Analysis of Marine Traffic – Phase 3: Predictive Accident Risk Models for Marine Traffic

Prepared by: Sara Rezaee, M.Sc., and Dr. Ronald Pelot

MARIN Risk Lab,
Dalhousie University
5269 Morris Street
Halifax, Nova Scotia B3J 1B6

March 31, 2014

PWGSC Contract Number: W7714-093795

Contract Scientific Authority: Sean Bourdon
Maritime Operational Research Team
DRDC - Centre for Operational Research and Analysis
613.992.8507

Defence Research and Development Canada
Contract Report
DRDC-RDDC-2014-C232

Inspiring Minds

Abstract

This document describes the results of Phase 3 of the Analysis of Marine Traffic (AMT) Project, conducted under contract W7714-093795. The purpose of this phase is to provide a review of geostatistical methods and make recommendations about which methods are more useful to determine the correlations between marine traffic densities (fishing and shipping) and safety incidents as well as predicting the likelihood of incidents occurring relative to traffic levels. The advantages and disadvantages of introduced methods are described and the recommended methods were applied to test cases for the purpose of comparison.

Résumé

Ce document décrit les résultats de la troisième phase du Projet d'analyse du trafic maritime (ATM), exécuté dans le cadre du contrat W7714-093795. Cette phase avait comme but de résumer des méthodes géostatistiques afin de recommander lesquelles sont les plus utiles pour déterminer les corrélations entre le trafic maritime (navires de pêche et navires de commerce) et les incidents de sécurité, ainsi que de prévoir la fréquence des incidents par rapport au volume du trafic. Les avantages et les inconvénients des méthodes sont discutés et une comparaison des méthodes recommandées est effectuée en utilisant des cas d'essai.

Executive Summary

Phase 3 of the Analysis of Marine Traffic (AMT) Project focuses on the review of spatial analysis methods to determine the relationships between safety incident occurrences and the associated traffic in Canadian waters and reveal patterns which can help to predict future risks.

Spatial analysis is a type of quantitative data analysis which includes geographically referenced data and requires analytical methods which can process geographic variables. The primary objective of spatial analysis is to reveal relationships between spatial features. Spatial analysis can draw on different forms depending on the purposes that the analyst wants to achieve through data analysis.

This study focuses on three main classes of geostatistical methods:

- Spatial Autocorrelation: Spatial autocorrelation indicates the extent to which the occurrence of one feature is influenced by similar features in the adjacent area. These methods can help to reveal the degree of spatial correlation between marine traffic densities and incident occurrences.
- Hot Spot Analysis: Hot spot analysis determines the high concentration of incidents or activity within a certain area. These methods can be extended to find the incident hot spots with respect to underlying traffic.
- Spatial Interpolation: Spatial interpolation, which includes a variety of methods, aims to predict the value of a target variable based on the available data. Methods such as Kriging and Splines are useful to interpolate traffic or incident occurrences in unsampled areas based on mathematical distributions built on sampled data. Other methods, including

Dual Kernel estimation and spatial regression, are useful for identifying patterns of incident occurrence based on the traffic amount in each specified region.

Advantages and limitations of each method are discussed and, based on the main goal of this study phase as described in the Statement of Work (Identifying spatial patterns between marine traffic densities and safety incident occurrences), Moran's I Autocorrelation Test, Risk-Adjusted Nearest Neighbour Hierarchical Clustering, Dual Kernel Estimation, and Spatial Regression are recommended as appropriate methods. These methods are then carried out using the outputs from AMT-Phase 1 for the vessel types 'fishing' and 'shipping', and the Canadian Coast Guard SISAR[1] Incident database for the year 2010.

Key findings:

- Moran's I Autocorrelation Test: Fishing incidents are positively spatially autocorrelated. The same statement is true for shipping incidents, fishing traffic, and shipping traffic as well. Fishing incidents and activities are mostly concentrated near the shoreline but shipping activities and incidents are more highly autocorrelated further than shore (between 20 nm to 200 nm from shore).
- Risk-Adjusted Nearest Neighbour Hierarchical Clustering: hot spots of fishing incidents related to fishing traffic do not dramatically change during different quarters of year 2010. Although most of the fishing incidents happen near shore, incident hot spots with respect to traffic happen farther from shore. Shipping incident clusters show that incidents are more clustered with respect to traffic in the Arctic which is due to the very small amount of traffic there. Shipping incidents in Canadian waters were re-examined excluding the Arctic region to circumvent this peculiarity.

[1] SISAR: CCG Search and Rescue Program Information Management System

- Dual Kernel estimation: The results for fishing incidents are in alignment with the results from the risk-adjusted nearest neighbourhood hierarchical clustering, where there is no significant change in spatial patterns during different quarters over the year, and the two most important hot spots occur in the same areas: one Southwest of Vancouver Island and the other North of Newfoundland and Labrador. When it comes to shipping, the probability of incidents relative to traffic decreases further from shore.
- Spatial Regression: Results indicate that North of Newfoundland and Southwest of Vancouver Island there is a strong relationship between fishing incidents and traffic which also showed up in two previous methods as well. This method is more sensitive to slight changes in incident numbers and traffic in different quarters of the year. Geographically Weighted Regression (GWR) estimations of shipping incidents based on traffic indicate the relationship decreases when the distance from shore increases, which was also the result from the Dual Kernel estimation.

Sommaire

La phase 3 du Projet d'analyse du trafic maritime (ATM) est axée sur un sommaire de méthodes d'analyse spatiale qui seraient utilisées pour déterminer les relations entre les incidents de sécurité et le trafic dans les eaux canadiennes et d'élucider les profils spatiaux qui pourraient aider à la prédiction de futurs incidents.

L'analyse spatiale est une étude quantitative de données codées suivant une grille géographique qui nécessite des méthodes analytiques capables de traiter des variables géographiques. L'analyse spatiale a comme but de découvrir les relations entre les entités spatiales. Ces analyses peuvent comprendre plusieurs techniques selon les besoins de l'analyste.

Cette étude se concentre sur trois catégories de méthodes géostatistiques :

- Autocorrélation spatiale : L'autocorrélation spatiale indique jusqu'à quel point la présence d'une caractéristique est influencée par des caractéristiques semblables à proximité immédiate. Ces méthodes peuvent servir à déterminer l'intensité de la corrélation spatiale entre les densités du trafic maritime et la présence d'incidents.
- Analyse de points chauds : L'analyse de points chauds sert à établir les zones où la fréquence des incidents est haute relatif au volume du trafic.
- Interpolation spatiale : L'interpolation spatiale inclut de nombreuses méthodes et vise d'estimer la valeur d'une variable selon les données disponibles. Des méthodes basées sur le krigeage et les courbes splinées sont utiles pour l'interpolation du trafic ou des incidents dans les zones sans données en construisant des distributions mathématiques conformes aux données observées. D'autres méthodes, incluant la méthode des noyaux

doubles et la régression spatiale, servent à identifier la distribution spatiale des incidents selon le trafic dans chaque région considérée.

Ce rapport discute des avantages et des contraintes de chaque méthode et, selon le critère principal de cette phase de l'étude tel qu'identifié dans l'énoncé de travail (Identification des relations spatiales qui existent entre le trafic maritime et les incidents de sécurité), le test d'autocorrélation « I » de Moran, l'analyse hiérarchique par grappes plus proches voisines ajustées en fonction du risque, l'estimation par la méthode des noyaux doubles, et la régression spatiale sont recommandées. Ces méthodes sont ensuite mises à l'épreuve en utilisant les données de la phase 1 du projet ATM pour les navires de pêche et de commerce ainsi que les incidents dans la base de données SISAR[2] de la Garde côtière canadienne pour l'année 2010.

Principaux résultats :

- Test d'autocorrélation « I » de Moran : Les incidents de sécurité pour les navires de pêche et les navires de commerce démontrent une autocorrélation spatiale positive. De même est vrai pour le trafic des deux genres de navires. L'activité et les incidents de sécurité pour les navires de pêche sont plutôt agglomérés dans la zone côtière, tandis que l'autocorrélation pour les navires de commerce est plus élevée dans la zone extracôtière (en particulier de 20 à 200 NM de la côte).
- Analyse hiérarchique par grappes plus proches voisines ajustées en fonction du risque : les points chauds des incidents pour les navires de pêche ne varient pas beaucoup d'une saison à l'autre pour l'année 2010. Même si la majorité des incidents ont lieu dans la zone côtière, le trafic suggère qu'il existe des points chauds plus loin de la côte. L'analyse des incidents pour les navires commerciaux indique que les incidents pour ce

[2] Système informatisé recherche et sauvetage

genre de navires sont beaucoup plus regroupés dans l'arctique, grâce à la rareté du trafic dans cette région. Pour cette raison, les incidents liés aux navires de commerces furent réexaminés en excluant l'arctique afin de contourner ce problème.

- Estimation à base de la méthode des noyaux doubles : Les résultats pour les incidents liés aux navires de pêche étaient conformes à ceux provenant de l'analyse hiérarchique par grappes plus proches voisines ajustées en fonction du risque. C'est-à-dire que les profils sont très semblables d'une saison à l'autre et les deux points chauds les plus importants se trouvent aux mêmes endroits, soit au sud-ouest de l'île de Vancouver et au nord de Terre-Neuve-et-Labrador. L'analyse démontre aussi que la probabilité d'incidents liés aux navires commerciaux (calculée en fonction du trafic) diminue en s'éloignant de la côte.
- Régression spatiale : Cette méthode indique qu'au nord de Terre-Neuve-et-Labrador et au sud-ouest de l'île de Vancouver il y a un lien important entre les incidents de navires de pêche et le trafic, tel qu'identifié avec les deux méthodes précédentes. Cette méthode est plus sensible aux changements dans le taux d'incidents et le trafic au cours des saisons de l'année 2010. Des estimations de régression géographiquement pondérées des incidents liés aux navires commerciaux basées sur le trafic démontrent que le lien s'affaiblit quand la distance de la côte augmente, un phénomène aussi observé des résultats de l'analyse avec la méthode des noyaux doubles.

Contents

List of Figures

List of Tables

List of Acronyms

AMT	Analysis of Marine Traffic
AOI	Area of Interest
ESRI	Environmental Systems Research Institute
GWR	Geographically Weighted Regression
IDW	Inverse Distance Weighting
LRIT	Long Range Identification and Tracking
SAR	Search and Rescue
SISAR	Search and Rescue Program Information Management System
VTOSS	Vessel Traffic Operations Support System

1 Introduction

The Analysis of Marine Traffic (AMT) project was initiated with an overall goal of quantifying information on the quantity and type of marine traffic data in and near Canadian waters for use by DRDC and Canadian Armed Forces in a planning environment. Specific end use purposes are best described by the text from the original Statement of Work (SOW) from DRDC:

> "Canada First Defence Strategy calls for the Canadian [Armed] Forces (CF) to work closely with federal government partners to ensure the constant monitoring of Canada's territory and all air and maritime approaches, including in the Arctic, in order to detect threats to Canadian security as early as possible. [Canadian Joint Operations Command (CJOC)] is the national military authority responsible for the conduct of all domestic operations. As such, [CJOC] will conduct operations to detect, deter, prevent, pre-empt and defeat threats and aggression aimed at Canada within its area of responsibility (AOR). The Command is also responsible for the effective operation of the federal maritime and aeronautical search and rescue (SAR) system.
>
> Defence Research and Development Canada (DRDC) Centre for Operational Research and Analysis (CORA) is conducting various research activities aiming to help Canada Command and subordinate organizations to improve how operations are planned and conducted. In order for DRDC CORA to accurately simulate marine traffic, surveillance activities, as well as SAR activities, a comprehensive understanding of marine traffic occurring inside and near Canadian waters is required. Analyses of seasonal and geographical patterns in marine traffic help to build this understanding. They also contribute to better estimate the risks of various types of undesired activities and marine incidents to occur inside the [CJOC] AOR, and will help [CJOC] to allocate resources for maximizing maritime domain awareness and operational effectiveness."

Within this overall plan, the project tasks for AMT Phase 3 are primarily focused on reviewing geostatistical methods to reveal spatial patterns and relationships between safety incident occurrences and marine traffic densities. This primary task is defined in Section 4.3 of the SOW, reiterated here:

> "The Contractor must investigate and review geostatistical analysis methods for identifying patterns in marine traffic, exploring the degree of correlation between marine traffic densities for different vessel types and activities of particular interests (e.g. safety incidents), and predicting the risk of certain activities or incidents to happen. In

> consultation with the Technical Authority (TA), the Contractor must compare the strengths and limitations of the reviewed methods, and identify those that are applicable to operational planning and the most appropriate to Canada Command's requirements.
>
> The TA will provide the Contractor with an initial set of geostatistical methods for review. Following this, the TA may request that the Contractor research and gather new or modified geostatistical analysis methods. The Contractor is not to begin gathering or researching new methods without prior written authorization from the TA."

This report focuses on the task of reviewing spatial analysis literature and recommending methods which are in alignment with AMT project purposes. The proposed methods are then carried out on the outputs from AMT-Phase 1 in combination with the Canadian Coast Guard Incidents from the year 2010 to show the effectiveness of these methods and to provide a means of comparisons between these methods. This report concludes with the potential uses for each method and recommendations for future work.

2 Scope of Analysis of Marine Traffic – AMT Phase 3

2.1 Area of Interest

The Area of Interest (AOI) for the third phase of the AMT project comprises Canadian waters covering out to 300 nautical miles from the Canadian shore in the North, and 1500 nautical miles offshore elsewhere, as indicated in Statement of Work. The resulting AOI was defined and mapped using GIS software (ArcMap) as shown in Figure Figure **2.1**.

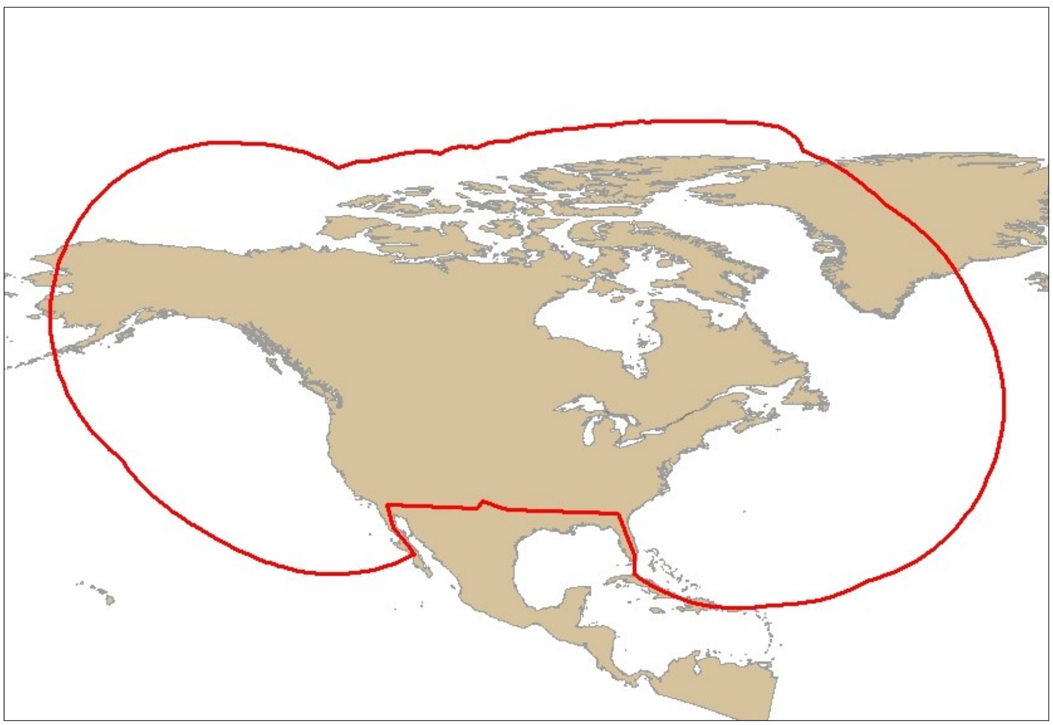

Figure 2.1. AMT-Phase 3-Area of Interest. Source: Hilliard and Pelot, 2012.

2.2 Time-Scope

The analyses were conducted for the year 2010 to be consistent with the timeframe of previous Phases of the AMT study.

2.3 Incident Data - SISAR

When a Search and Rescue (SAR) Coordination Centre receives a report of an incident, they dispatch the most available and suitable SAR resource(s) to provide assistance and a record in their SISAR incident database is created. The SISAR (Search and Rescue Program Information Management System) database includes detailed information about each incident, such as time and location, type of vessel, type of incident, severity, etc. This study is restricted to CCG maritime incidents classified as M1, M2 or M3 (see Section 6).

2.4 Activity Data - Outputs from AMT-Phase 1

Output from Phase 1 of the Analysis of Marine Traffic (AMT) project falls into three major traffic groupings−Shipping, Fishing, and Pleasure/Recreational Boating−corresponding to the data sources drawn upon and the methods required for assessment of traffic using these sources. Since original data on Pleasure/Recreational Boating Traffic data is not collected countrywide, in AMT Phase 1 a model was developed to estimate traffic level distribution using the pleasure and recreational boating incidents relative to populated places. Thus correlating the Pleasure/Recreational traffic from Phase 1 with the incident data from which it was derived is meaningless, and is consequently omitted from this study.

The first and largest grouping of traffic processed in Phase 1 comprises a set of diverse vessels, developed from traffic reporting information systems (i.e. LRIT[3] and VTOSS[4] data). This group includes merchant shipping, tankers, and government (non-military), research, passenger, and tug/service vessels. The second grouping includes fishing vessel traffic, which was modeled using fishing catch-effort records (Hilliard & Pelot, 2012).

---

[3] Long Range Identification and Tracking

[4] Vessel Traffic Operations Support System

Traffic modelling outputs are in the form of ESRI ASCII and Shapefiles providing the minimum traffic levels, instantaneous density estimates, or vessel presence for two grid resolutions at yearly and quarterly temporal aggregations. The grid resolutions are as follows:

- Uniform Size Grids: 0.5 degree grids across the AOI.
- Variable Size Grids: 0.1 degree grids up to 20 nm from shore, 0.25 degree grids from 20 nm to 200 nm from shore and finally, 0.5 degree grids for the remainder of the AOI.

The quarterly day of year bounds are shown in TableTable 2.1.

Table 2.1. Quarterly Date Bounds

Quarter	**Start**		**End**		**Days**
	Day of Year	**Date**	**Day of Year**	**Date**	
Q1	355	Dec 21	78	Mar 19	89
Q2	79	Mar 20	171	Jun 20	93
Q3	172	Jun 21	264	Sep 21	93
Q4	265	Sep 22	354	Dec 20	90

Figure 2.2 and Figure 2.3 shows the AMT Uniform and Variable Size Grids, respectively.

