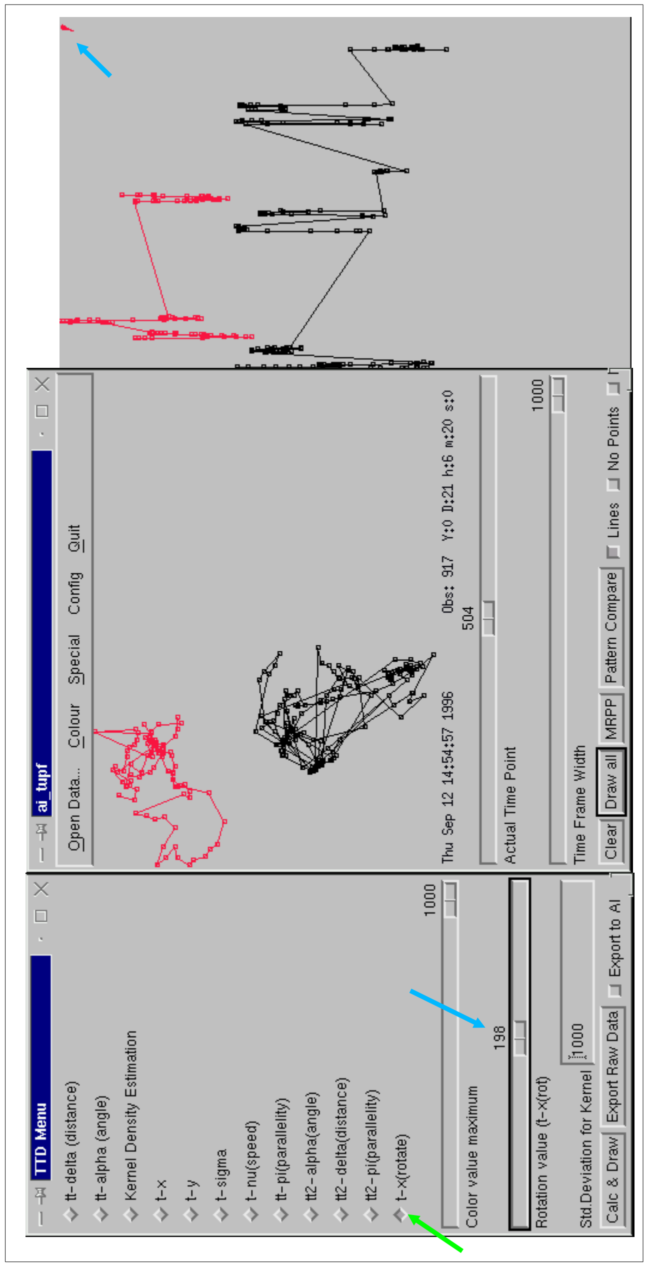


Figure A.5: The prototype application TUPF: The T-r plot is illustrated. To activate the T-r plot the appropriate button must be selected in the window on the right (green arrow). After that, the angle from which the T-r plot is calculated can be defined and changed by moving the scalebar pointed to by the blue scalebar. In the window on the right side the T-r plot is shown and updated immediately. The little arrow pointed to by the blue arrow on the right indicates graphically the rotation angle.

Appendix B

Glossary of Time

A variety of words, expressions and definitions are used in describing time and temporal aspects. The glossary provided here should help understanding them. It was partly adopted from the 'National Institute of Standards and Technology', Boulder CO.

Ambiguous time

Condition of having more than one possible value. For example, if a 24 hour clock displays a time of 15 hours, 5 minutes and 8 seconds, it is ambiguous as to the day, month, and year.

Clock

A device for maintaining and displaying time.

Coordinated Universal Time or Universal Time Coordinated (UTC)

A coordinated time scale, maintained by the 'Bureau International des Poids et Mesures (BIPM)', which forms the basis of a coordinated dissemination of standard frequencies and time signals. NOTE: A UTC clock has the same rate as a 'Temps Atomique International (TAI)' clock or international atomic time clock but differs by an integral number of seconds called leap seconds. The UTC scale is adjusted by the insertion or deletion of seconds (positive or negative leap seconds) to ensure approximate agreement with UT1 (also known as the Julian Date).

