During the years 2013-2015 I conducted my Masters degree in ICT (Computer Information Systems) at the University of Malta. The first year of this Masters course involved taught study units whilst the second year was considered to be the research part. This part was concluded with the presentation of my final dissertation. This research was partially funded by the Master It! Scholarship Scheme (Malta). This scholarship is part-financed by the European Union – European Social Fund (ESF) under Operational Programme II – Cohesion Policy 2007-2013, “Empowering People for More Jobs and a Better Quality of Life”.

My thesis dealt with the ever changing area of databases and database systems and was entitled “Challenges of Indexing Multi-Dimensional Persistent Data”. Modern systems saw an exponential growth in the demand for high dimensional data; an example of which is spatial data. The increase in this type of data pushes for a different solution some kind of solution to be able to retrieve it efficiently compared to single dimensional data. Indexes are one of the main options which can help in efficient data retrieval.

However the selection of an appropriate index for a specific use case is a complex task and only approximate solutions can be attained. This is mainly due to the variety of factors which effect the performance of an indexing structure and its costs. These include namely data characteristics, types of queries and memory parameters.

Environmental weather data is considered as the main use case throughout this research. A domain expert from the Maltese meteorological office, identified the generic datasets required and the main operational and tactical queries involved in meteorology with a local context. This expertise provided the basis for the selection of adequate meteorological data sets and how queries need to be developed.

The analysis process involved the execution of many queries on raster and vector data. The performance of such queries was evaluated before and after indexing structures were introduced. Besides, additional adjustments such as the usage of partial or expressional indexes and other memory tweaking were taken into account. To ensure that all queries are treated equally the data server was restarted before every query.

After considering the above factors a top-down, holistic approach is adopted to select appropriate indexes for meteorological queries based on the previous analysis evaluation and a cost benefit analysis. Although query optimisation per statement might be used the procedure adopted for this research was adopted for the set of data queries as a whole. The analysis showed that particular indexes are more targeted towards particular query types. Moreover, the functions chosen when formulating an SQL query are vital as there are certain functions which do not make use of indexes. It was noted that the raster format was more suitable with restricted range locations (i.e from vector geometries). Furthermore, provided that queries are focused on particular areas, indexing a subset rather than the whole data set was considered beneficial.
Surely without the aid of the Master It! Scholarship Fund, it would not have been so easy to continue to further pursue my studies in my preferred area. Such funds surely assist students to pursue their dreams and contribute their share in enhancing the future of the Maltese society.