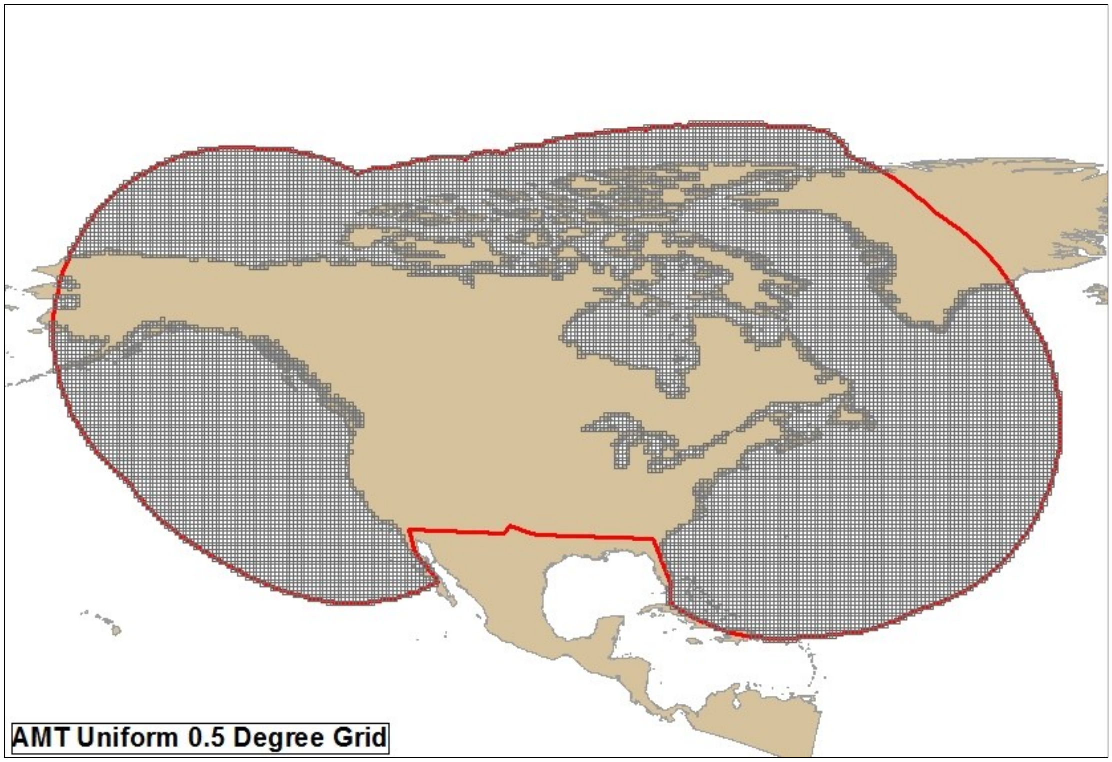

Figure 2.2. AMT Uniform 0.5 Degree Grid. Source: Hilliard and Pelot, 2012.

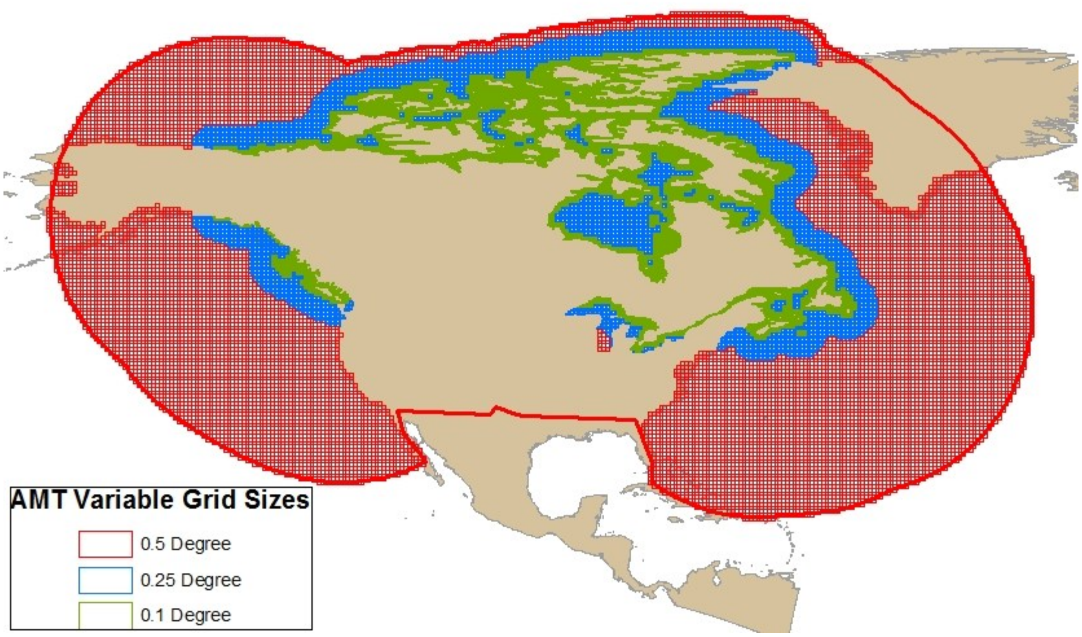

Figure 2.3. AMT Variable Grid Sizes. Source: Hilliard and Pelot, 2012.

Since this report is aimed at finding the relationship between incident occurrence and traffic, an adjustment to the AOI is needed by defining a boundary based on the Canadian Search and Rescue (SAR) Areas. Only the grids which fall within this boundary are studied since the SISAR incident data is only collected in those areas. Figure 2.4 shows the boundaries of Canadian SAR areas.

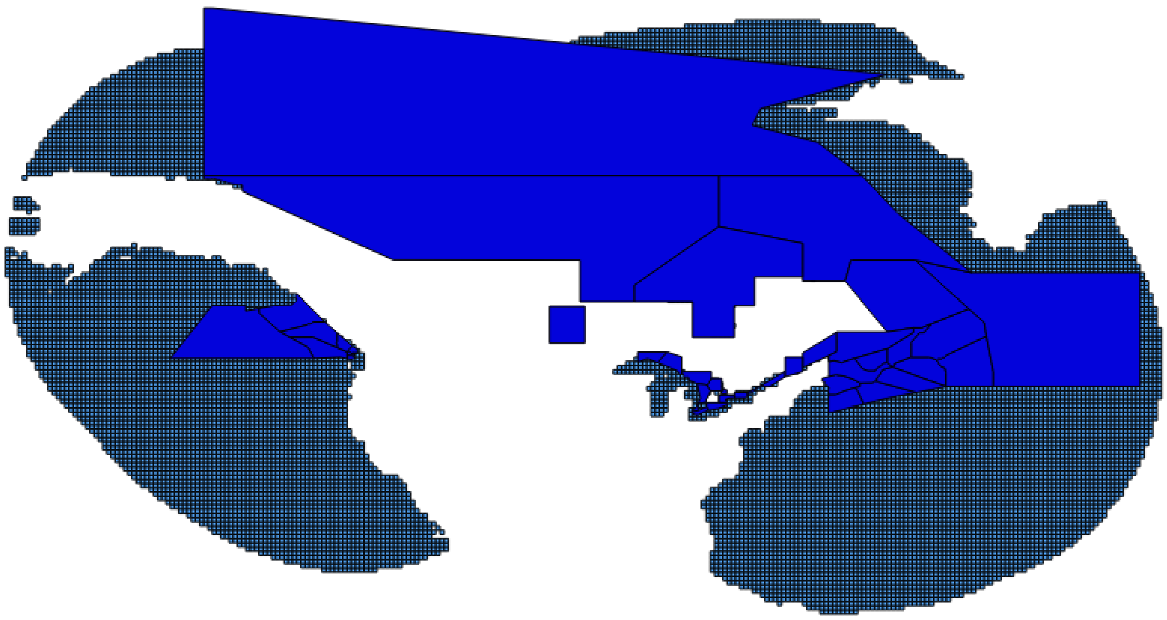

Figure 2.4. Canadian SAR Areas

To conduct the geostatistical analyses, both the number of fishing incidents in each grid as well as the estimated traffic level are associated with the centroid of each grid cell.

3 Spatial Autocorrelation

3.1 Introduction

Spatial Autocorrelation indicates the extent to which the occurrence of one feature is influenced by similar features in the adjacent area. If the existence of one feature attracts similar features to occur in its neighbourhood, the spatial autocorrelation is positive (i.e. the spatial distribution is characterized by groups of similar entities). Conversely, if the existence of one feature leads to repelling similar features from its neighbourhood, the spatial autocorrelation is negative (i.e. the distribution is represented by a scatter pattern). If neither attraction nor repulsion is observed, there is no spatial autocorrelation among the features. It is crucial to examine the spatial dependency to see if the inclusion of a spatial term is needed; if there is no spatial relationship between incidents and the location of an incident is independent of the location of any other incident, then spatial modeling is irrelevant and useless. In other words, spatial dependency should be viewed as a source of information rather than something to be corrected.

There are a number of formal statistics that attempt to measure spatial autocorrelation. The most common methods are Moran's I or Geary's C statistics. There exists a core relationship between Geary's C and Moran's I spatial autocorrelation coefficient in that they often yield similar results. However comparisons between Moran's I and Geary's C suggest that the former performs better because its variance is less affected by the distribution of the sample data (Cliff and Ord, 1981).

To find Moran's I index (Moran, 1950) for any variable, X_i, a mean of all the observations in the study area is calculated and then the deviation of each observation from that mean is also calculated. The final step is to compare the value of deviation from the mean at any location with the deviation value at all other locations:

$$I = N\sum_{i}\sum_{j} W_{ij}\ (X_i\ -\ \bar{X})(X_j - \bar{X}) / (\sum_{i}\sum_{j} W_{ij})(\sum_{i} X_i - \bar{X})^2$$

- N: Number of cases
- X_i : the variable value at particular location (i)
- X_j : the variable value at particular location (j) where (i≠j)
- $\bar{X}$: The mean of the variable, and
- W_{ij} ; the weight that applies to the comparison between location i and location j.

In the initial formulation of Moran's I, W_{ij} was a contiguity matrix; if the geographic area j was adjacent to the geographic area i, W_{ij} was equal to one, otherwise 0. Cliff and Ord (1973) improved the formula to accept any type of weight.

Moran's I or weighted Moran's I is similar to a correlation coefficient since it compares the sum of cross-products of values at different locations and it varies between -1 and +1. When nearby points have similar values, the cross product is high and when they have dissimilar values, the cross product is low. In other words, a high value of I indicates more spatial autocorrelation than a low one. However, an index value equal to zero doesn't indicate a lack of spatial dependence, but instead a number which is negative and very close to zero (Levine, 2002).

The theoretical expected value of I, E(I), depends on the sample size N:

E(I)= -1/(N-1)

Values of I above E(I) which differ from it significantly, show positive spatial autocorrelation and values below E(I) and significantly different from it indicate negative spatial autocorrelation.

3.2 Testing the significance of Moran's I

After calculating Moran's I index, a test of significance should be carried out. The test of significance compares the empirical distribution with the theoretical one by taking the theoretical standard deviation Z(I) into account:

$Z(I)= [I-E(I)]/S_{E(I)}$

where I is the calculated Moran's index, E(I) is the value of Moran's index if the data was distributed randomly over the study area and $S_{E(I)}$ is the standard deviation of the data. Cliff and Ord (1981) assumed that the expected value and the variance of Moran's I can be derived under the assumption of randomness and normality. Randomness means that each observed value could have occurred at any location, and there is not a relationship between the values and their spatial characteristics, and normality assumes that the standardized variable, Z(I), has a sampling distribution which follows a standard normal distribution. Thus, if Z(I) is greater than 1.96, a null hypothesis of spatial randomness can be rejected at a 5% confidence level.

3.3 Test cases

Table Table **3.1** presents the result of Moran's I spatial autocorrelation calculation for Fishing and Shipping Incidents and Activities for each grid resolution. In this study, weight is defined to be distance based (i.e. the inverse distance between locations i and j) which is the most common method to represent the correlations between observations over a study area (Levine, 2002):

$W_{ij=} 1/d_{ij}$

where d_{ij} is the distance between locations i and j.

Defining weight as the inverse of the distance may lead to unreliable results when the distance between two areas is very small and consequently weight can approach infinity. This situation can arise in adjacent geographic areas. Therefore an adjustment should be applied for small distances so that the maximum weight cannot be greater than a specified threshold. In this study, the adjustment scales distances to one nautical mile (i.e., minimal distance is considered one nautical mile.)

Table 3.1. Moran's I Index for Fishing and Shipping Incidents and Activities

Activity type and grid size	Variable	Moran I	Spatially random (expected) I	Standard deviation of I	Z value
Fishing 0.1	Incidents	0.043952	-0.000100	0.000349	126.392389
	Activity	0.343368	-0.000100	0.000351	979.601330
Fishing 0.25	Incidents	0.038648	-0.000237	0.000784	49.614984
	Activity	0.074994	-0.000237	0.000786	95.771749
Fishing 0.5	Incidents	0.040242	-0.000627	0.001456	28.073789
	Activity	0.115969	-0.000627	0.001529	76.280861
Fishing Uniform	Incidents	0.042677	-0.000284	0.000663	64.768351
	Activity	0.117060	-0.000284	0.000705	166.510289
Shipping 0.1	Incidents	0.024117	-0.000059	0.000279	86.693507
	Activity	0.087701	0.000059	0.000276	317.734389
Shipping 0.25	Incidents	0.003795	-0.000189	0.000691	5.765388
	Activity	0.03123	-0.000189	0.000477	65.906085
Shipping 0.5	Incidents	0.010535	-0.000107	0.000274	38.833811
	Activity	0.066506	-0.000107	0.000280	238.326171
Shipping Uniform	Incidents	0.015614	-0.000160	0.000326	48.328389
	Activity	0.042987	-0.000160	0.000320	135.001732

The results demonstrate that fishing incidents are positively spatially correlated in the study area. One can conclude that geographic areas with many incidents tend to be located close to other incident-rich geographic areas and, conversely, geographic areas with few or no incidents tend to be located close to other geographic areas with few or no incidents. These indices provide statistical evidence of the existence of fishing incident clusters in the study area, although for precise characterization more detailed cluster and hot spot analyses should be invoked. High Z values confirm the significance of spatial autocorrelation between fishing incidents for each grid size. Fishing traffic is also positively spatially correlated, but compared to incidents, since the indices are greater, it suggests that traffic is more clustered than incidents.

Shipping incidents and shipping traffic results can be interpreted in a similar manner to fishing incidents and fishing traffic, respectively. In other words, shipping incidents are positively spatially autocorrelated and so is shipping traffic, but shipping traffic levels are slightly more concentrated.

Fishing incidents and activities are mostly concentrated near the shore, whereas shipping activities and incidents are more autocorrelated further from shore (between 20 nm to 200 nm from shoreline).

This method investigated the spatial autocorrelation for incidents or traffic respectively for each category of vessel. This type of analysis is crucial to carry out in the initial steps of the study, since it helps to understand the data distributions over the study area, leading to more detailed hot spot analyses if necessary.

4 Hot Spot Analysis

Hot spot analysis determines high concentrations of incidents or activity within a certain area. There are many different statistical techniques designed to identify hot spots and most of these techniques fall under the general label of clustering. Cluster Analysis groups a set of events in such a way that objects in the same group (called a cluster) are more similar to each other than they are to objects in other groups. Spatial clustering needs georeferenced data (i.e. geographical location of observations) to classify them into clusters.

Hot spot analysis techniques can be grouped as following (Shahrabi, 2006):

- Point location techniques
- Hierarchical techniques
- Partitioning techniques
- Risk-based techniques

4.1 Point Location Techniques

This method, also known as "Quadrat Count Methods," simply counts the number of events that occur in each quadrat and examines the density or intensity of points over the study area. The first step is to create uniform-size grids (quadrats) to cover the study area and then to build a frequency distribution based on the number of events in each grid. The choice of grid size in this method is very important since very large quadrats may lead to an unacceptably coarse description of the data while very small ones may result in having no incidents in most of the quadrats (Nicholson, 1998).

When the objective is to find hot spot areas instead of hot spot points, it is better to use the fuzzy mode of quadrat counting. In this approach a search radius is applied around each quadrat, so the incidents which are occurring around or near one particular location are also included in the count. This is the case most of the time, since it is very rare to have many events happening at exactly one point. But the problem with this method is that some incidents may be counted more than once, particularly in areas with high concentrations of incidents.

4.2 Hierarchical Techniques

Hierarchical hot spot analysis examines the spatial characteristics of observations to determine clusters. Generally these techniques are agglomerative, meaning they start with many groups (first-order clusters) and reduce the number of clusters by amalgamating them based on some spatial characteristics such as closeness. The process is continued until either only one cluster remains or the grouping criterion fails. In other words, the first order clusters may be clustered into second order clusters and second order clusters may be grouped into third-order clusters, and so on.

Hierarchical techniques mostly differ based on the clustering criteria. The most common method is the nearest neighbour method. This method works in a way that it compares the distance between all pairs of points with a pre-determined threshold. Points which are closer than the threshold are grouped into one cluster. After clustering all the points, the center of each cluster is considered to be a point which would then be analyzed for generating second-order clusters. The number of points in each cluster and the threshold should be determined by the analyst. Since the choice of parameters are arbitrary, if they were not chosen properly, then the method may fail to detect small hot spots.

4.3 Partitioning Techniques

These techniques attempt to group points into a specified number of clusters which is usually defined by the analyst. The K-means method, which is the most important partitioning method, aims to find the best location of the centers of K clusters (where K is selected by the analyst) and assigns each observation to the relevant cluster in such a way that the distance between the point and the center of the cluster is minimum.

Setting the number of clusters before the analysis gives the analyst full control over the size and number of clusters but also makes the method somewhat arbitrary, perhaps not representing all the actual hot spots over the entire study area.

4.4 Risk-Based Methods

Risk-based clustering techniques identify clusters relative to an underlying base 'at risk' variable, such as activity (traffic in our case) or population. The risk-adjusted nearest neighbour hierarchical clustering technique dynamically adjusts the threshold distance in the nearest neighbour techniques according to the distribution of this second underlying variable.

For the purpose of this phase of the AMT project, wherein the probability of incidents relative to traffic is needed rather than the volume of incidents alone, this method seems to be the most useful among other hot spot analysis methods. This method defines clusters of incident points that are closer than what would be expected on the basis of activity levels. The threshold distance is dynamically altered such that in those areas with high activities the threshold is small, and within regions with low activity the threshold is larger.