Database Time

→ Transaction Time

Date

A unique instant defined in a specified time scale. NOTE: The date can be conventionally expressed in years, months, days, hours, minutes, seconds, and fractions. Also, Julian Date (JD) and Modified Julian Date (MJD) are useful dating measures (→ Julian Date and Modified Julian Date).

DUT1

The approximate time difference between UT1 and UTC, expressed to the nearest 0.1s. DUT1 = UT1 + or - UTC. NOTE: DUT1 may be regarded as a correction to be added to UTC to obtain a better approximation to UT1. The values of DUT1 are given by the International Earth Rotation Service (IERS) in integral multiples of 0.1s.

Ephemeris Time (ET)

An astronomical time scale based on the orbital motion of the earth around the sun (→ Terrestrial Time).

Epoch

Epoch signifies the beginning of an era (or event) or the reference date of a system of measurements.

Event Time

The instant at which an event occurred.

GPS Time

GPS Time is measured in weeks and seconds from 24:00:00, January 5, 1980 and is steered to within one microsecond of UTC. GPS Time has no leap seconds and is ahead of UTC by several seconds. Time in Universal Coordinated Time (UTC) is computed from GPS Time using the UTC correction parameters sent as part of the navigation data bits. At the transition between 23:59:59 UTC on December 31, 1998 and 00:00:00 UTC on January 1, 1999, UTC was retarded by one-second. GPS Time is now ahead of UTC by 13 seconds. The ten-bit Week Number parameter can only represent integer week values from 0 to 1023. At 00:00:00 GPS Time on August 22, 1999 the transmitted Week Number parameter changed from 1023 (1111111111 in bits) to 0 (0000000000 in bits). The rollover moment took place on August 21 at 23:59:47 UTC.

Greenwich Mean Time (GMT)

A 24 Hour system based on mean Solar time plus 12 hours at Greenwich, England. Greenwich Mean Time can be considered approximately equivalent to Coordinated Universal Time (UTC), which is broadcasted from all standard time and frequency radio stations. However, GMT is now obsolete and has been replaced by UTC.

Instant

A specific time.

International Atomic Time or Temps Atomique International (TAI)

An atomic time scale based on data from a worldwide set of atomic clocks. It is the internationally agreed upon time reference conforming to the definition of the second, the fundamental unit of atomic time in the 'International System of Units (SI)'. It is defined as the duration of 9 192 631 770 periods of the radiation corresponding to the transition between two hyperfine levels of the ground state of the cesium - 133 atom. The TAI is maintained by the 'Bureau International des Poids et Mesures (BIPM)' in France. Although TAI was officially introduced in January 1972, it has been available since July 1955. Its epoch was set so that TAI was in approximate agreement with UT1 on 1 January 1958 (→ second).

Interval

The duration between two instants read on the same time scale.

Julian Day

Obtained by counting days from the starting point of noon on 1 January 4713 B.C. (Julian Day zero). One way of telling what day it is with the least possible ambiguity. NOTE: The Julian Date is conventionally referred to UT1, but may be used in other contexts, if so stated.

Julian Date (JD)

The Julian Day number followed by the fraction of the day elapsed since the preceding noon (1200 UT). Example: The date 1900 January (1) 0.5 day UT corresponds to JD = 2 415 020.

Julian Day Number (JDN)

The number of a specific day from a continuous day count having an initial origin of 1200 UT on 1 January 4713 B.C., the start of Julian day zero. Example: The day extending from 1900 January (1) 0.5 day UT to 1900 January 1.5 days UT has the number 2 415 020.

Leap second

An intentional time step of one second used to adjust UTC to ensure approximate agreement with UT1. An inserted second is called a positive leap second, and an omitted second is called a negative leap second. A positive leap second is presently needed about once per year.

Local Time

Time within a timezone.

Logical Time

→ Event time

Mean Solar Time

Mean Solar Time is simply apparent solar time corrected for the effects of orbital eccentricity and the tilt of the Earth's axis relative to the ecliptic plane; that is, corrected by the equation of time which is defined as the hour angle of the true Sun minus the hour angle of the mean Sun.

Modified Julian Day (MJD)

Equal to the Julian day. Shifted so its origin occurs at midnight on 17 November 1858. The MJD differs from the Julian date by exactly 2 400 000.5 days.

