2. Statement of the Problem

"Cancer is really hard to go through and it’s really hard to watch someone you love go through and I know because I have been on both sides of the equation”. Cynthia Nixon on her and her mother’s battle with breast cancer.

Cancer has been recognised as a major public health problem throughout the world. According to the GLOBOCAN (2012) estimates, 14.1 million new cancer cases and 8.2 million cancer related mortality have occurred in 2012. Prevalence estimates show that, in the previous five years 32.6 million people were diagnosed with cancer. Based on these estimates, it is predicted that 19.3 million new cancer cases per year would occur by 2020 and further reported that half of all cancers and cancer deaths would proportionally increase in less developed regions of the world. In gender based cancers the reproductive and breast cancer dominates the rest in women.

Breast cancer is being diagnosed as the most common cancer in women and accounts for 25 to 31% of all cancers in women in India (NCRP-ICMR, 2013). According to ICMR- PBCR (2013) report, Bangalore city tops the chart with 36.6 new breast cancer cases for every one lakh population than the rest of the metropolitan cities in India. Bangalore has been known as the Breast Cancer Capital of India. Even a woman as young as 18 years has been diagnosed with breast cancer and cancer in young tend to be more aggressive. The rising incidence of breast cancer has been attributed to changing lifestyle and related activities. The risk of developing breast cancer for women in their lifetime in urban area is 1 in 22 and in rural area it is relatively lower being 1 in 60 (Pink India Statistics, 2014).

The information on the epidemiology of breast cancer is very limited except for few reports. There is no central cancer registry to provide comprehensive nationwide data. The existing cancer registries in India represent less than 5% of the total population and are predominantly from urban areas and majority of rural areas remain uncovered and hence would not reflect the real situation of the problem (Agarwal and Ramakant, 2008).

The challenges in the existing cancer disease surveillance systems are:
i) Improper local level cancer surveillance system that limited to only comparisons at national and global levels,

ii) Enabling users to make sense of the available data,

iii) Exploiting the richness and complexity of novel data sources at the societal scale,

iv) Creating new methods which can scale upto massive quantities of data and can integrate information from large number of data sources and

v) Making the cancer disease surveillance system more interactive and integrated.

The need of the day is a technology for targeting possible risk areas and groups. Statistical methods are used to understand the possible factors for the risk of breast cancer. Recently, GIS is used as a tool to assess and forecast areas and populations under risk. Literature review revealed that the application of GIS in the existing cancer registration system is limited in and around Mysore.

Within India there exists diversity in ethnicity, culture, religion, economy and variation in the health infrastructure and cancer services. Sporadic epidemiological studies on breast cancer do provide useful information on breast cancer incidence and care. However, major hindrance in the collection of data from the individual hospital registries leads to the lack of uniformity in data recording and storage.

The tendency and quality of the data collection changes as cancers tend to spread geographically including smaller risk zones; and as it tends to expand geometrically in terms of age groups, socio-economic status or particular geographical zones, the inclusion of newer cases from a more extensive population leads to further queries. Such situations call-for extensive data survey including environmental, geographical, socio-economical and accessibility factors to understand the expansion of cancer. Moreover, the methodologies for handling such data are unique in nature referring to the particular data type and integrating them may show several conclusions or may tend to deviate the nature of disease understanding. Sufficient and ready-to-adapt methodologies which consider the nature of the data, the expected risk and the factors responsible for the increase/decrease of risk should be devised. Additionally, user-friendly techniques that will assist latest medical advancements are also required. GIS coupled with spatial statistics and intelligent techniques are one of the state-of-art tools that help the data
managers, users and decision makers, all together in a single platform, to demonstrate the current disease distribution, estimated disease risk and predicted risk at both spatial and temporal levels and also its geometric dimensions.

The distinctive nature of the diseases, its variations based on the population and variations based on the tendency of changing variables such as age, socio-economic status, etc. should be visualised more meaningfully in a GIS environment. When variations are compared with national and international statistics, the regional variations of a particular study area should be considered as a subset and not as a separate entity. In other words, geographical scaling should be precise for national, international and regional level disease variations. Hence for these, apt statistical methods in GIS should be adopted.

Moreover, disease clusters in smaller regions and smaller populations – at-risk should not alter the disease variations when compared at the larger scales. Small area estimates and their associations tend to influence the disease rates dynamically when compared to larger areas.

Therefore methodologies and approaches chosen for a disease study, especially cancer, should not only highlight the prominent zones of disease burden, but should also evaluate the associations of other factors with cancer, display variations and show influences on different sectors of the population at risk.

Though several spatial and statistical methods are available, selected methods and spatial statistical approaches are chosen for the present study based on the data and the region of the study. Through this, it is proposed to delineate the risk based on several categories so as to improve the present disease understanding and suggest alternate measures. Therefore, keeping in view of the existence of the diverse pattern and regional uniqueness of breast cancer occurrence and paucity of published data from southern Karnataka, the present study demonstrates a more adaptable knowledge exploration of breast cancer using spatio-temporal modelling and intelligent techniques for better policy generation and decision making for control of breast cancer.

With this background information, the analytical frame work of the study is described in Figure 2.1
and the specific objectives of the present study are:

i) To study the spatio-temporal epidemiological pattern of breast cancer in southern districts of Karnataka

ii) To evaluate the use of multinomial logistic regression to assess the influence of socio-economic status and location of residence on breast cancer incidences in southern Karnataka

iii) To estimate the age adjusted breast cancer incidence rate to identify high risk age groups on a regionalised scale.

iv) To experiment the influence of spatially explicit variables on breast cancer to identify the variation of disease risk using spatio-temporal outcomes in General Linear Model; and

v) To use Poisson Probability Mapping of Breast Cancer for small estimates in a region.