4.4.1 Test cases

The risk-adjusted nearest neighbour hierarchical clustering was applied to fishing incidents (annual and quarterly) and shipping incidents (annual) using the uniform size grids of 0.5 degree, by using the "CrimeStat II" software.

Table 4.1 summarizes the results of the Risk-Adjusted Clustering for fishing (yearly and quarterly) and shipping (yearly) incidents.

Table 4.1. Risk-Adjusted Clustering Results

	Number of 1st Order Clusters	**Number of 2nd Order Clusters**	**Number of 3rd order clusters**
Fishing-Year 2010	38	4	
Fishing-1st Quarter 2010	25	3	
Fishing 2nd Quarter 2010	30	4	
Fishing 3rd Quarter 2010	26	3	
Fishing 4th Quarter 2010	25	4	
Shipping Year 2010	48	8	1
Shipping Year 2010-without Arctic	10		

Since the number of clusters is large, to make them clearly visible only the second and in one case third order clusters have been represented in Figures Figure 4.1Figure 4.2Figure 4.3Figure 4.4Figure 4.5, **Error! Reference source not found.** for fishing annual incidents, fishing quarterly incidents and annual shipping incidents, respectively. The first order clusters and the characteristics of each cluster can be found in Appendix A.

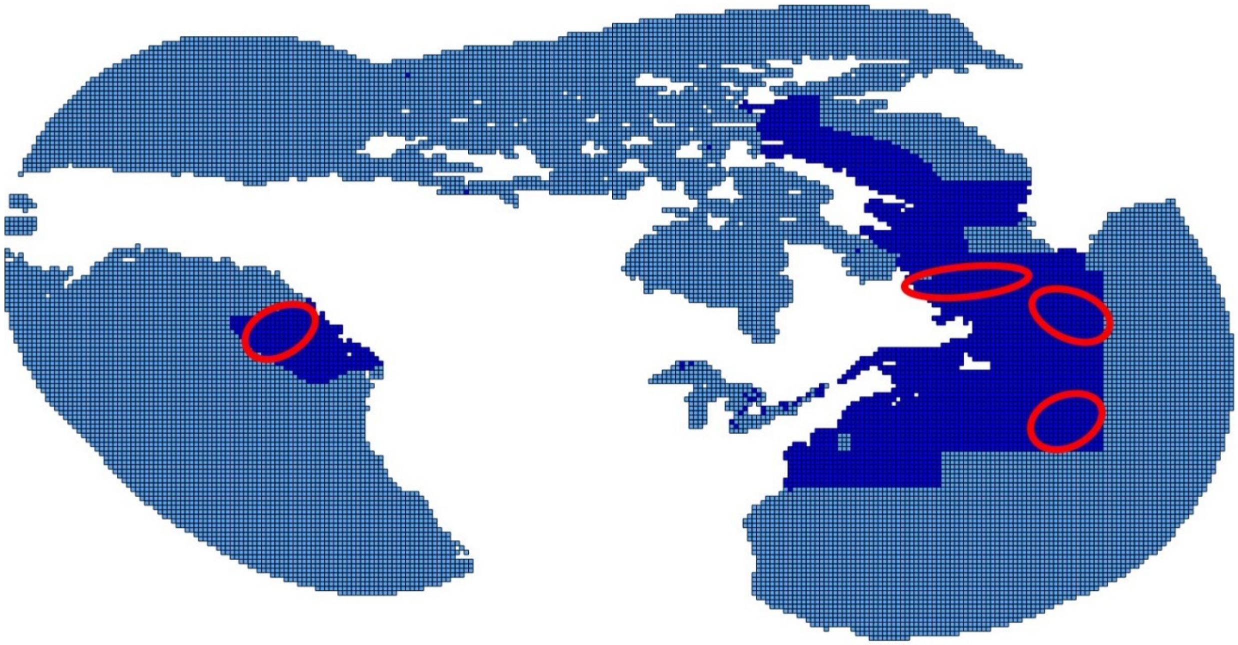

Figure 4.1. Second Order Risk-Adjusted Clusters-Fishing-Year 2010

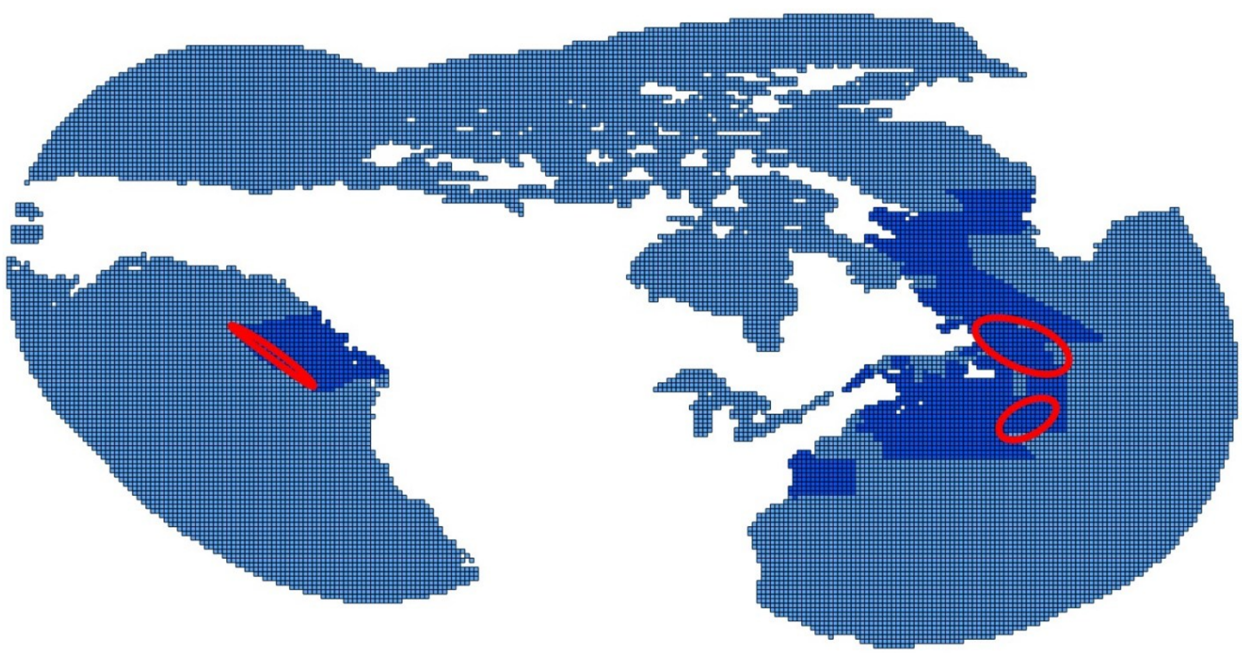

Figure 4.2. Second Order Risk-Adjusted Clusters-Fishing-1st Quarter-2010

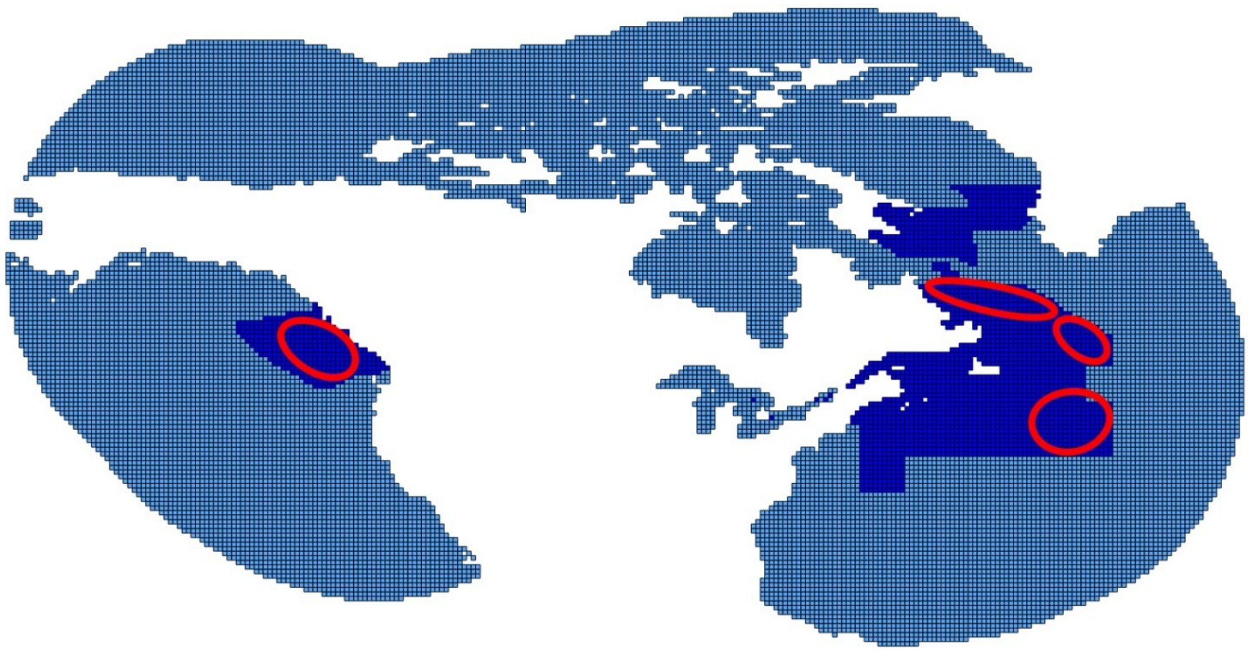

Figure 4.3. Second Order Risk-Adjusted Clusters-Fishing-2nd Quarter-2010

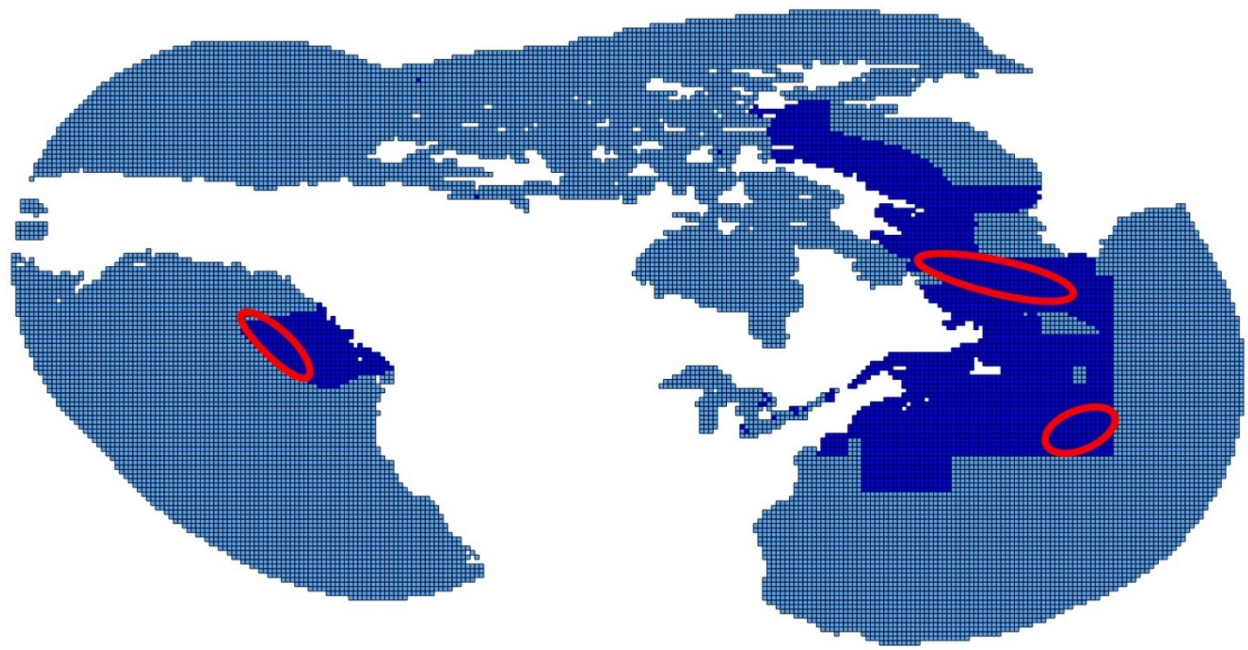

Figure 4.4. Second Order Risk-Adjusted Clusters-Fishing-3rd Quarter-2010

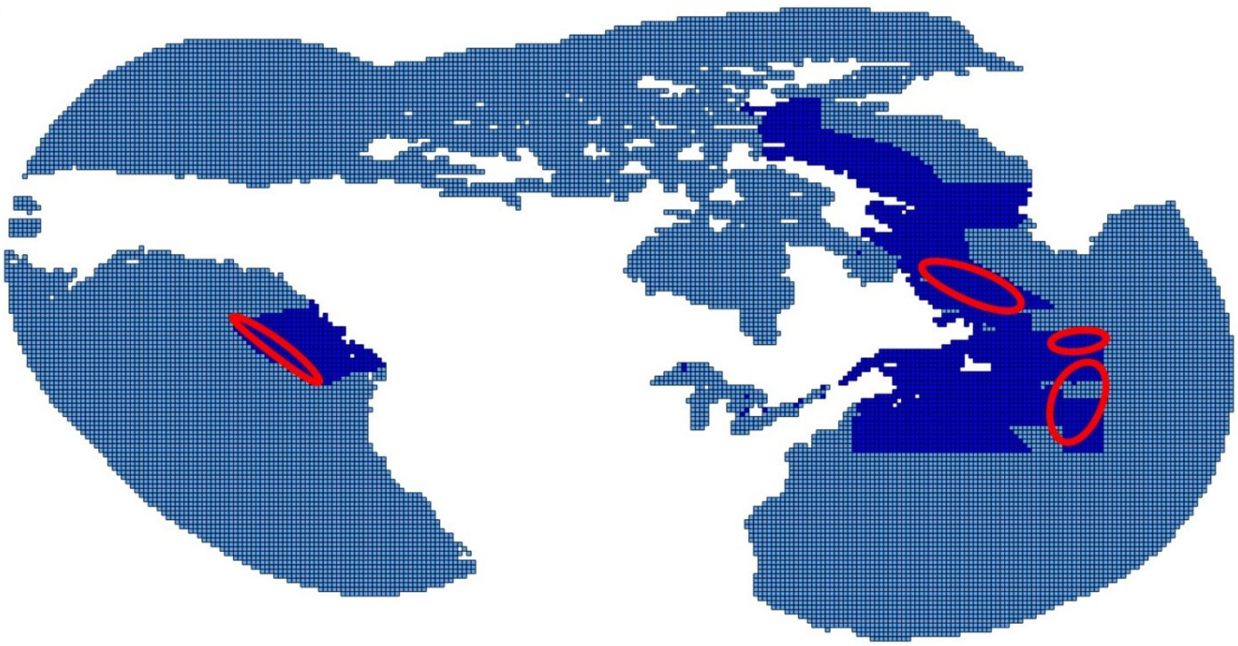

Figure 4.5. Second Order Risk-Adjusted Clusters-Fishing-4th Quarter-2010

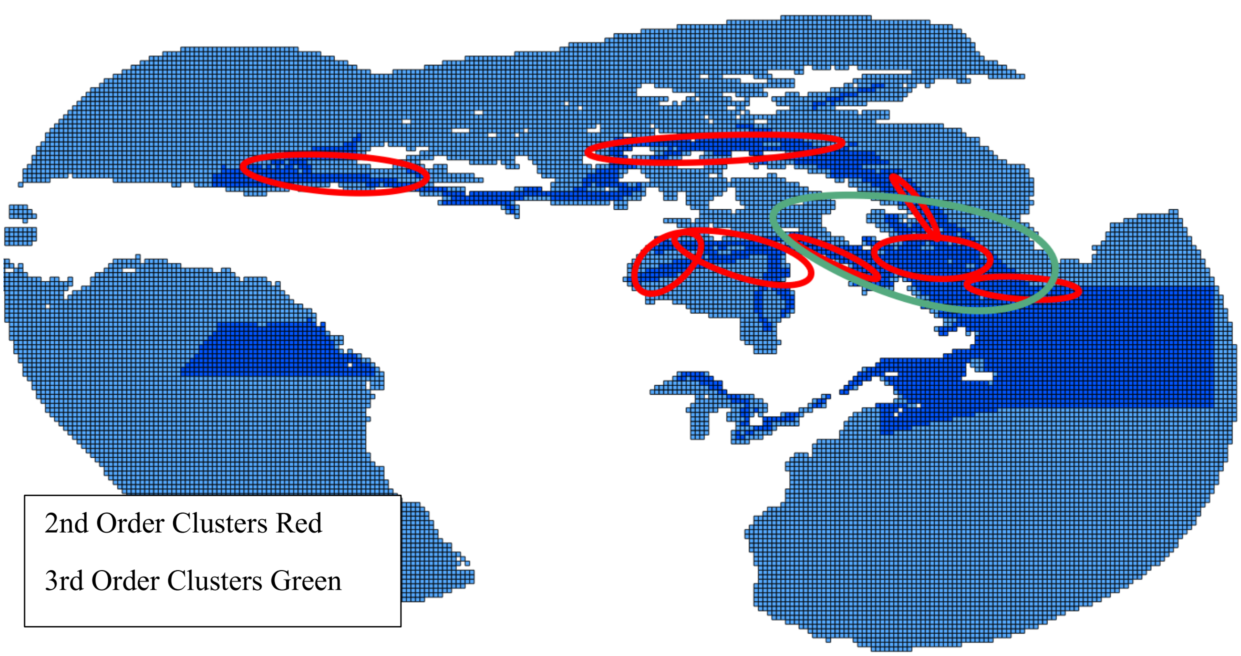

Figure 4.6. Second and Third Order Risk-Adjusted Clusters-Shipping-Year 2010

Based on the maps, the hot spots of fishing incidents relative to fishing traffic do not change dramatically during the different quarters of the year. Although most of the fishing incidents happen near shore, incident hot spots with respect to traffic happens further from shore.