Modified Julian Date (MJD)

Julian date less 2 400 000.5

Period

A duration of time defined by two instants.

Phase

A measure of a fraction of the period of a repetitive phenomenon, measured with respect to some distinguishable feature of the phenomenon itself.

Physical Time

→ Transaction Time

Proper time

The local time, as indicated by an ideal clock, in a relativistic sense. NOTE: Proper time is distinguished from a coordinated time which involves theory and computations. If a time scale is realized according to the proper time concept, it is called a proper time scale. Examples (a) for proper time: the second is defined in the proper time of the cesium atom; (b) for proper time scale: a time scale is produced in a laboratory and not transmitted outside the laboratory.

Second

A second is the time it takes for 9,192,631,770 cycles of a frequency that resonates with an atom of cesium. It is the basic unit of time or time interval in the International System of Units (SI) which is equal to 9 192 631 770 periods of radiation corresponding to the transition between the two hyperfine levels of the ground state of cesium-133 as defined at the 1967 Conference Generale des Poids et Mesures.

Sidereal time

The measure of time defined by the apparent diurnal motion of the vernal equinox; hence, a measure of the rotation of the Earth with respect to the reference frame that is related to the stars rather than the sun. Two types of sidereal time are used in astronomy: mean sidereal time and apparent sidereal time. One sidereal day is equal to about 23 hours, 56 minutes, and 4.090 seconds of mean solar time. Also, 366.2422 mean sidereal days equal 365.2422 mean solar days.

Standard Time

Terrestrial Time (TT)

The new 1991 International Astronomical Union replacement for what was once called Ephemeris Time. On 1 January 1997, TT = TAI + 32.184 seconds, and the length of the second is chosen so that it agrees with the International Second (SI) on the geoid. The TT scale differs from the old Ephemeris Time in its conceptual definition. Practically, however, it is realized by means of International Atomic Time (TAI).

Time Range

→ Period

Transaction Time

The instant when a database record was updated.

Truncated Julian Day (TJD)

The JDN 2 440 000.5 occurred on 24 May 1968 and defines the origin of the TJD time scale used in the PB5 time code. NOTE: The TJD is used by the scientific community for recording astronomical and historical events and for archival data storage and is useful in the space sciences area. The TJD has an epoch of 24 May 1969 with a repetition period (recycle time) of 10000 days (27.379 years) and recycled on 9 October 1995. The TJD is currently equal to MJD minus 50000. TJD = MJD truncated to four digits.

Universal Time (UT)

Universal Time is the mean solar time of the prime meridian plus 12 hours, determined by measuring the angular position of the Earth about its axis. The UT is sometimes designated GMT, but this designation should be avoided. Communicators use the designation (Z) or (Zulu). Time-keepers should use UTC of the national standard, for example, UTC(USNO) rather than GMT.

Universal Time Coordinated (UTC) or Coordinated Universal Time

→ Coordinated Universal Time

Universal Time (UT) Family

Universal Time (UT) is the general designation of time scales based on the rotation of the Earth. In applications in which a precision of a few tenths of a second cannot be tolerated, it is necessary to specify the form of UT such as UT1 which is directly related to polar motion and is proportional to the rotation of the Earth in space. The UT1 is further corrected empirically for annual and semiannual variations in the rotation rate of the earth to obtain UT2.

UT0

UT0 measures UT with respect to the observers meridian (position on earth) which varies because of polar motion.

UTC

→ Universal Time Coordinated

Valid Time

The time period at which a database entry is valid, i.e. the factum is true in the real world.

World Time

→ Event Time

Bibliography

Aebischer, N. J. and Robertson, P. A. (1993). Compositional analysis of habitat use from animal radio-tracking data. *Ecology*.

Ahn, I. and Snodgrass, R. (1986). Performance evaluation of a temporal database management system. In *Proceedings of the SIGMOD International Conference*.

Alfarmanesh, H., McLeod, D., Knapp, D., and Parker, A. (1985). An extensible object-oriented approach to databases for vlsi/cad. In *Proceedings of the Sixth International Conference on Very Large Data Bases*.

Andersen, M. (1992). Spatial analysis of two-species interactions. **Oecologia*.

Ariav, G. (1986). A temporally oriented data model. *ACM Transactions on Database Systems*.

Arlinghaus, S. L., editor (1996). *Practical Handbook of Spatial Statistics*. CRC Press.

Aspinall, R. J. and Pearson, D. M. (1995). *Innovations in GIS II*, chapter Describing and managing uncertainty of categorical maps in GIS. Taylor and Francis.

Augustin, N. H., Mugglestone, M., and Buckland, S. (1996). An autologistic model for the spatial distribution of wildlife. *Journal of Applied Ecology*.

Bagg, J. and Ryan, N. (1997). *Modelling historical change in southern Corsica: temporal GIS development using an extensible database system*, chapter 3. Taylor and Francis.

Beer, R. (1990). *Intelligence as Adaptive Behaviour: An experiment in Computational Neuroenthology.* Academic Press.

Beller, A., Giblin, T., Le, K., Litz, S., Kittel, T., and Schimel, D. (1991). Temporal gis prototype for global change research. In *GIS/LIS 1991 Proceedings*, volume 2, pages 752–765.

Bertin, J. (1977). *La Graphique et le Traitement Graphique de l'Information.* Flammarion, Paris.

Biondini, M. E., Mielke, P. W., and Redente, E. F. (1988). Permutation techniques based on euclidian analysis spaces: a new and powerful statistical method for ecological research. *Coenoses*.

Borbely, A., Neuhaus, H., Mattmann, P., and Waser, P. (1981). [long term recording of motor activity: its use in research and clinical situations]. *Schweiz Med Wochenschr*, 111(23):730–5.

Borbely, A. and Tobler, I. (1996). Sleep regulation: relation to photoperiod, sleep duration, waking activity, and torpor. *Prog Brain Res.*

Bucher, F. (1998). *Regionen-basierte Wahl und Parametrisierung von Interpolationsmodellen*. PhD thesis, Geographical Institute, University of Zurich.

Buma, M., Moskal, J., and Liang, D. (1998). Ethovision multi-pro: improved animal identification during automatic multi-object tracking. In *Measuring Behavior '98*, pages 103–104, Groningen.

Burrough, P. A. and Frank, A. U., editors (1996). *Geographic objects with indeterminate boundaries*. Taylor and Francis.

Camargo, C., Visconti, M., and Castrucci, A. (1999). Physiological color change in the bullfrog, rana catesbeiana. *J Exp Zool*.

Carpenter, S. and Chaney, J. (1983). Scale of spatial pattern: four methods compared. *Vegetatio*, 53:153–160.

Cederlund, G. and Lemnell, P. (1980). Activity recording of radio-tagged animals. *Biotelemtry Patient Monitoring*.

Cole, J. P. and King, C. A. M. (1968). *Quantitative Geography*. Wiley, London.

Couclelis, H. (1996). Geographic illusion systems: Towards a (very partial) research agenda for gis in the information age. Position Paper, NCGIA Initiative 19. Specialist Meeting on GIS and Society, http://www.geo.wvu.edu/i19/papers/couclelis.html.

Cressie, N. A. C. (1993). *Statistics for spatial data*. Wiley.

David, F. and Moore, P. (1954). Notes on contagious distributions in plant populations. *Annals of Botany*, 18:47–53.

de Bono, E. (1972). *Laterales Denken für Führungskräfte*. Rowolt, Reinbek bei Hamburg.

Diggle, P. J. (1983). *Statistical Analysis of Spatial Point Pattern*. Academic Press.

Dixon, K. and Chapman, J. (1980). Harmonic mean measure of animal activity areas. *Ecology*.

Doguwa, S. I. and Upton, G. J. G. (1989). Simulations to determine the mean and variance of the point-object analogue of the clark-evans statistic. *Biometrical Journal*.

Douglas, J. (1975). Clustering and aggregation. *Sankhyā*, 37:398–417.

Dunn, J. and Gipson, P. (1977). Analysis of radio telemetry data in studies of home range. *Biometrics*.

Evans, P. J. and Evans, F. C. (1954). Distance to nearest neighbour as a measure of spatial relations in populations. *Ecology*.

Fisher, R., Thornton, H., and Mackenzie, W. (1922). The accuracy of the plating method of estimating the density of bacterial populations. *Annals of Applied Biology*, 9:325–359.