Shipping incident clustering shows that incidents with respect to traffic are more clustered in the Arctic region. The reason for this could be the small amount of traffic up there and a high number of incidents relative to this slight traffic. To eliminate the effect of this “distortion” and provide a more realistic picture of shipping incidents in Canada, it was decided to find the clusters of incidents related to shipping traffic without considering the Arctic region. The Arctic region for the purposes of our study can be defined based on the AOI of AMT-Phase 2. FigureFigure 4.7 depicts the boundaries for the AMT Arctic area:

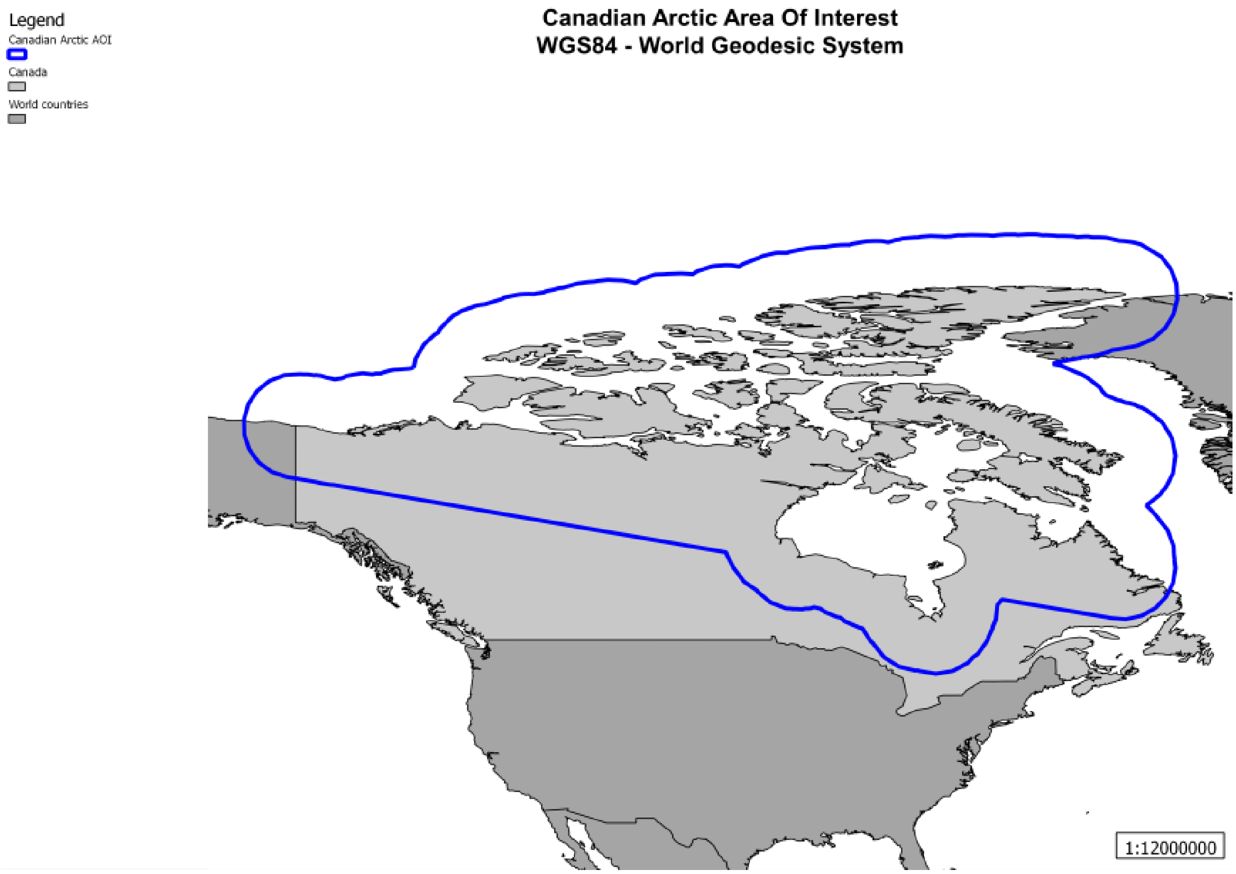

Figure 4.7. Canadian Arctic AOI for the AMT study. Source: Engler and Pelot, 2012, Etienne, et al. 2013.

Error! Reference source not found.8 shows first order clusters of shipping incidents related to he traffic amount without considering Arctic region. The results represent ten first order clusters, the characteristics of each cluster can be found in Appendix A.

Figure 4.8. First Order Risk-Adjusted Clusters-Shipping-without Arctic-Year 2010

5 Spatial Interpolation

Spatial interpolation is the procedure of estimating the value of properties at unsampled locations within the area covered by existing observations. These methods are useful to interpolate traffic or incident occurrences in unsampled areas based on mathematical distributions built on sampled data or to identify patterns of incident occurrence based on the traffic amount in each specified region.

5.1 The inverse distance weighting

The inverse distance weighting, or inverse distance weighted (IDW), method estimates the values of an attribute at unsampled points using a linear combination of values at sampled points weighted by an inverse function of the distance from the point of interest to the sampled points. The assumption is that sampled points closer to the unsampled point are more similar to it in their values than those further away (Bartier and Keller, 1996).

5.2 Kriging

Kriging is an exact statistical method, which means it generates an estimate that is the same as the observed value at a sampled point (Cressie,1990). Hemyari and Nofziger (1987) define Kriging as follows:

"Kriging is a form of weighted averaging in which the weights are chosen such that the error associated with the [predictor] is less than for any other linear sum."

Kriging estimators are all variants of the following equation:

$$\dot{Z}(x_0)-\mu=\sum^{n} \lambda_i [Z(x_i)-\mu(x_0)]$$

Where μ is a known stationary mean, assumed to be constant over the whole domain and calculated as the average of the data. The parameter λ_i is Kriging weight for the observed point at location i and it is estimated by minimizing the variance of the attribute which is going to be estimated in unsampled areas; n is the number of sampled points used to make the estimation and depends on the size of search window; and $\mu(x_0)$ is the mean of samples within the search window.

Simple Kriging assumes that the mean, variance and covariance are constant over the region which is a restricting assumption. Ordinary Kriging uses local mean instead of global mean to overcome the restrictions in simple Kriging.

5.3 Splines

Spline interpolation aims to create a smooth surface over the study area by means of a spline interpolation function. Generally speaking a spline is a curve that connects at least two specific points and smoothing splines are the functions which minimize the mean squared error, subject to a constraint on the average curvature. The simplest spline is linear function. In Spatial Spline interpolation, there is almost a universal consensus that cubic is the optimal degree for splines (Schumaker,1981).

Splines can generate sufficiently accurate surfaces from only a few sampled points when there is only one output and input variable. However in case of multiple variables, Kernel estimations, which will be described in the following section, are easier to both program and analyze mathematically. Splines usually work best for gently varying surfaces like a temperature field.

5.4 Kernel Density Estimation

In this method a symmetrical surface is placed over each data point and then based on a mathematical function, the distance from the point to a reference location is evaluated. Next, it sums the value of all surfaces at all reference locations. The underlying density distribution is estimated by summing the individual Kernel density estimators at all locations to produce a smooth cumulative continuous density function. The Kernel estimation is applied to a limited search distance which is called bandwidth and it is the key factor in the smoothness of the Kernel estimation. Bandwidth can be fixed or adaptive. Adaptive bandwidth adjusts the search radius in a way that in areas with high concentration, the bandwidth is narrow and in more sparse areas it is larger.

The Kernel estimator can be a uniform distribution, negative exponential distribution, quadratic distribution, triangular distribution, and normal distribution. There are two versions of the Kernel density method: Single and Dual. The single Kernel is used to interpolate individual points such as incident locations and the Dual Kernel density interpolates points with respect to a baseline distribution (e.g. incidents with respect to traffic). First and second variables are interpolated to the same reference grid (geographic unit) and density can be calculated in different ways such as ratio of densities, log ratio of densities, or relative difference in densities.

Kernel estimation differs from Hot Spot techniques by generalizing the incident occurrence over the entire study area and not just providing statistical summaries for the incidents. But the main advantage for this method lies in determining the spatial spread of the risk (probability) of an accident. The spread of risk can be defined as the area around a defined cluster in which there is an increased likelihood for an accident to occur based on spatial dependency. Since this project aims to uncover the relationships between incident occurrences and related traffic, Dual Kernel estimation is one of the recommended methods that can help to achieve this goal.

5.4.1 Test Cases

Dual Kernel estimation can be applied through the “CrimeStat II” software. The first layer is the location of incidents and the second layer is the location of activities (traffic levels). The outcome is the ratio of densities which can be derived by dividing the Kernel estimate for incidents by the Kernel estimate for activity. The Kernel density estimator which has been used in this study is the normal distribution:

$$g(X_i) = \sum \{\frac{1}{2\pi * h^2} e^{-[\frac{d_{ij}^2}{2*h^2}]}\}$$

where:

$g(X_i)$ is the Kernel function at point location i,

d_{ij} is the difference between location i and any reference point in the study area, and

h is the standard deviation of the Normal distribution (the bandwidth).

In this study the bandwidth is chosen to be adaptive and contains at least 50 points.

Figure 5.1,Figure 5.2,Figure 5.3,Figure 5.4,Figure 5.5Figure 5.6 represent Dual Kernel estimations for fishing annual incidents, fishing quarterly incidents and annual shipping incidents, respectively. Darker colors represent greater values (i.e. higher occurrence of incidents relative to traffic).

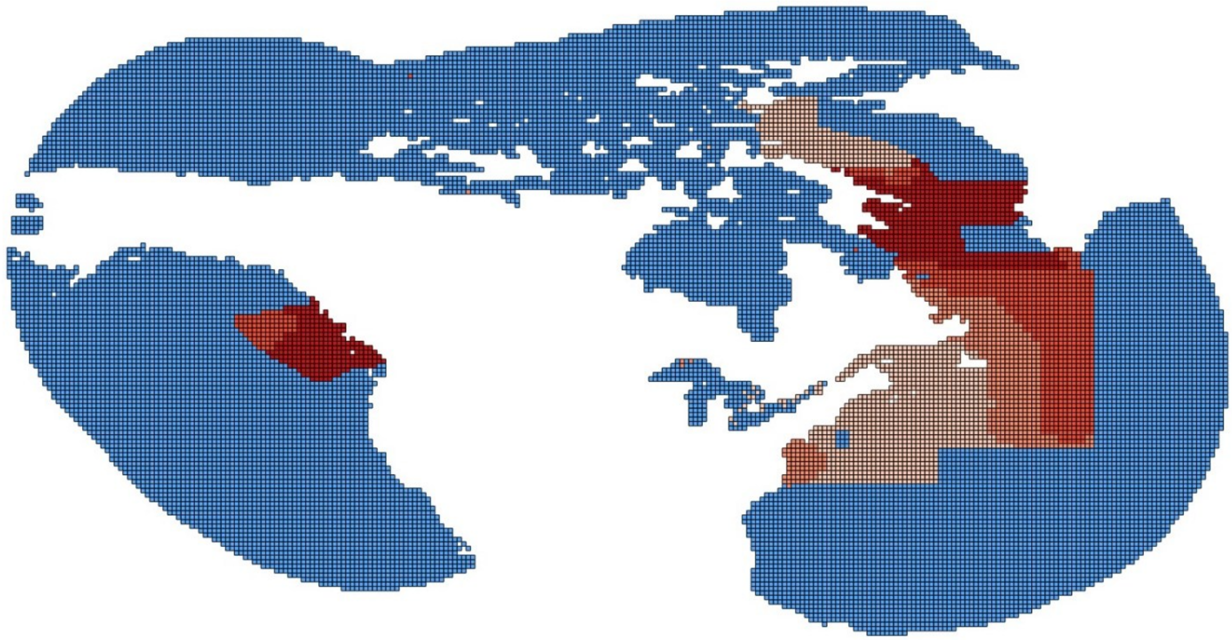

Figure 5.1. Dual kernel estimation-Fishing-Year 2010

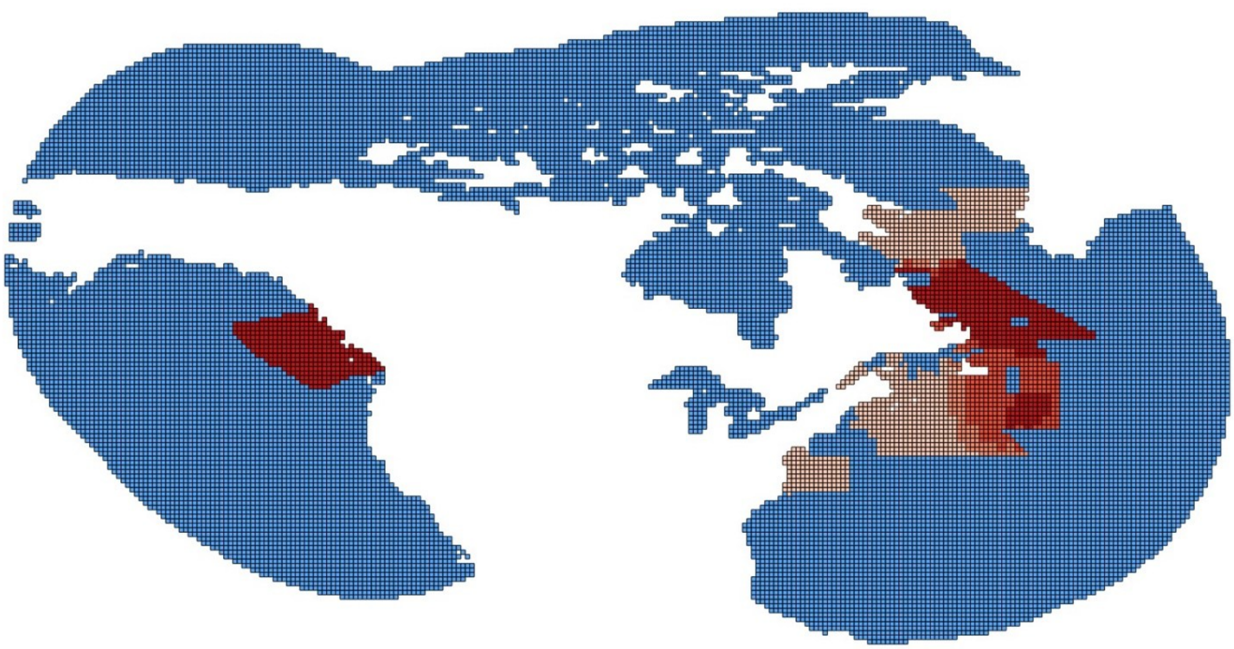

Figure 5.2. Dual kernel estimation-Fishing-1st Quarter-2010

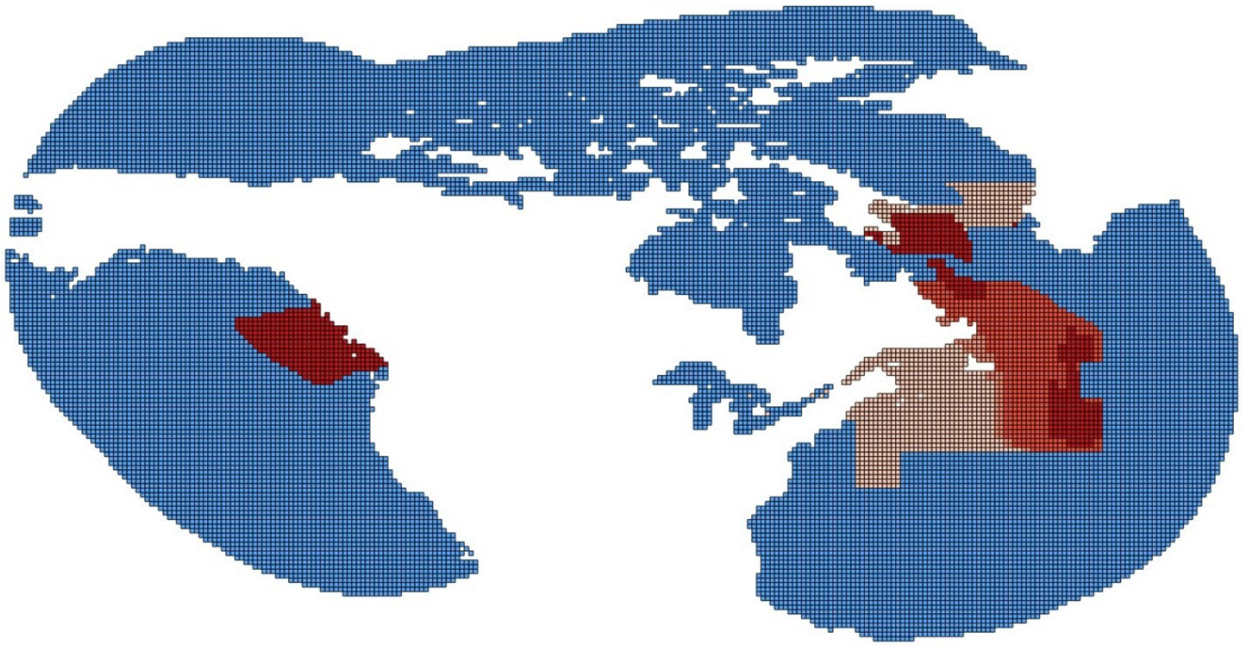

Figure 5.3. Dual kernel estimation-Fishing-2nd Quarter-2010

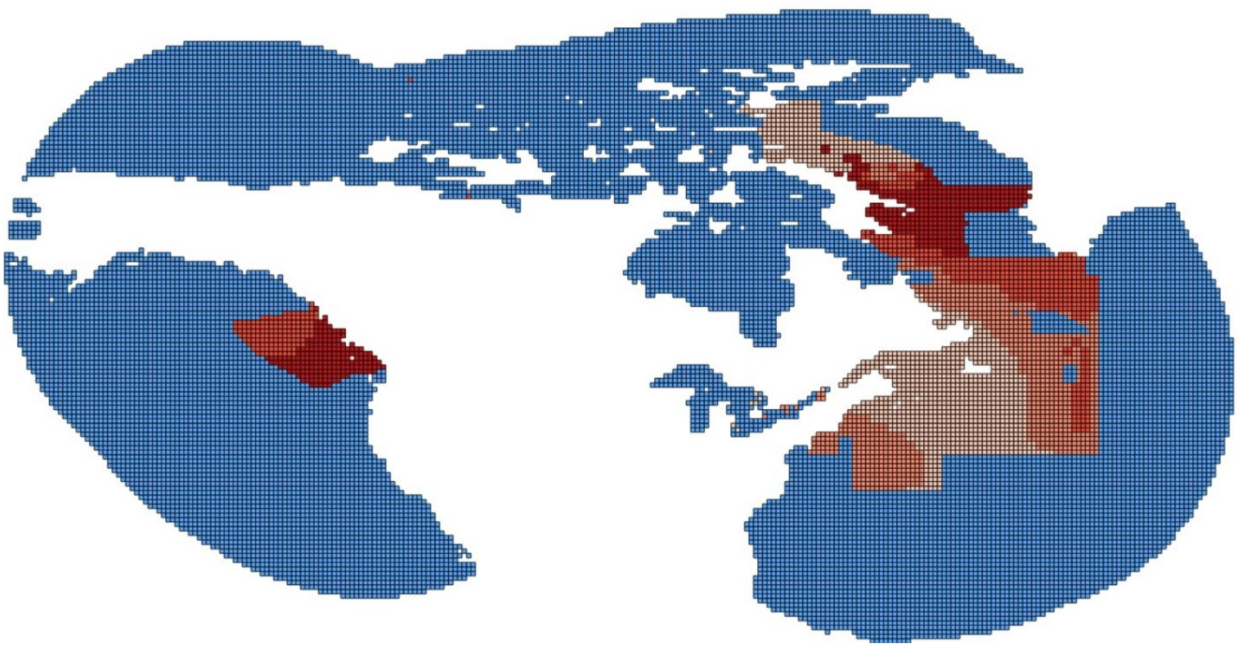

Figure 5.4. Dual kernel estimation-Fishing-3rd Quarter-2010

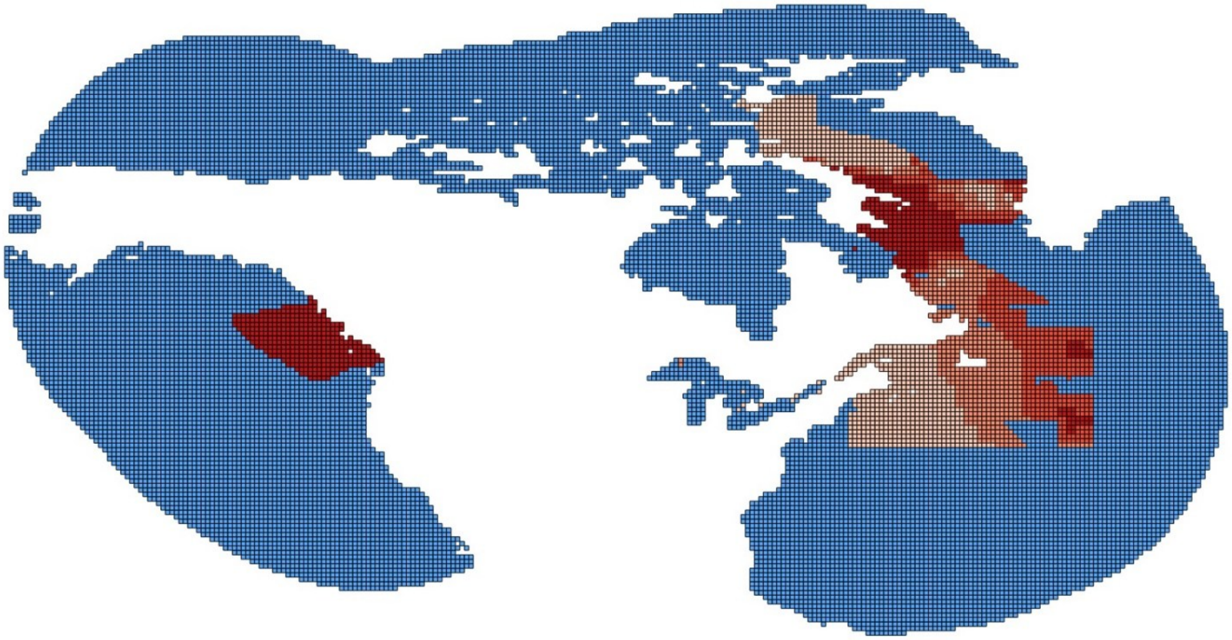

Figure 5.5.Dual kernel estimation-Fishing-4th Quarter-2010

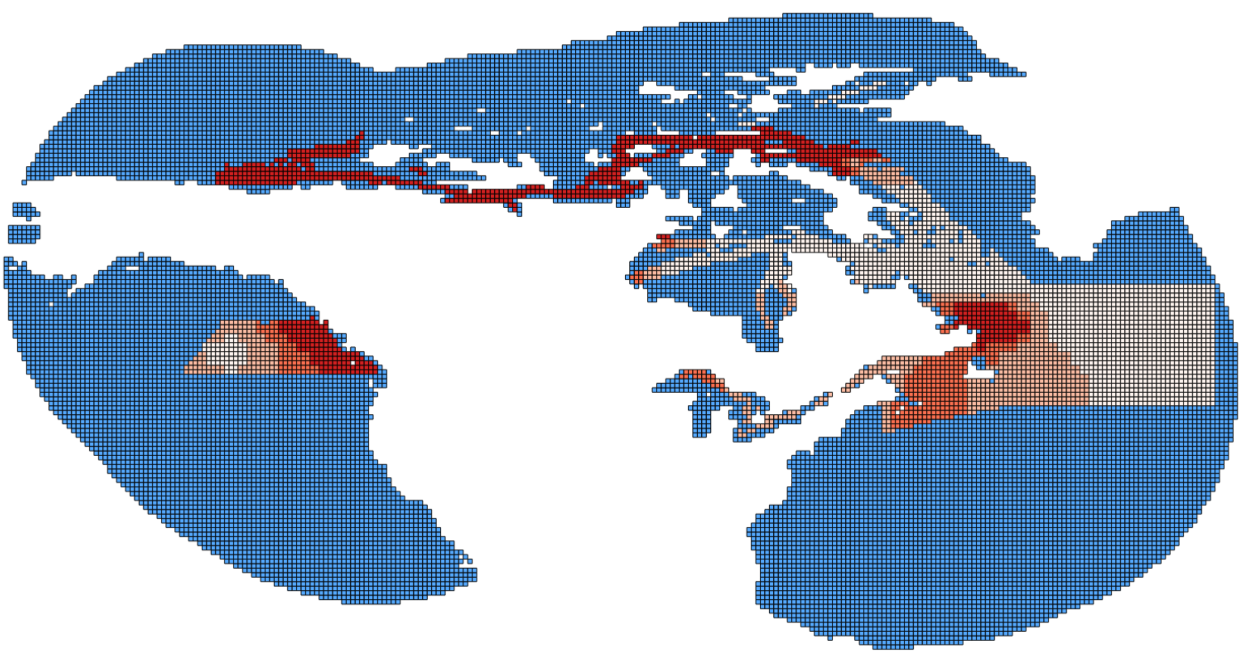

Figure 5.6. Dual kernel estimation-Shipping-Year 2010

The results of the fishing analysis are in alignment with the results from the Risk-adjusted nearest neighbourhood hierarchical clustering, whereby there is not a significant change in spatial patterns during different quarters over the year and yielding the same two most important hot spots, Southwest of Vancouver Island and North of Newfoundland and Labrador. When it comes to shipping the probability of incidents occurring relative to traffic is decreased by going further from shore which is in alignment with the results from the Clustering Method as well, and high incident rates with respect to traffic in Arctic showed up using this method too. When interpreting Shipping incident rates, one should consider that the number of shipping incidents are very few comparing to the number of grids with shipping traffic. In other words, a large number of grids in the Shipping study area have zero incident occurrences which may cause biased results. If it is needed to get more accurate results of shipping, it is better to analyze data for smaller regions or with coarser grids for more than one year data.

5.5 Spatial Regression

The main idea behind spatial modeling is to explore how the relationship between a dependent variable and independent variable(s) might vary geographically. Instead of assuming that a single model can be fitted to the entire study region, these models search for geographical differences in the relationships. Geographically Weighted Regression (GWR) is a refinement to classical regression. GWR model can be expressed in matrix notation:

$$y_i = X_i\beta_i + \varepsilon_i$$

where X_i is a row vector of explanatory variables and β_i is a column vector of regression coefficients at location i. In GWR, a spatial structure is specified in the model through applying weights to the data. The weights, which are applied to the outcome variable and the covariates, are calculated from a kernel function that typically assigns more weight to observations that are spatially closer to the data point (i^{th} location) where the model is estimated. The introduction of the weights into the model follows from the assumption of spatial autocorrelation, where observations more proximate in space are thought to be more similar. One common way in determining weights is to use the bisquare function which is defined by

$wij = [1 - d_{ij}^2/d^2]^2$ if $dij < d;$

$wij = 0$ otherwise.

Where j represents a specific point in space at which data are observed and i represents any point in space for which parameters are estimated, and d_{ij} is the geographic distance between i and j (Brundon et al., 1996).

The kernel function used to calculate the weights in the GWR setting takes as input distances between all locations, conveniently in the form of a distance matrix. The kernel function has a bandwidth parameter that determines the spatial range of the kernel. The bandwidth parameter must be selected a priori or estimated from the data (fixed or adaptive). One way to choose the bandwidth is to cross-validate different bandwidths and choose the best one which minimizes the root mean square prediction error for the geographically weighted regressions.

The main disadvantage of geographically weighted regression is that it is time consuming. When data sets are very big, it is better to use modified GWR methods such as gridded GWR. It can also be improved to fit more complex regressions over the study area.

5.5.1 Test Cases

GWR can be coded through the "R project" software. The outputs are the estimation of the regression model for each data point (i.e. centroid). The dependent variable is the number of incidents in each grid square, and the predictor is the related traffic for that grid. Since reporting the result for each grid centroid (e.g. 3609 points for Shipping) is not helpful in providing general insight into the situation, it was decided to classify coefficients for each vessel category into 4 groups (quartiles) and represent the results based on these quartiles. For example if a grid in the Annual Fishing dataset has a coefficient which is equal to 0.7 it is in 2nd quartile and if it is equal to 8.5 it falls into the 4th quartile. Table 5.1 summarizes the quartiles of results of GWR estimations for fishing and shipping incidents.

Table 5.1. Spatial Regression Results

	Coefficients				
	Min	**1st Quartile**	**2nd Quartile**	**3rd Quartile**	**Maximum**
Fishing-Year 2010	0	0.059000	0.089650	0.292200	18.250000
Fishing-1st Quarter	0	0.000000	0.000000	0.007190	83.820000
Fishing-2nd Quarter	0	0.000000	0.000141	0.015030	77.500000
Fishing-3rd Quarter	0	0.020920	0.021350	0.022500	0.494000
Fishing-4th Quarter	0	0.000000	0.001574	0.047770	167.600000
Shipping-Year 2010	0	0.007546	0.011107	0.013990	0.023584

Figure 5.8Figure 5.9Figure 5.10Figure 5.11, Figure 5.12 represent GWR estimations for fishing annual incidents, quarterly incidents and annual shipping incidents for uniform size grids (0.5 degree), respectively.