Gadia, S. (1988). Homogeneous relational model and query languages for temporal databases. *ACM Trans. on Database Systems*.

Gatrell, A. and Rowlingson, B. (1994). *Spatial point process modelling in a GIS environment*. Taylor and Francis, London.

Getis, A. (1964). Temporal land use pattern analysis with the use of nearest neighbour and quadrat methods. *Anns Ass. Am. Geogrs*.

Getis, A. (1984). Interaction modeling using second-order analysis. *Environment and Planning A*.

Getis, A. and Franklin, J. (1987). Second-order neighborhood analysis of mapped point patterns. **Ecology*.

Good, T. (1997). (badger setts in the sihlwald). Master's thesis, Zool.Inst. Univ. ZH.

Goodchild, M., Haining, R., and Wise, S. (1991). Integrating gis and spatial data analysis: problems and possibilities. *International Journal of Geographical Information Systems*.

GRASS (1993). *GRASS 4.1 User's Reference Manual*. USACERL.

Greig-Smith, P. (1952). The use of random and contiguous quadrats in the study of the structure of plant communities. *Ann. Bot.*

Güting, R. H., Bḧlen, M. H., Erwig, M., Jensen, C. S., Lorentzos, N. A., Schneider, M., and Vazirgiannis, M. (1998). A foundation for representing and querying moving objects. Technical Report CH-98-3, Chorochronos Research Network, http://www.dbnet.ntua.gr/ choros/TRs/1998/3/report.ps.gz.

Güttinger, R. (1997). *Jagdhabitate des Grossen Mausohrs (Myotis myotis) in der modernen Kulturlandschaft*. BUWAL.

Haase, P. (1995). Spatial pattern analysis in ecology based on ripley's k-function: Introduction and methods of edge correction. *Journal of Vegetation Science*.

Haining, R. and Wise, S. (1991). Gis and spatial data analysis. Report on the Sheffield Workshop. Discussion Paper 11. University of Sheffield.

Hammond, R. and McCullagh, P. S. (1978). *Quantitative Techniques in Geography: an Introduction*. Clarendon.

Healey, C. G. (1997). On the use of perceptual cues and data mining for effective visualization of scientific datasets. In *Graphics Interface '97*.

Heisey, D. (1985). Analyzing selection experiments with log-linear models. *Ecology*.

Hudson, J. and Fowler, P. (1966). *The concept of Pattern in Geography.* Dept. of Geography, Univ. of Iowa.

James, F. C. and McCulloch, C. E. (1990). Multivariate analysis in ecology and systematics: panacea or pandora's box? *Annu. Rev. Ecol. Syst.*, 21:129–166.

Johnson, D. (1980). The comparison of usage and availability measurements for evaluating resource preference. *Ecology*, 61:65–71.

Kemp, K. K. (1993). Environmental modeling with gis: A strategy for dealing with spatial continuity. Technical report, NCGIA.

Kemp, Z. and Kowalczyk, A. (1994). *Innovations in GIS 1*, chapter 7. Incorporating the temporal dimension in a GIS. Taylor and Francis.

Kienast, F., Frank, C., and Leu, R. (1991). Analyse raum-zeitlicher Daten mit einem Geographischen Informationssystem. Technical report, Berichte der Eidgenoessischen Forschungsanstalt für Wald, Schnee und Landschaft.

Kim, J., Yoo, H., and Lee, Y. (1990). Design and implementation of a temporal query language with abstract time. *Information Systems.*

King, L. J. (1962). *Statistical Analysis in Geography.* Englewood Cliffs: Prentice-Hall.

Knox, E. (1964). The detection of space-time interactions. *Applied Statistics.*

Korte, G. B. (1997). *The GIS Book.* Onword Press, Santa Fe, NM.

Kruuk, H. (1987). Foraging and spatial organisation of the european badger (meles mesls). *Behav. Ecol. Sociobiol.*, 4:75–89.

Kurt, F. (1991). *Das Reh in der Kulturlandschaft: Sozialverhalten und Oekologie eines Anpassers.* Parey.

Langran, G. (1988). Temporal gis design tradeoffs. pages 890–899, Falls Church, VA (ACSM).