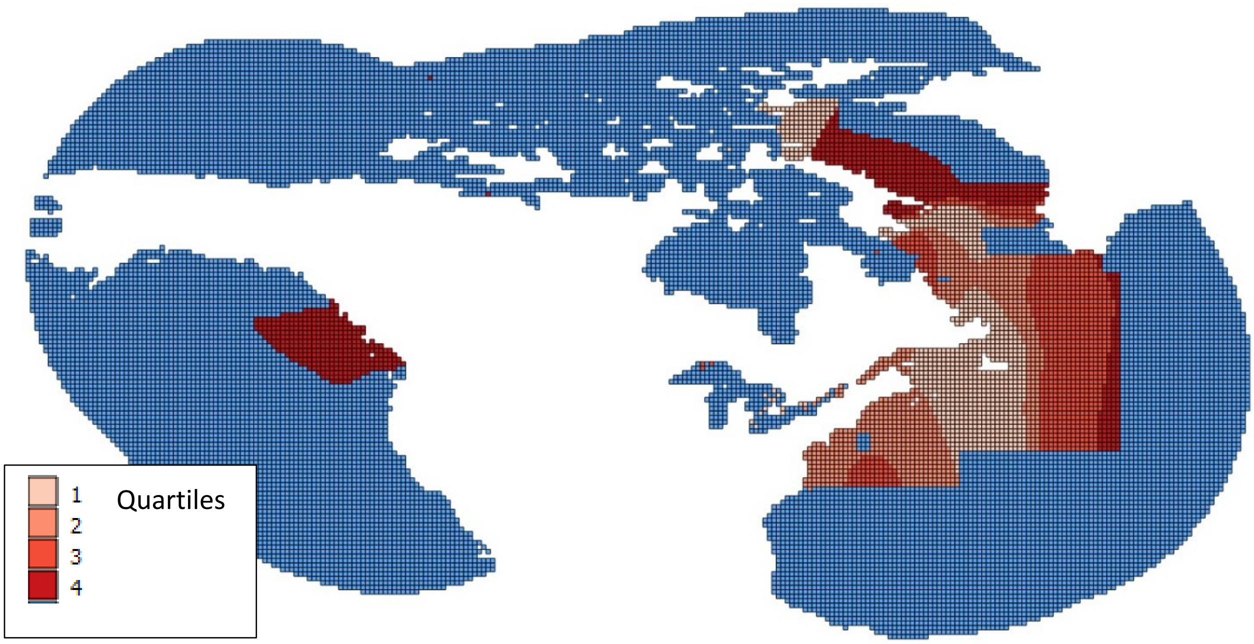

Figure 5.7. Spatial Regression Results-Fishing-Year 2010

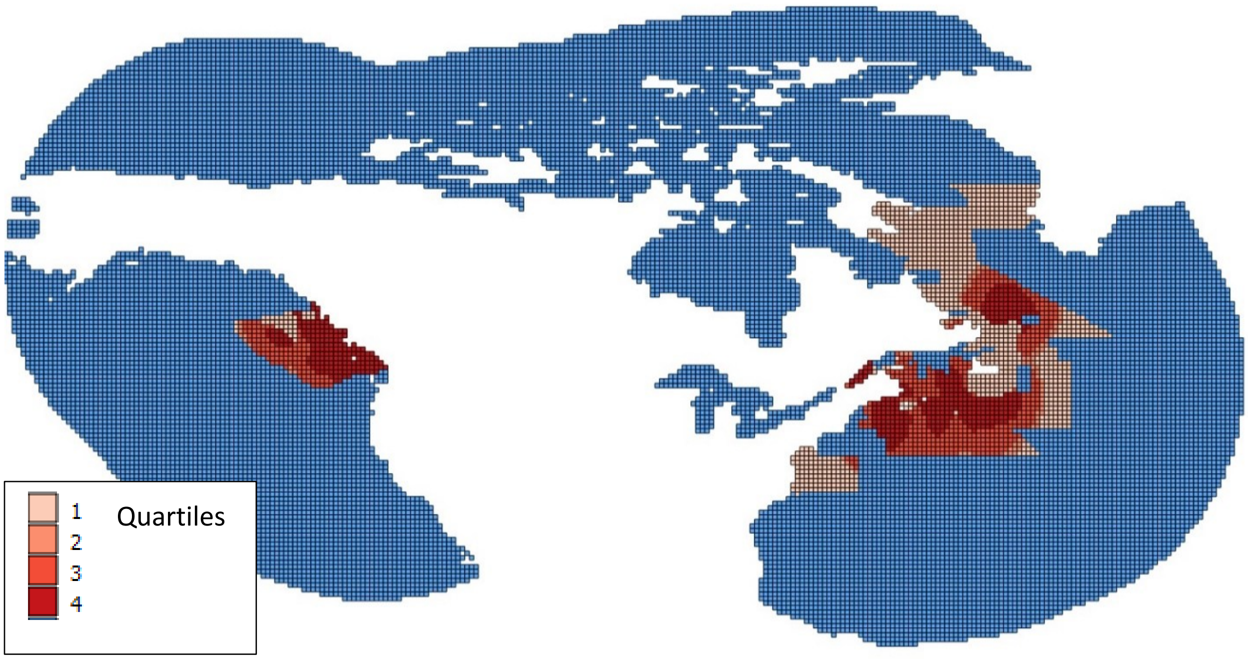

Figure 5.8. Spatial Regression Results-Fishing-1st Quarter-2010

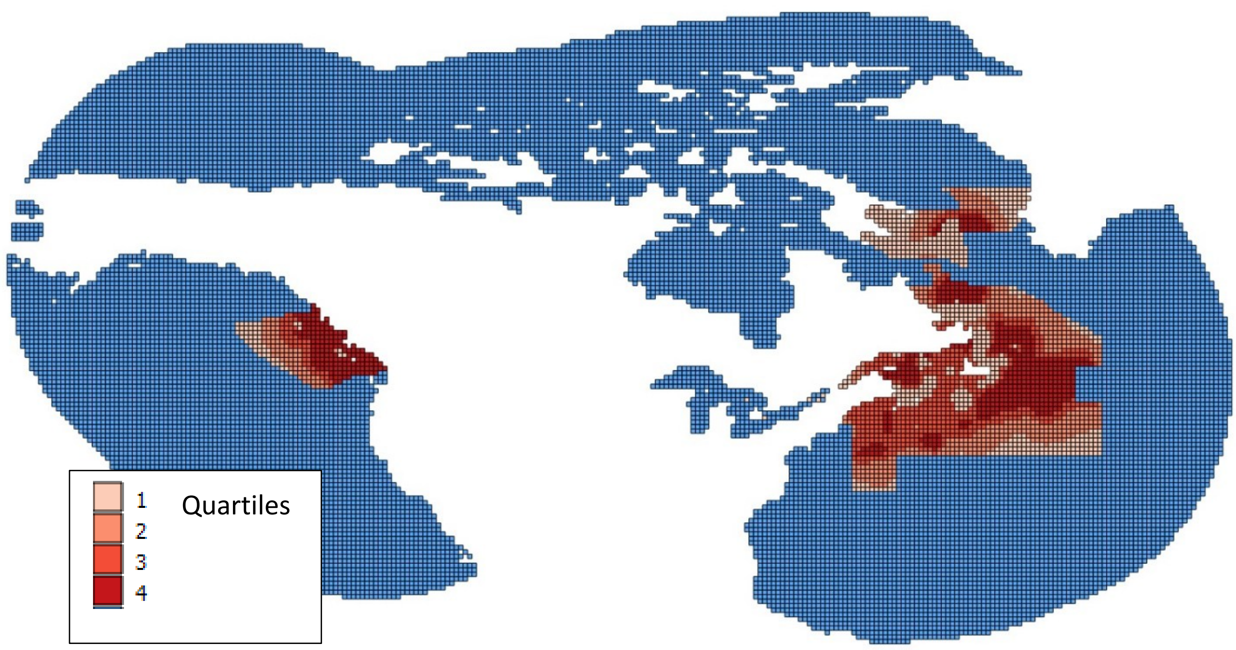

Figure 5.9. Spatial Regression Results-Fishing-2nd Quarter-2010

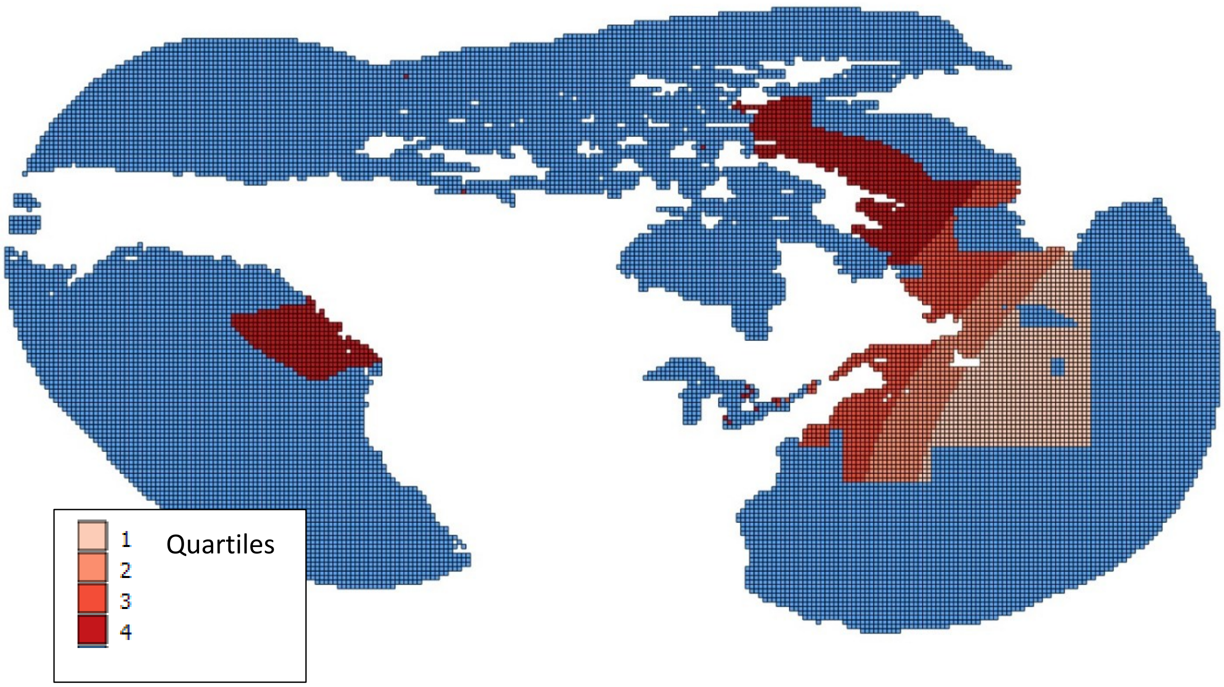

Figure 5.10. Spatial Regression Results-Fishing-3rd Quarter-2010

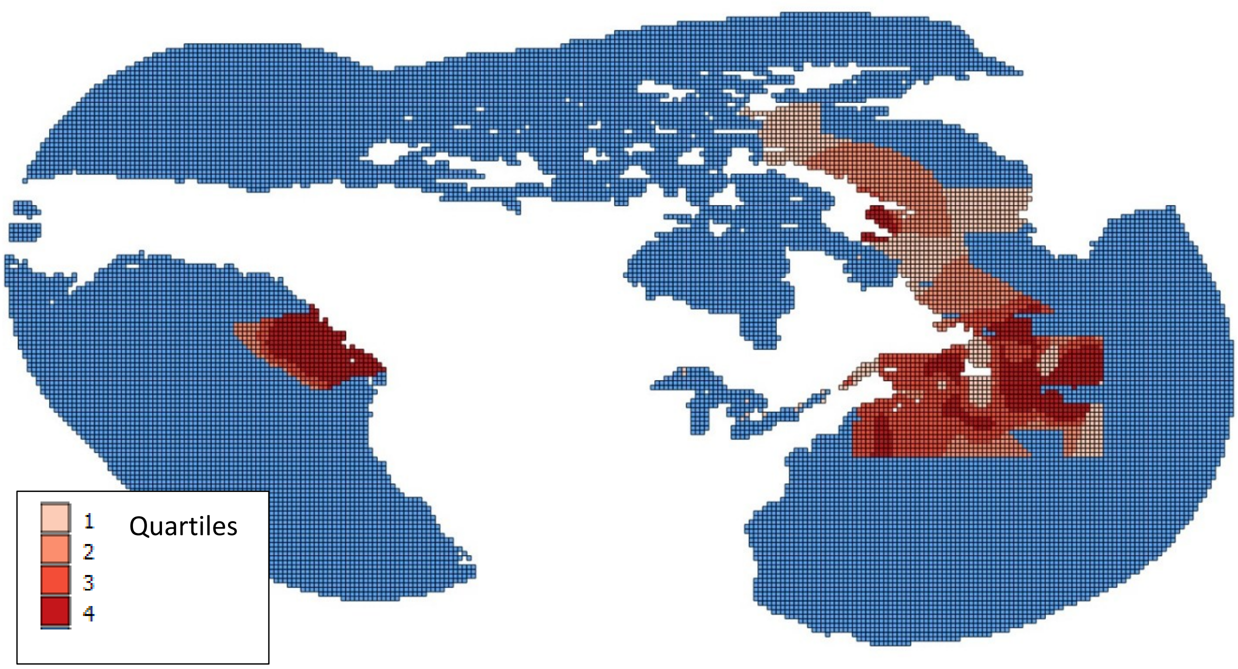

Figure 5.11. Spatial Regression Results-Fishing-4th Quarter-2010

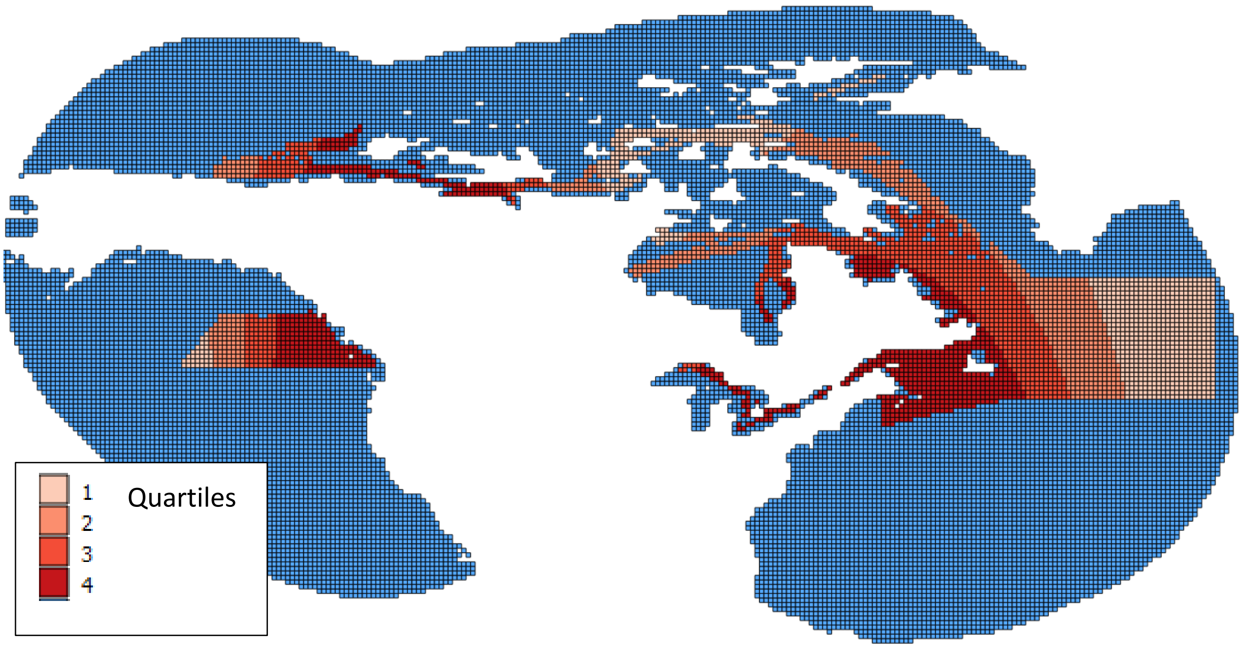

Figure 5.12. Spatial Regression Results-Shipping-Year 2010

Spatial Regressions helps to predict across the entire AOI how incident rates can change in each grid if there is a change in the amount of traffic in that grid. Results indicate that North of Newfoundland and Southwest of Vancouver Island there is a strong relationship between fishing incidents and traffic, which also showed up in the Dual kernel estimation and Risk-adjusted nearest neighbour clustering. But how the coefficients change in different locations in different quarters of the year is more sensitive than in the other methods.

GWR estimations of shipping incidents based on traffic indicate the incident rate decreases when the distance from shore increases, which was also the case in Dual Kernel estimation. As mentioned before, the number of shipping incidents is very few, meaning that almost 90% of grids with shipping traffic do not have incidents occurring in them which may lead to biased results.

6 Summary and extensions

The choice of spatial analysis method is dependent on what the purpose of analysis is. Spatial autocorrelation methods are useful to give insight about the data and examine if there exist any spatial relationships in the data over the entire study area. Moran's I is the most robust measurement of spatial autocorrelation.

When it comes to answer "who needs help the most", hot spot analyses are the most useful methods. When the concentration of incidents alone is important, Nearest Neighbour Hierarchical clustering and Single Kernel estimations are very helpful in finding precise hot spot areas. But when the purpose is to find the concentration of incidents relative to traffic, Risk-Adjust Nearest Neighbour Hierarchical Clustering or Dual Kernel Estimation should be carried out. For allocation purposes when it is not feasible to assign a SAR boat to each small hot spot, K-means clustering is the best method, since it is simple and produces medium-sized hot spots areas.

To determine in which regions incident rates are more sensitive to changes in traffic, Spatial Regression can be applied. The outputs of this method provide specific regression fits for each specific region and it can be used to predict incident rates in case of increases or reductions in the traffic amount. Spatial Regression is the extension of regular regression, and one can improve the outcomes by changing the regression distribution and examining more complicated relationships between traffic and incidents. It is also possible to include other factors such as vessel types or vessel lengths in the regression to get more realistic outputs if these attributes are suspected to affect the incident occurrences.

A valuable extension to this study would be to discriminate between low impact and high impact incidents. This report mainly focuses on the probability of incidents with respect to traffic, but a risk analysis can also consider the severity levels of consequences. The Canadian Coast Guard classifies incidents according to their type and level of severity:

- M - Maritime Incidents (M1, M2, M3, M4)
- A – Aeronautical Incidents (A1, A2, A3, A4)
- H – Humanitarian Incidents (H1, H2, H3, H4)
- U – Unknown Incidents (U4).

Maritime incidents are strictly associated with the distress of vessels; in other words, only distress involving the vessel as a whole is recorded as a "Maritime Incident" (Canadian Coast Guard, 2001).

Maritime incidents are sub-classified according to the level of their severity as follows (Canadian Coast Guard, 2000):

- M4- False alarms and hoaxes: Situations that cause the SAR system to react which proves to be unjustified or fabricated, such as a mistaken report of a flare.
- M3- Incidents resolved in the uncertainty phase (Non-Distress): No distress or perceived appreciable risk to life apparent. (General calls for assistance). An Uncertainty phase exists when:

  1. There is doubt regarding the safety of a vessel or the person on board;

  2. A vessel has been reported overdue at destination; or

  3. A vessel has failed to make an expected position report.

- M2- Potential Distress incidents: The potential exists for a distress incident if timely action is not taken; i.e., immediate responses are required to stabilize a situation in order to prevent distress. This incident exist when:
    1. There is apprehension regarding the safety of a vessel or the person on board;
    2. Following the uncertainty phase, attempts to establish contact with the vessel have failed and inquiries addressed to the other appropriate sources have been unsuccessful; or
    3. Information has been received indicating that the operational efficiency of a vessel is impaired but not to the extent of being a distress situation.
- M1-Distress incidents: Distress phase exists when:
    1. A vessel or a person is threatened by grave and imminent danger and requires immediate assistance (Life-threatening situation was judged to be present or close at hand at some point during the incident);
    2. Following the previous phase, further unsuccessful attempts to establish contact with the vessel and more widespread unsuccessful inquiries point to the high probability that the vessel is in distress; or
    3. Information is received which indicates that the operating efficiency of the vessel has been impaired to the extent that a distress situation is very likely.

Incidents classified as M4 do not create any real demand on SAR resources, so they can be excluded from the analysis from the viewpoint of this research.

Table 6.1 shows the distribution of different classes of fishing incidents in year 2010.

Table 6.1. Fishing Incidents severity distribution-Year 2010

Type	Number	Percentage
M1	82	8%
M2	129	12%
M3	827	80%

The small numbers of M1 and M2 incidents, compared to M3 incidents, can reduce the robustness of statistical analyses. One way to overcome this problem is to group M1 and M2 distress incidents into a single category of serious consequence events (Severity=1) and assign Severity=0 to the remaining ones (i.e. M3). Another reason to aggregate M1 and M2 is that the Canadian Coast Guard defines the emergency level for each incident when assigning a SAR unit to help. Emergency levels are Distress (for M1 and M2 types) and Non-Distress (M3 types).

Since the test cases for this project span one year (year 2010), even grouping M1 and M2 for fishing will lead to very small number of incidents (211) compared to the number of traffic grids (3524) and consequently potentially biased results. Should the methods in this study be applied to a larger dataset (i.e. more years), one could carry out all the mentioned methods but substituting the number of incidents with the number of Distress and Non-Distress incidents and determine the hot spots of severe incidents with respect to traffic or the correlations between the severity levels of incidents and traffic in different regions. Table 6.2 summarizes methods and required variables for spatial analysis of consequences of incidents with respect to traffic.

Table 6.2. Methods and needed variables for spatial analysis of consequences of incidents with respect to traffic.

Moran's I test	Moran's I index for Distress Incidents	
Risk-Adjusted Nearest Neighbour Clustering Method	**Primary Variable:** Number of Distress Incidents in each Grid	**Secondary Variable:** Amount of Traffic in each Grid
Dual-Kernel Estimation	**Primary Variable:** Number of Distress Incidents in each Grid	**Secondary Variable:** Amount of Traffic in each Grid
Spatial Regression	**Dependent Variable:** Number of Distress Incidents in each Grid	**Predictor:** Amount of Traffic in each Grid
	Dependent Variable: Distress/Non-Distress	**Predictor:** Amount of Traffic in grids which have at least one incident happened

References

- Bartier, P.M., & Keller, C.P. (1996). Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW). Computers & Geosciences, 22(7), 795-799.
- Brunsdon, C., Fotheringham, A.S., & Charlton, M.E. (1996). Geographically weighted regression: a method for exploring spatial nonstationarity.Geographical analysis, 28(4), 281-298.
- Canadian Coast Guard (2000). National Search and Rescue Manual: Emergency Services Resources. Canada.
- Canadian Coast Guard (2001). Safety and Environmental Response Systems: Maritime Search And Rescue Incidents Annual Report 2001. Canada.
- Cliff, A.D., & Ord, J.K. (1973). Spatial autocorrelation (Vol. 5). London: Pion.
- Cliff, A.D., & Ord, J.K. (1981). Spatial processes: models & applications (Vol. 44). London: Pion.
- Cressie, N. (1990). The origins of kriging. Mathematical Geology, 22(3), 239-252.
- Etienne, L., Pelot, R., & Engler, C. (2013). Analysis of Marine Traffic along Canada's Coasts – Phase 2, Part 2: A spatio-temporal simulation model for forecasting marine traffic in the Canadian Arctic in 2020. DRDC CORA CR2013-214, March 2013.
- Engler, C., & Pelot, R. (2012). Analysis of Marine Traffic along Canada's Coasts – Phase 2, Part 1: Factors Influencing Arctic Traffic. DRDC CORA CR 2012-065, August 2012.
- Hemyari, P., & Nofziger, D.L. (1987). Analytical solution for punctual kriging in one dimension. Soil Science Society of America Journal, 51(1), 268-269.

- Hilliard, R.C., & Pelot, R., (2012). Analysis of Marine Traffic along Canada's Coasts – Phase 1. DRDC CORA CR 2012-060, Jan 2012.
- Levine, N. (2002). CrimeStat II. A spatial statistics program for the analysis of crime incident locations. Ned Levine and Associates and the National Institute of Justice, Washington, DC, USA.
- Moran, P.A.P. (1950). Notes on continuous stochastic phenomena. Biometrika, 37:17
- Nicholson, A. (1998). Analysis of spatial distributions of accidents. Safety science, 31(1), 71-91.
- Shahrabi, J. (2006) Spatial and temporal analyses of maritime fishing and shipping traffic and incidents. Doctoral Thesis. Dalhousie University.
- Schumaker, L.L. (1981). Spline functions: basic theory (Vol. 1981). New York: Wiley.
- Team, R.C. (2005). R: A language and environment for statistical computing. R Foundation for Statistical Computing.

7 Appendix A

The following section shows the results of Risk-Adjusted Clustering for Fishing incidents (Yearly and Quarterly) and Shipping incidents (Yearly, with and without the Arctic region) with respect to traffic.