Langran, G. (1990). Tracing temporal information in an automated nautical charting system. *Cartography and GIS.*

Langran, G. (1992a). Preparing for a temporal forest gis. *The Compiler.*

Langran, G. (1992b). *Time in Geographic Information Systems.* Taylor and Francis.

Langran, G. (1993). Issues of implementing a spatiotemporal system. **Int. J. Geogr. Information Systems.*

Langton, C. (1989). *Artificial Life.* Addison-Wesley.

Legrendre, P. and Fortin, M.-J. (1989). Spatial pattern and ecological analysis. **Vegetatio.*

Lippert-Stephan, E.-M. (1996). Exploratory visualization in environmental data analysis.

Lloyd, M. (1967). Mean crowding. *Journal of Animal Ecology*, 36:1–30.

London, M. G. (1997). Time and the mahabharata. http://www.thefirmament.com/Skeets/Research/Papers/mahabharata.html.

Lum, V., Dadam, P., Erbe, R., Guenaver, J., Pistor, P., Walch, G., Werner, H., and Woodfill, J. (1984). Designing dbms support for the temporal dimension. In *Proceedings SIGMOD'84 Conference*, pages 115–130.

MacLennan, M. J. e. (1991). The use of a geographic information system for second-order analysis of spatial point patterns.

McClean, S. A., Rumble, M. A., King, R. M., and Baker, W. L. (1998). Evaluation of resource selection methods with different definitions of availability. *Journal of Wildlife Management*.

Mead, R. (1974). A test for spatial pattern at several scales using data from a grid of contigous quadrats. *Biometrics*.

Melton, J. and Simon, A. (1993). *Understanding the New SQL: A Complete Guide*. Morgan Kaufmann.

Miller, E. J. (1997). *Towards a 4D GIS: Four-dimensional interpolation utilizing kriging*, chapter 13. Taylor and Francis.

Morisita, M. (1959). Measuring of the dispersion and analysis of distribution patterns. *Memoires of the Faculty of Science, Kyushu University, Series E. Biology.*, 2:215–235.

Moutka, T. and Penttinen, A. (1994). Detecting small-scale spatial patterns in lotic predator-prey relationships: Statistical methods and a case study. **Can. J. Fish. Aquat. Sci.*

Neft, D. (1966). *Statistical Analysis for Areal Distributions*. Regional science research Institute.

Neu, C. W., Byers, C. R., and Peek, J. M. (1974). A technique for analysis of utilization-availability data. *J. Wildl. Manage.*

of Psychopharmacology, S. and Sleep Research, Pharmacological Institute, U. o. Z. (1998). Rest-activity cycle recorded for years. http://www.unizh.ch/phar/sleepcd/demo/rhythms/chap3/ab0033.htm.

O'Kelly, M. (1994). *Spatial Analysis and GIS.*, chapter Spatial analysis and GIS. Taylor and Francis, London.

Olsen, E., Ramsey, R., and Winn, D. S. (1994). A modified fractal dimension as a measure of landscape diversity. A Modified Fractal Dimension as a Measure of Landscape Diversity.

Openshaw, S. (1987). A mark 1 geographical analysis machine for the automated analysis of point data sets. *International Journal of GIS*.

Openshaw, S. (1991). *Handling Geographical Information.*, chapter A spatial analysis research agenda., pages 18–37. Longman, London.

Openshaw, S. (1994). *Spatial Analysis and GIS.*, chapter Exploratory space-time-attribute pattern analysers. Taylor and Francis., London.

Openshaw, S. and Perrée, T. (1996). *Innovations ins GIS 3*, chapter User-centred intelligent spatial analysis of point data., pages 119–134. Taylor and Francis.

Perry, J. and Hewitt, M. (1991). A new index of aggregation for animal counts. *Biometrics*.

Perry, J. N. (1995). Spatial analysis by distance indices. *Journal of Animal Ecology*.

Peuquet, D. (1988). Representations of geographic space: Toward a conceptual synthesis. *Annals of the Association of American Geographers*, 78:375–394.

Peuquet, D. (1994). It's about time: A conceptual framework for the representation of temporal dynamics in geographic information systems. *Annals of the Association of American Geographers*.

Peuquet, D. and Wentz, E. (1994). An approach for time-based analysis of spatiotemporal data. In *SDH 94*, pages 489–504.