Table 77.1. Risk-Adjusted Nearest Neighbour-Fishing Incidents-Year 2010

Order	Cluster	Mean X	Mean Y	Rotation	X-Axis	Y-Axis	Area	Points	Density
1	1	-66.49063	70.46214	48.29669	1566.42600	164.58319	809925.79313	667	0.000824
1	2	-49.84792	59.55833	12.55524	15.04676	122.49080	5790.23856	240	0.041449
1	3	-62.09127	62.01720	0.00000	0.10000	0.10000	0.03142	189	6016.056760
1	4	-57.14130	58.38043	0.00000	0.10000	0.10000	0.03142	46	1464.225455
1	5	-59.71875	57.40625	0.00000	0.10000	0.10000	0.03142	16	509.295810
1	6	-57.56818	56.70455	0.00000	22.14994	0.10000	6.95861	11	1.580776
1	7	-63.00000	58.83333	0.00000	0.10000	0.10000	0.03142	12	381.971858
1	8	-53.09211	57.25000	0.00000	0.10000	0.10000	0.03142	19	604.788775
1	9	-60.52273	56.43182	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	10	-51.01923	56.48077	0.00000	0.10000	0.10000	0.03142	13	413.802846
1	11	-49.70455	55.29545	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	12	-46.00000	55.09211	0.00000	0.10000	0.10000	0.03142	38	1209.577550
1	13	-42.45833	57.04167	0.00000	0.10000	0.10000	0.03142	12	381.971858
1	14	-42.97727	54.65909	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	15	-46.25000	53.36765	0.00000	0.10000	0.10000	0.03142	17	541.126799
1	16	-43.58333	53.25000	0.00000	0.10000	0.10000	0.03142	15	477.464822
1	17	-45.61364	52.02273	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	18	-42.90000	52.15000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	19	-44.02273	51.11364	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	20	-44.52273	49.97727	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	21	-43.00000	46.75000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	22	-44.17308	45.71154	0.00000	0.10000	0.10000	0.03142	13	413.802846
1	23	-44.38043	44.03261	0.00000	0.10000	0.10000	0.03142	23	732.112727
1	24	-46.07075	42.14623	0.00000	0.10000	0.10000	0.03142	53	1687.042372
1	25	-130.15000	54.65000	11.87030	13.01952	39.60491	1619.92065	10	0.006173
1	26	-50.25000	42.29545	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	27	-48.63462	40.98077	0.00000	0.10000	0.10000	0.03142	13	413.802846
1	28	-43.11000	40.75000	0.00000	0.10000	0.10000	0.03142	25	795.774704
1	29	-132.16667	54.20833	38.79884	39.34176	21.71581	2683.98283	12	0.004471
1	30	-46.31667	39.95000	0.00000	0.10000	0.10000	0.03142	15	477.464822
1	31	-73.88636	38.88636	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	32	-136.96154	52.71154	0.00000	0.10000	0.10000	0.03142	26	827.605692
1	33	-135.46875	51.40625	0.00000	0.10000	0.10000	0.03142	16	509.295810
1	34	-134.18333	50.41667	0.00000	0.10000	0.10000	0.03142	15	477.464822
1	35	-132.21429	50.03571	0.00000	0.10000	0.10000	0.03142	14	445.633834
1	36	-130.63462	47.98077	0.00000	0.10000	0.10000	0.03142	13	413.802846
1	37	-74.30000	37.20000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	38	-74.25000	36.00000	0.00000	0.10000	0.10000	0.03142	10	318.309881
2	1	-45.45291	54.04731	56.79162	225.10872	148.62554	105107.97231	10	0.000095
2	2	-57.27014	57.79496	78.29699	114.45521	258.41197	92917.61834	7	0.000075
2	3	-46.13365	42.26666	38.79766	177.23268	236.80059	131848.88337	7	0.000053
2	4	-133.52410	52.23808	27.76505	142.62459	227.06685	101741.47631	6	0.000059

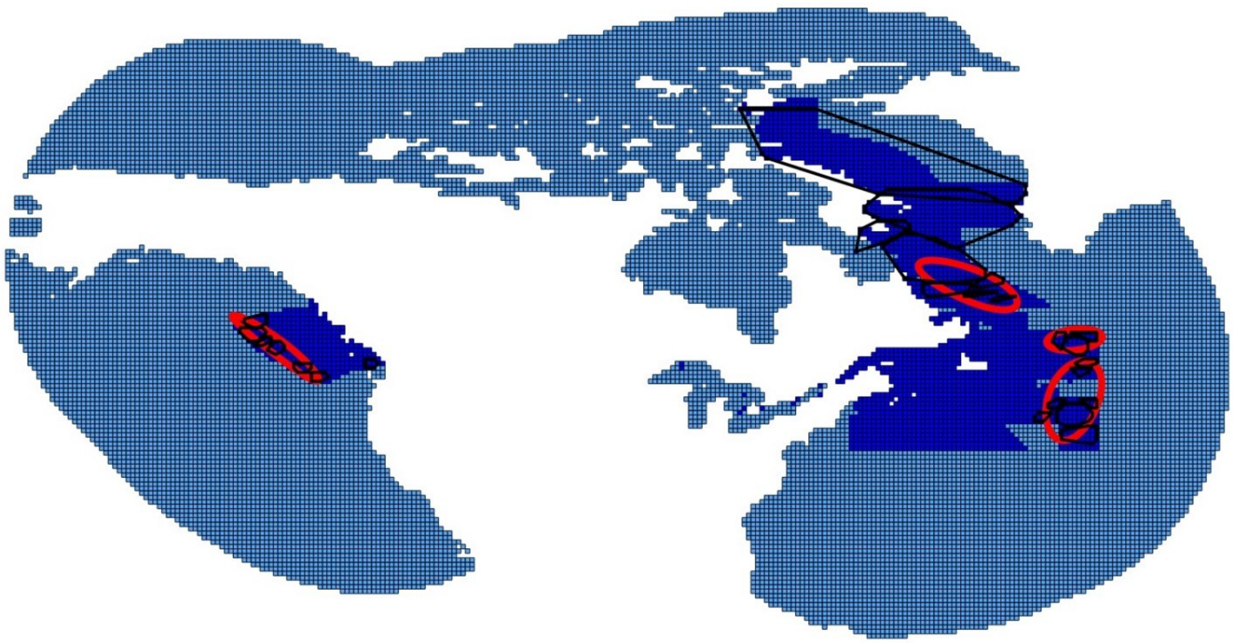

Figure 7.1. Risk-Adjusted Nearest Neighbour-Fishing Incidents-Year 2010

Table 77.2. Risk-Adjusted Nearest Neighbour-Fishing Incidents-1st Quarter-2010

Displaying 28 ellipse(s) starting from 1

Order	Cluster	Mean X	Mean Y	Rotation	X-Axis	Y-Axis	Area	Points	Density
1	1	-59.79942	64.42442	0.00000	0.10000	0.10000	0.03142	344	10949.859923
1	2	-55.60938	57.16295	0.00000	0.10000	0.10000	0.03142	224	7130.141345
1	3	-57.65909	54.72727	0.00000	0.10000	0.10000	0.03142	44	1400.563479
1	4	-63.57143	59.96429	0.00000	0.10000	0.10000	0.03142	14	445.633834
1	5	-51.01190	53.11905	0.00000	0.10000	0.10000	0.03142	42	1336.901502
1	6	-55.47222	52.47222	0.00000	0.10000	0.10000	0.03142	18	572.957787
1	7	-46.45270	53.27703	0.00000	0.10000	0.10000	0.03142	37	1177.746561
1	8	-52.39706	51.63235	0.00000	0.10000	0.10000	0.03142	17	541.126799
1	9	-49.39286	51.03571	0.00000	0.10000	0.10000	0.03142	14	445.633834
1	10	-51.38636	50.52273	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	11	-49.21429	49.50000	0.00000	0.10000	0.10000	0.03142	14	445.633834
1	12	-48.77381	44.86905	0.00000	0.10000	0.10000	0.03142	42	1336.901502
1	13	-48.52273	47.99242	0.00000	0.10000	0.10000	0.03142	33	1050.422609
1	14	-51.16667	43.38889	0.00000	0.10000	0.10000	0.03142	18	572.957787
1	15	-48.41667	42.87500	0.00000	0.10000	0.10000	0.03142	12	381.971858
1	16	-51.15000	45.50000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	17	-138.00000	53.29167	0.00000	0.10000	0.10000	0.03142	12	381.971858
1	18	-52.43182	42.52273	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	19	-136.75000	52.25000	0.00000	0.10000	0.10000	0.03142	17	541.126799
1	20	-53.55000	41.60000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	21	-74.55000	36.45000	0.00000	0.10000	0.10000	0.03142	15	477.464822
1	22	-136.00000	51.00000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	23	-132.40385	49.25000	0.00000	0.10000	0.10000	0.03142	13	413.802846
1	24	-134.15000	49.80000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	25	-130.29545	47.34091	0.00000	0.10000	0.10000	0.03142	11	350.140870
2	1	-51.27880	51.58653	43.34360	272.50187	151.30334	129529.25652	9	0.000069
2	2	-50.91483	43.45928	44.76314	113.11605	196.58551	69859.52114	6	0.000086
2	3	-134.59988	50.48876	49.40286	308.42368	27.60138	26744.12880	6	0.000224

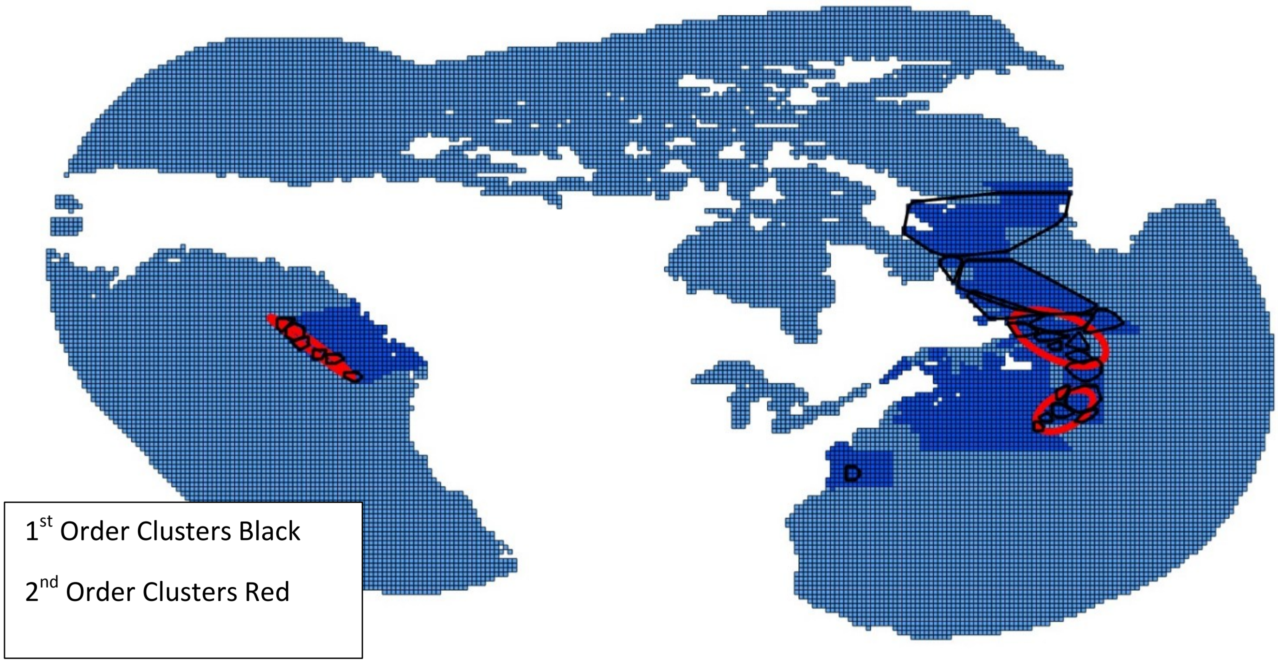

Figure 7.2. Risk-Adjusted Nearest Neighbour-Fishing Incidents-1st Quarter-2010

Table 77.3. Risk-Adjusted Nearest Neighbour-Fishing Incidents-2nd Quarter-2010

Order	Cluster	Mean X	Mean Y	Rotation	X-Axis	Y-Axis	Area	Points	Density
1	1	-59.87425	64.48653	0.00000	10.98897	0.10000	3.45229	334	96.747501
1	2	-59.32353	58.77941	0.00000	0.10000	0.10000	0.03142	34	1082.253597
1	3	-56.68939	57.03788	0.00000	0.10000	0.10000	0.03142	33	1050.422609
1	4	-58.89706	56.04412	0.00000	0.10000	0.10000	0.03142	17	541.126799
1	5	-54.04412	56.25000	0.00000	0.10000	0.10000	0.03142	17	541.126799
1	6	-61.25000	56.66667	0.00000	0.10000	0.10000	0.03142	12	381.971858
1	7	-43.23077	51.65385	0.00000	0.10000	0.10000	0.03142	26	827.605692
1	8	-51.85000	56.81000	0.00000	0.10000	0.10000	0.03142	25	795.774704
1	9	-45.91129	42.10484	0.00000	0.10000	0.10000	0.03142	93	2960.281898
1	10	-46.53571	51.80357	0.00000	0.10000	0.10000	0.03142	28	891.267668
1	11	-50.44231	55.05769	0.00000	0.10000	0.10000	0.03142	26	827.605692
1	12	-46.75000	53.59000	0.00000	0.10000	0.10000	0.03142	25	795.774704
1	13	-44.97222	50.25000	0.00000	0.10000	0.10000	0.03142	18	572.957787
1	14	-48.85000	54.35000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	15	-43.51190	43.91667	0.00000	0.10000	0.10000	0.03142	21	668.450751
1	16	-46.50000	45.91667	0.00000	0.10000	0.10000	0.03142	12	381.971858
1	17	-47.11364	44.70455	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	18	-50.79167	42.41667	0.00000	0.10000	0.10000	0.03142	12	381.971858
1	19	-49.39583	40.20833	0.00000	0.10000	0.10000	0.03142	24	763.943716
1	20	-43.44565	40.29348	0.00000	0.10000	0.10000	0.03142	23	732.112727
1	21	-134.87353	52.27941	0.00000	0.10000	0.10000	0.03142	85	2705.633993
1	22	-131.25000	54.30556	12.84844	18.21229	57.22684	3274.26688	27	0.008246
1	23	-130.71875	52.40625	0.00000	0.10000	0.10000	0.03142	16	509.295810
1	24	-132.46053	49.77632	0.00000	0.10000	0.10000	0.03142	38	1209.577550
1	25	-128.64474	52.69737	60.54757	85.17284	18.80688	5032.31537	19	0.003776
1	26	-130.06081	47.97973	0.00000	0.10000	0.10000	0.03142	37	1177.746561
1	27	-129.66667	50.83333	0.00000	0.10000	0.10000	0.03142	12	381.971858
1	28	-125.23333	49.23333	56.22438	65.31010	44.88610	9209.62865	30	0.003257
1	29	-127.71429	51.00000	57.58836	59.33566	27.84489	5190.52407	14	0.002697
1	30	-127.48333	47.98333	0.00000	0.10000	0.10000	0.03142	15	477.464822
2	1	-129.81060	50.84946	57.68938	240.98645	148.25476	112240.90205	10	0.000089
2	2	-55.16830	56.37447	22.52554	286.59269	95.27171	85778.59732	8	0.000093
2	3	-46.66714	42.79446	27.38036	212.53648	236.94430	158208.43515	7	0.000044
2	4	-45.37218	51.82435	60.94451	179.83165	94.67807	53489.11513	4	0.000075

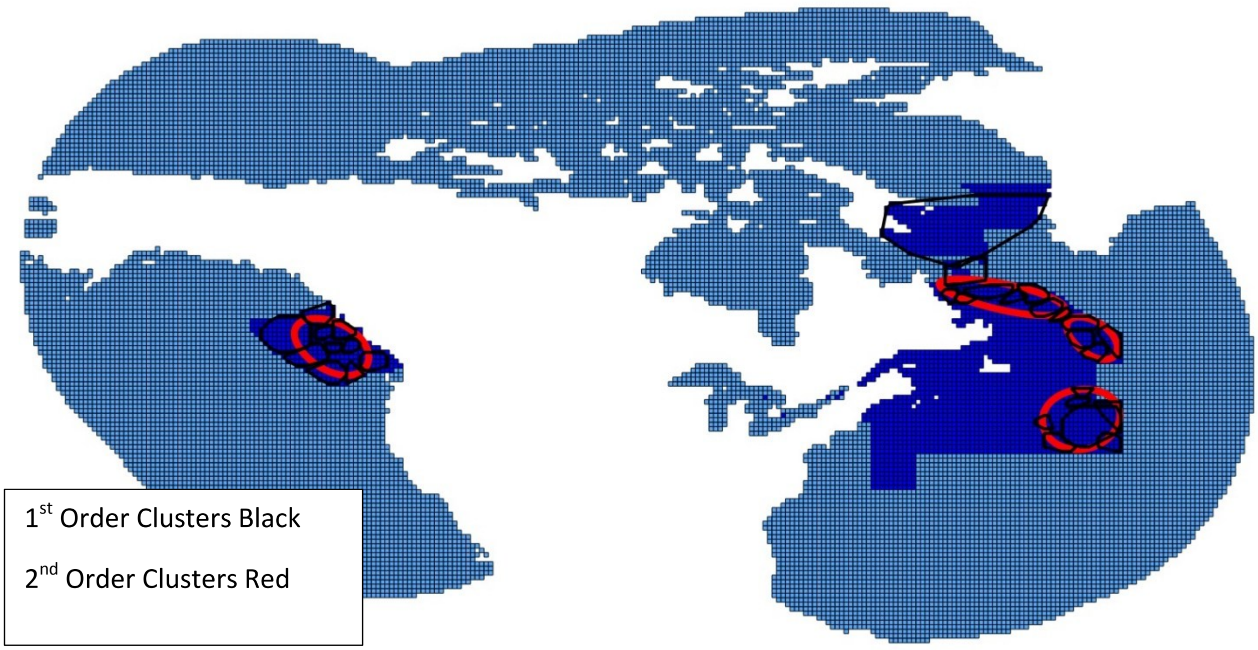

Figure 7.3. Risk-Adjusted Nearest Neighbour-Fishing Incidents-2nd Quarter-2010

Table 7.4. Risk-Adjusted Nearest Neighbour-Fishing Incidents-3rd Quarter-2010

Displaying 29 ellipse(s) starting from 1

Order	Cluster	Mean X	Mean Y	Rotation	X-Axis	Y-Axis	Area	Points	Density
1	1	-68.75499	71.92465	0.00000	0.10000	0.10000	0.03142	501	15947.325062
1	2	-60.19141	66.05469	0.00000	0.10000	0.10000	0.03142	128	4074.366483
1	3	-59.88450	61.94006	0.00000	0.10000	0.10000	0.03142	171	5443.098973
1	4	-47.67672	59.15517	0.00000	0.10000	0.10000	0.03142	116	3692.394625
1	5	-66.20833	63.73611	0.00000	0.10000	0.10000	0.03142	36	1145.915573
1	6	-63.28125	60.09375	0.00000	0.10000	0.10000	0.03142	16	509.295810
1	7	-58.99074	59.12037	0.00000	0.10000	0.10000	0.03142	27	859.436680
1	8	-55.12500	58.87500	0.00000	0.10000	0.10000	0.03142	28	891.267668
1	9	-51.97727	57.70455	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	10	-48.11364	56.78030	0.00000	0.10000	0.10000	0.03142	33	1050.422609
1	11	-43.32143	57.34524	0.00000	0.10000	0.10000	0.03142	21	668.450751
1	12	-52.02273	56.79545	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	13	-44.30000	55.30000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	14	-45.21341	41.45732	0.00000	0.10000	0.10000	0.03142	41	1305.070514
1	15	-43.12500	40.21875	0.00000	0.10000	0.10000	0.03142	16	509.295810
1	16	-43.46053	43.06579	0.00000	0.10000	0.10000	0.03142	19	604.788775
1	17	-47.91667	40.45000	0.00000	0.10000	0.10000	0.03142	15	477.464822
1	18	-47.11364	42.75000	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	19	-49.56818	40.43182	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	20	-43.71667	44.58333	0.00000	0.10000	0.10000	0.03142	15	477.464822
1	21	-135.25000	53.75000	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	22	-136.40385	51.75000	0.00000	0.10000	0.10000	0.03142	13	413.802846
1	23	-137.20000	52.90000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	24	-67.50000	37.25000	0.00000	0.10000	0.10000	0.03142	12	381.971858
1	25	-134.33333	49.79167	0.00000	0.10000	0.10000	0.03142	12	381.971858
1	26	-131.25000	48.25000	0.00000	0.10000	0.10000	0.03142	11	350.140870
2	1	-54.63398	58.80808	25.85195	338.12149	120.42936	127924.88613	8	0.000063
2	2	-45.73058	41.85100	52.75159	140.37992	228.16812	100625.91278	7	0.000070
2	3	-134.88744	51.28833	59.11626	292.32971	80.95394	74346.55441	5	0.000067