Pissinou, N., Snodgrass, R., Elmasri, R., Mumick, I., Oezsu, M., Pernici, B., Segev, A., Theodoulidis, B., and Dayal, U. (1994). Towards an infrastructure for temporal databases: Report of an invitational arpa/nsf workshop. In *Sigmod record*.

Ramachandran, B., MacLeod, F., and Dowers, S. (1994). Modelling temporal changes in a gis using an object-oriented approach. In *SDH 94*, pages 518–537.

Renshaw, E. and ED., F. (1984). The discription of spatial pattern using two-dimensional spectral analysis. *Vegetatio*.

Ripley, B. (1976). The second-order analysis of stationary point processes. *J.Appl. Probab.*, 13:225–266.

Ripley, B. D. (1988). *Statistical Inference for Spatial Processes*. Cambridge University Press, Cambridge.

Rowlingson, B. and Diggle, P. (1993). Splancs: Spatial point pattern analysis code in s-plus. *Computers and Geosciences*.

Samuel, M. and Garton, E. (1985). Home range: a weighted normal estimate and tests of underlying assumptions. *J. Wildl. Manage.*

San, E. D. L. (1997). Habitatwahl, Nahrungsspektrum und Sozialorganisation des Dachses (Meles meles L.) in einer offenen Kulturlandschaft des schweizerischen Mittellandes (Knonaueramt, ZH). Master's thesis, ETH Zürich.

Schmitt, D. (1996). Kopplung des Raster-GIS IDRISI mit einem Fuzzy-Klassifikationsmodell zur Modellierung eines unscharfen räumlichen Entscheidungsprozesses. In LORUP, E. and STROBL, J., editors, *IDRISI GIS 96*, number 25 in Salzburger Geographische Materialien. Selbstverlag des Instituts für Geographie der Universität Salzburg.

Seaman, J. W. and Jaeger, R. G. (1990). Statisticae dogmaticae: a critical essay on statistical practice in ecology. *Herpetologica*, 46(3):337–346.

Shepherd, I. D. H. (1995). *Innovations in GIS II*, chapter Putting time on the map: dynamic displays in data visualisation and GIS. Taylor and Francis.

Smith, D. M. (1975). *Patterns in Human Geography*. Penguin.

Snodgrass, R. (1987). The temporal query language tquel. *ACM Trans. Database Systems*.

Snodgrass, R. (1997). Response to mad-220. ANSI X3H2-97-010, ISO/IEC JTC1/SC21/WG3 DBL MAD- 2??

Snodgrass, R. and Ahn, I. (1985). A taxonomy of time in databases. In *Proceedings of the SIGMOD '85 conference*.

Snodgrass, R., Böhlen, M. H., Jensen, C. S., and Steiner, A. (1996a). Adding transaction time to sql/temporal. change proposal. ISO/IEC JTC1/SC21/WG3 DBL MAD-147r2, ANSI X3H2-96-502r1.

Snodgrass, R., Böhlen, M. H., Jensen, C. S., and Steiner, A. (1996b). Adding valid time to sql/temporal. 1996. change proposal. ISO/IEC JTC1/SC21/WG3 DBL MAD-146r2, ANSI X3H2-96-501r2.

Snodgrass, R. T., Ahn, I., Ariav, G., Batory, D., Clifford, J., Dyreson, C. E., Elmasri, R., Grandi, F., Jensen, C. S., Käfer, W., Kline, N., Leung, K. K. T. Y. C., Lorentzos, N., Roddick, J. F., Soo, A. S. M. D., and Sripada, S. M. (1994). Tsql2 language speci cation. ftp://ftp.cs.arizona.edu/tsql/tsql2/spec.pdf.

Stoms, D., Davis, F., Cogan, C., Painho, M., Duncan, B., and Scepan, J. (1993). Geographic analysis of california condor sighting data. **Conservation Biology*.

Studer, R. (1986). Modeling time aspects of information systems. In *Proceeings of the International Conference on Data Engineering*.

Swihart, R. and Slade, N. (1985). Testing for independence of observations in animal movements. *Ecology*.