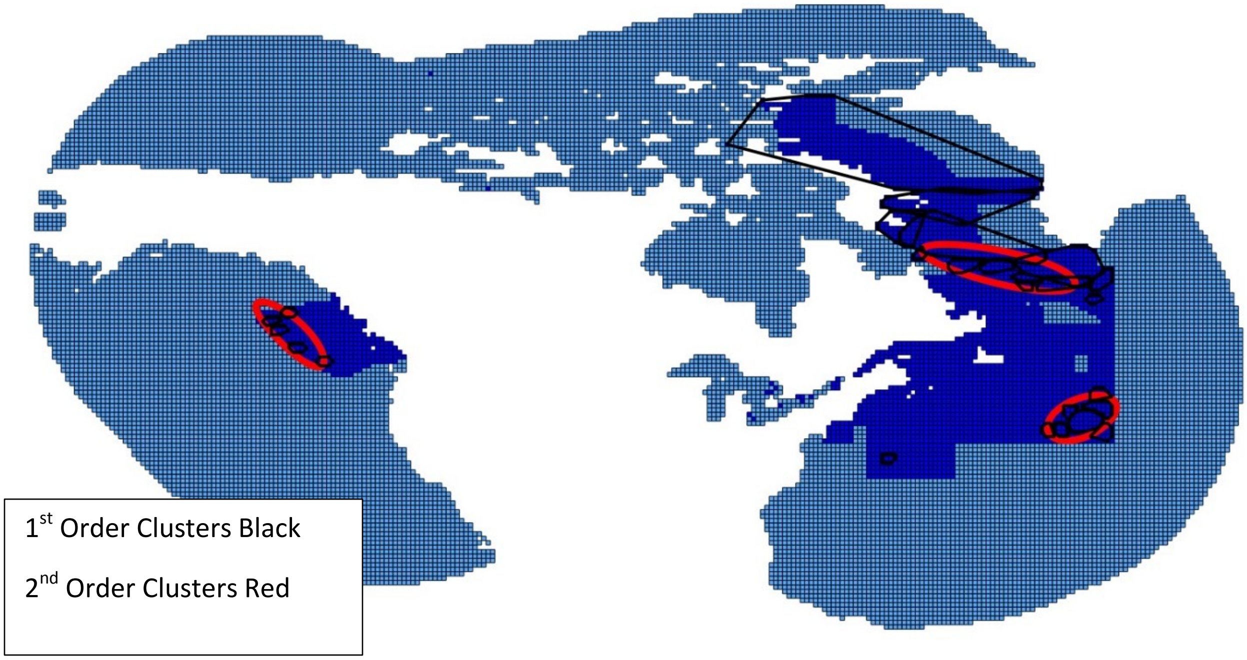

Figure 7.4. Risk-Adjusted Nearest Neighbour-Fishing Incidents-3rd Quarter-2010

Table 7.5. Risk-Adjusted Nearest Neighbour-Fishing Incidents-4th Quarter-2010

Displaying 29 ellipse(s) starting from 1

Order	Cluster	Mean X	Mean Y	Rotation	X-Axis	Y-Axis	Area	Points	Density
1	1	-69.14116	72.04252	0.00000	0.10000	0.10000	0.03142	441	14037.465773
1	2	-59.31227	65.71520	0.00000	0.10000	0.10000	0.03142	273	8689.859764
1	3	-60.71667	60.32576	0.00000	0.10000	0.10000	0.03142	165	5252.113044
1	4	-65.97500	63.52500	5.27600	0.23522	122.59215	90.59302	40	0.441535
1	5	-59.23750	57.03750	0.00000	0.10000	0.10000	0.03142	40	1273.239526
1	6	-54.17857	57.75000	0.00000	0.10000	0.10000	0.03142	21	668.450751
1	7	-55.15000	56.48333	0.00000	0.10000	0.10000	0.03142	15	477.464822
1	8	-52.75000	56.00000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	9	-44.79651	42.98256	0.00000	0.10000	0.10000	0.03142	43	1368.732490
1	10	-44.57000	50.75000	0.00000	0.10000	0.10000	0.03142	25	795.774704
1	11	-43.32692	51.94231	0.00000	0.10000	0.10000	0.03142	13	413.802846
1	12	-46.43182	51.52273	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	13	-44.05645	40.81452	0.00000	0.10000	0.10000	0.03142	31	986.760633
1	14	-43.34091	48.88636	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	15	-48.45000	42.80000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	16	-136.37500	53.50000	0.00000	0.10000	0.10000	0.03142	16	509.295810
1	17	-46.85000	44.25000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	18	-136.85526	52.06579	0.00000	0.10000	0.10000	0.03142	19	604.788775
1	19	-42.95000	44.40000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	20	-44.05000	47.95000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	21	-135.45000	51.10000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	22	-134.06818	50.52273	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	23	-123.65000	48.75000	75.24875	47.84140	20.45189	3073.88179	10	0.003253
1	24	-129.32143	47.21429	0.00000	0.10000	0.10000	0.03142	14	445.633834
1	25	-131.25000	48.32692	0.00000	0.10000	0.10000	0.03142	13	413.802846
2	1	-44.92770	44.58335	12.66190	140.38886	314.76002	138823.21686	7	0.000050
2	2	-133.88665	50.45495	50.57477	328.11963	47.07246	48523.14679	6	0.000124
2	3	-56.40655	57.51932	42.36606	260.80035	103.16265	84524.09425	5	0.000059
2	4	-44.77625	51.40501	77.22098	79.86768	136.41428	34227.94304	3	0.000088

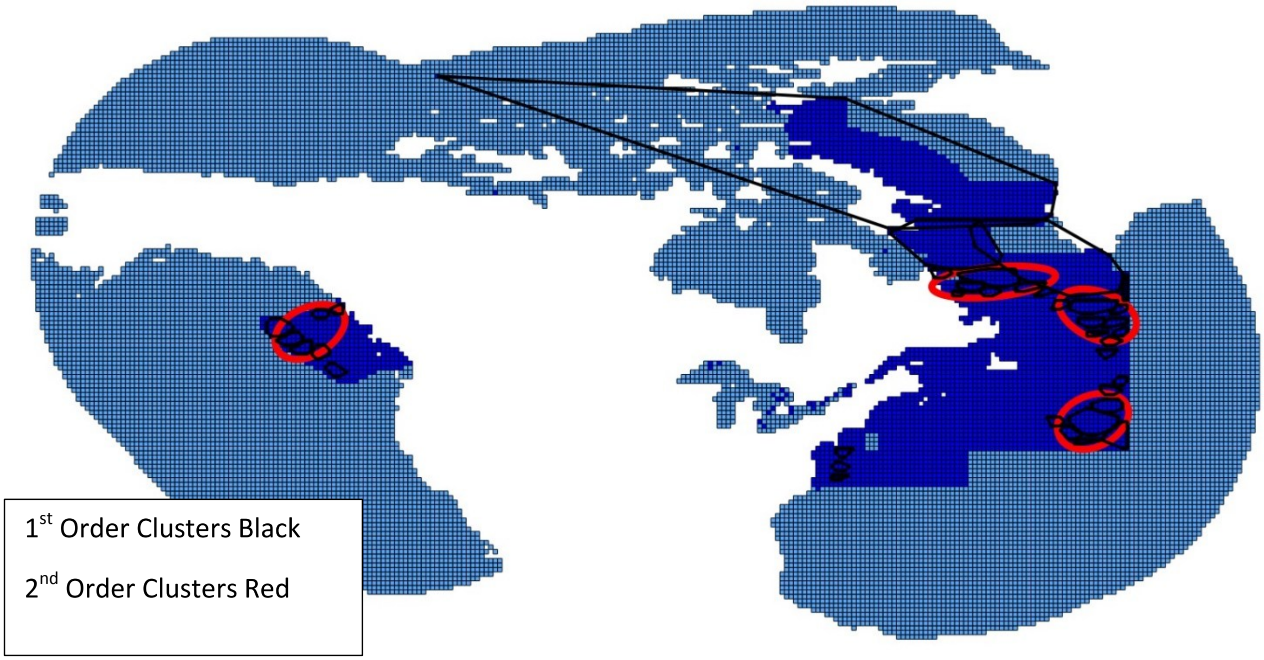

Figure 7.5. Risk-Adjusted Nearest Neighbour-Fishing Incidents-4th Quarter-2010

Table 77.6. Risk-Adjusted Nearest Neighbour-Shipping Incidents-Year 2010

Order	Cluster	Mean X	Mean Y	Rotation	X-Axis	Y-Axis	Area	Points	Density
1	1	-70.67500	79.62500	0.00000	0.10000	0.10000	0.03142	20	636.619763
1	2	-74.05392	73.42157	0.00000	0.10000	0.10000	0.03142	102	3246.760791
1	3	-82.29839	74.99194	0.00000	0.10000	0.10000	0.03142	31	986.760633
1	4	-84.60366	73.72561	0.00000	0.10000	0.10000	0.03142	41	1305.070514
1	5	-134.49390	71.54268	0.00000	0.10000	0.10000	0.03142	82	2610.141028
1	6	-68.43644	70.94492	0.00000	0.10000	0.10000	0.03142	59	1878.028301
1	7	-92.48864	74.14773	0.00000	0.10000	0.10000	0.03142	44	1400.563479
1	8	-133.52778	70.08333	0.00000	0.10000	0.10000	0.03142	27	859.436680
1	9	-64.18333	68.91667	0.00000	0.10000	0.10000	0.03142	30	954.929644
1	10	-127.17593	73.45370	0.00000	0.10000	0.10000	0.03142	27	859.436680
1	11	-127.13095	70.51190	0.00000	0.10000	0.10000	0.03142	21	668.450751
1	12	-62.36905	67.22619	0.00000	0.10000	0.10000	0.03142	21	668.450751
1	13	-95.38636	72.38636	0.00000	0.10000	0.10000	0.03142	22	700.281739
1	14	-61.76613	65.44355	0.00000	0.10000	0.10000	0.03142	31	986.760633
1	15	-60.21000	64.01000	0.00000	0.10000	0.10000	0.03142	25	795.774704
1	16	-97.19048	69.69048	0.00000	0.10000	0.10000	0.03142	42	1336.901502
1	17	-57.51471	63.27941	0.00000	0.10000	0.10000	0.03142	17	541.126799
1	18	-55.70313	59.43750	0.00000	0.10000	0.10000	0.03142	32	1018.591621
1	19	-59.15323	61.75000	0.00000	0.10000	0.10000	0.03142	31	986.760633
1	20	-122.45588	70.10294	0.00000	0.10000	0.10000	0.03142	17	541.126799
1	21	-56.05556	61.19444	0.00000	0.10000	0.10000	0.03142	18	572.957787
1	22	-85.91667	62.25000	0.00000	0.10000	0.10000	0.03142	36	1145.915573
1	23	-111.21053	68.23684	0.00000	0.10000	0.10000	0.03142	38	1209.577550
1	24	-52.64474	59.38158	0.00000	0.10000	0.10000	0.03142	19	604.788775
1	25	-100.17500	68.47500	0.00000	0.10000	0.10000	0.03142	20	636.619763
1	26	-118.01667	69.98333	0.00000	0.10000	0.10000	0.03142	15	477.464822
1	27	-89.25000	60.57000	0.00000	0.10000	0.10000	0.03142	25	795.774704
1	28	-90.22059	62.95588	0.00000	0.10000	0.10000	0.03142	17	541.126799

1	29	-63.84091	63.79545	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	30	-61.75000	61.16667	0.00000	0.10000	0.10000	0.03142	12	381.971858
1	31	-106.91667	68.41667	0.00000	0.10000	0.10000	0.03142	12	381.971858
1	32	-58.89286	59.32143	0.00000	0.10000	0.10000	0.03142	14	445.633834
1	33	-92.75000	59.65625	0.00000	0.10000	0.10000	0.03142	16	509.295810
1	34	-55.21429	57.92857	0.00000	0.10000	0.10000	0.03142	14	445.633834
1	35	-80.77174	62.75000	0.00000	0.10000	0.10000	0.03142	23	732.112727
1	36	-64.33333	61.80556	0.00000	0.10000	0.10000	0.03142	18	572.957787
1	37	-76.16176	63.25000	0.00000	0.10000	0.10000	0.03142	17	541.126799
1	38	-74.34091	62.61364	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	39	-78.60294	59.89706	0.00000	0.10000	0.10000	0.03142	17	541.126799
1	40	-49.70000	57.85000	0.00000	0.10000	0.10000	0.03142	10	318.309881
1	41	-66.31250	60.71875	0.00000	0.10000	0.10000	0.03142	16	509.295810
1	42	-59.84091	57.15909	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	43	-68.25000	59.31667	0.00000	0.10000	0.10000	0.03142	15	477.464822
1	44	-69.06818	60.38636	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	45	-68.88636	61.47727	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	46	-70.65909	61.88636	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	47	-72.15909	62.38636	0.00000	0.10000	0.10000	0.03142	11	350.140870
1	48	-46.95000	57.90000	0.00000	0.10000	0.10000	0.03142	10	318.309881
2	1	-61.15023	61.63241	16.30633	214.81528	155.72709	105094.31592	11	0.000105
2	2	-71.77317	61.63990	45.73831	223.70588	42.10680	29592.35439	6	0.000203
2	3	-51.12726	58.26504	7.82632	228.03466	84.31541	60402.88729	4	0.000066
2	4	-127.13352	70.94632	15.66804	233.95770	139.25771	102354.39162	6	0.000059
2	5	-85.76619	73.73464	81.09567	100.45647	274.78360	86719.86478	5	0.000058
2	6	-62.77284	67.19547	74.97091	248.31277	17.81941	13900.87747	3	0.000216
2	7	-81.76378	61.63235	37.63351	293.14023	139.65195	128609.27888	3	0.000023
2	8	-90.74020	61.06071	17.68896	102.26724	244.80666	78651.97649	3	0.000038
3	1	-61.70587	62.18320	38.61170	563.82792	343.54498	608527.29098	4	0.000007

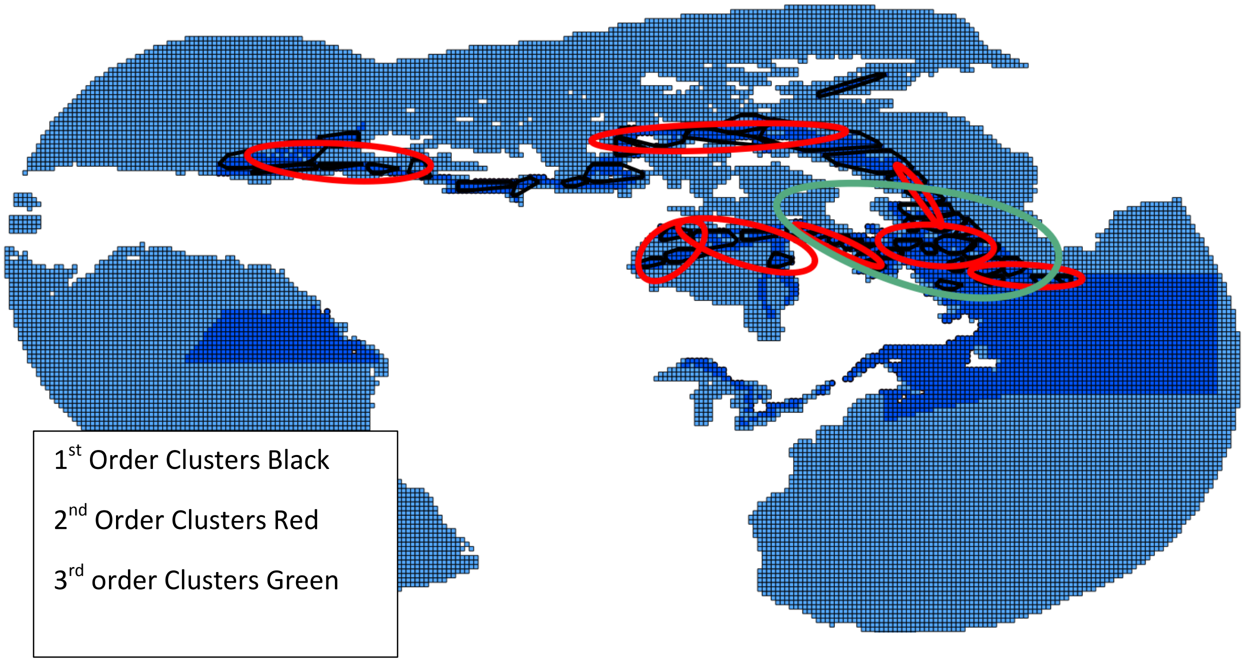

Figure 7.6. Risk-Adjusted Nearest Neighbour-Shipping Incidents-Year 2010

Table 7.7. Risk-Adjusted Nearest Neighbour-Shipping Incidents-Without Arctic Region-Year 2010

Order	Cluster	Mean X	Mean Y	Rotation	X-Axis	Y-Axis	Area	Points	Density
1	1	-129.41667	53.91667	0.00000	70.72447	0.10000	22.21875	6	0.270042
1	2	-131.00000	54.00000	34.06604	13.10987	22.73841	936.50105	6	0.006407
1	3	-133.10714	53.96429	0.00000	0.10000	0.10000	0.03142	7	222.816917
1	4	-85.66667	47.66667	0.00000	0.10000	0.10000	0.03142	6	190.985929
1	5	-33.25000	45.75000	0.00000	0.10000	0.10000	0.03142	5	159.154941
1	6	-32.12500	45.75000	0.00000	0.10000	0.10000	0.03142	4	127.323953
1	7	-87.25000	48.37500	0.00000	0.10000	0.10000	0.03142	4	127.323953
1	8	-88.37500	48.62500	0.00000	0.10000	0.10000	0.03142	4	127.323953
1	9	-55.75000	52.37500	0.00000	0.10000	0.10000	0.03142	4	127.323953
1	10	-33.25000	52.37500	0.00000	0.10000	0.10000	0.03142	4	127.323953