Tansel, A., Clifford, J., Gadia, Jajodia, S., Segev, A., and Snodgrass, R. E. (1993). Temporal databases: Theory, design, and implementation. *Database Systems and Applications Series. Benjamin/Cummings, Redwood City,CA.*

Tobler, W. (1989). *Accuracy of Spatial Databases*, chapter Frame independent spatial analysis. Taylor and Francis.

Tomppo, E. (1986). Models and methods for analysing spatial patterns for trees. *Commun. Inst. Forest. Fenn.*

Turner, D. (1980). A multi-variate analysis of roe deer (capreolus capreolus) population activity. *Rev. suisse Zool.*

Upton, G. and B.Fingleton (1985). *Spatial Data Analysis by Example. Point Pattern and Quantitative Data*. John Wiley and Sons.

Vckovski, A. (1997). Interoperable and distributed geoprocessing.

Wachowicz, M. and Healey, R. G. (1994). *Innovations in GIS 1*, chapter 8. Towards temporality in GIS. Taylor and Francis.

Warrik, G. D. and Cypher, B. L. (1998). Factors affecting the spatial distribution of san joaquin kit foxes. *Journal of Wildlife Management*.

White, G. and Garrot, R. (1990). *Analysis of Wildlife Radio-Tracking Data*. Academic Press.

Winfree, A. T. (1980). *The Geometry of biological time*. Springer.

Worboys, M., editor (1994a). *Innovations in GIS I*. Taylor and Francis.

Worboys, M. F. (1994b). Unifying the spatial and temporal components of geographical information. In *SDH 94*, pages 505–517.

Worton, B. (1989). Kernel methods for estimating the utilisation distribution in home-range studies. *Ecology*.

Xingping, L. (1996). Application of la-m model in the testing of spatial pattern for insect popoulation. *Forest Research*.

Yamada, Y. (1996). On the topographical origin of some remaining snow patterns, "yukigata". In *International Snow Science Workshop 1996*.

Yearsley, C., Worboys, M., Story, P., Jayawardena, D., and Bofakos, P. (1994). *Innovations in GIS 1*, chapter 6. Computational support for spatial information handling: models and algorithms. Taylor and Francis.

Yearsley, C. M. and Worboys, M. F. (1995). *Innovations in GIS*, chapter A deductive model of planar spatio-temporal objects. Taylor and Francis.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8:338–353.

Acknowledgements

This doctoral dissertation was conducted under the supervision of Prof. Dr. K. Brassel and Dr. sc. Techn. B. Allgöwer at the University of Zurich, Geographical Institute. I would like to thank my supervisors, Prof. Dr. B. Nievergelt and Prof. Dr. R. Weibel for their encouraging ideas and comments which added considerably to the quality of this thesis.

I would like to express my sincerest thanks to a number of researchers and organisations who offered their data for testing my methodological ideas. My warm thanks go to K. Hindenlang, A. Rainho, the Woodstork and CORA projects.

I would like to express my gratitude to R. Haller who contributed to this study not only by numerous discussions and critical remarks, but also supporting me through all stages of it.

My warm thanks go to Dr. O. Wigger, who generously supported me in diverse problems with computers. In this context I would like to thank all the members of the Spatial Data Handling Division and the Wildlife Research and Conservation Biology group at the University of Zurich for their support and helpful comments during this thesis.

Last but not least I want to thank my parents for making my studies possible in the first place and E. Ramseier for supporting me in many ways during most of the last three and a half years.

Curriculum vitae

Stephan Imfeld
Born December 2, 1967 in Zug, Switzerland
Citizen of Lungern (OW), Switzerland

Education

1980 – 1987	High School in Zug, concluded with "Matura" exam type "C" .
1987 – 1990, 1992 – 1996	Studies at the University of Zurich in Zoology. Minors in Systematic Botany and Biomathematics.
1990 – 1992	System administrator in a private company.
1994 – 1995	Assistant at the Department of Geography, Spatial Data Handling Division, University of Zurich.
1996	University diploma in Zoology with the thesis "Tages- und Jahreszeitliche Verteilungsmuster des Rehs *C.capreolus* im Sihlwald" advised by Prof. Dr. B. Nievergelt.
1996 – 1998	Unix system administrator at the Department of Geography, University of Zurich.
1996 – 2000	Dissertation "Time, Points and Space - Towards a Better Analysis of Wildlife Data in GIS" advised by Prof. Dr. K. Brassel.