Visualization of Differences across Port State Control Regimes by means of Correspondence Analysis

Sabine Knapp[1] and Michel van de Velden
Econometric Institute, Erasmus University Rotterdam

Econometric Institute Report 2007-32

Abstract

This article is based on a combined dataset of 183,819 port state control (PSC) inspections from various port state control regimes around the world. In theory, there should be no differences in treatments of inspections to determine if a vessel is substandard or not no matter where the inspection was performed. This article uses correspondence analysis in order to investigate whether there are differences and if so, visualizes them for easier interpretation. The results of this analysis confirm that treatment of vessels across port state control regimes differ although some groupings of regimes across the ownership groups can be found. Association can be found between the Paris MoU and the Viña del Mar Agreement and between the Caribbean MoU and the Indian Ocean MoU while the USCG and AMSA are always apart from the rest. With respect to the distribution of deficiency codes towards regimes, only areas such as structural safety, radio communications or safety and cargo operational areas show similarities while ISM related deficiencies only show strong association with one regime. The differences clearly show that there is room for harmonization across port state control inspections in all inspection areas. Harmonization could be achieved by increasing cooperation amongst the regimes. A good starting point would be the acceleration of the harmonization of inspection procedures, combined training of port state control officers and the use of combined datasets across regimes, in particular in the concept of the development of the Global Integrated Ship Information System (GISIS) of the International Maritime Organization (IMO).

The authors would like to acknowledge the following port state control regimes for providing their inspection data: Paris MoU, Caribbean MoU, Viña del Mar Agreement on PSC, USCG, Indian MoU and the Australian Maritime Safety Authority (AMSA) and for Lloyd's Register Fairplay for providing information on ownership of a vessel.

[1] Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR, Rotterdam, The Netherlands, email: knapp@few.eur.nl or vandevelden@few.eur.nl. This working paper does not reflect the view of the European Maritime Safety Agency (EMSA).

1. Introduction

The maritime industry is regulated by a complicated legal framework based on international law which is characterized by a limited legal enforcement in case of non compliance of a flag state. The legal loopholes create distortion to competition. Flag states are to be seen as the first line of defense in eliminating sub-standard vessels followed by the second line of defense, the port states that perform port state control inspections (PSC). One continuous question of discussion within the industry is to determine possible areas of harmonization across port state control regimes.

Port state control allows port states to inspect a vessel calling a port under its jurisdiction. In theory, inspections should be carried out in a uniform way so that no structural and significant differences exist between treatments of vessels of the various port state control regimes, in particular the determination of a substandard vessel. However, in practice, this uniformity may not be achieved. Uniform inspections would imply that ships that are similar in some sense, receive similar inspection reports regardless of the inspecting authority. Little cooperation can be found between the port state control regimes with reference to combined training efforts, the sharing of inspection information in a centralized database such as the Global Integrated Ship Information System (GISIS) of the International Maritime Organization (IMO) for the benefit of the port state control officers, flag states or classification societies. Another topic of continuous discussion amongst IMO member states is the harmonization of targeting criteria across the regimes.

There are many ways to classify ships. We could, for example, use physical or functional properties of a vessel, such as the ship type, age, size or type of cargo, etc. Perhaps less specific, but potentially influential properties of the safety quality of a vessel is influenced by the owner and in second line, the registry. The relationship between these important characteristics and ship inspection data has been addressed by Knapp and Franses (2007). By using binary logistic regression, it was shown that the main differences across port state control regimes is due to the port states and the effects of the deficiency codes towards the probability of detention.

In this paper, we use the same dataset, which comprises inspection data from five different port state control regimes split into six datasets, to study the relationship between inspection data and ship ownership by using a correspondence analysis (CA). The aim of this paper is two-fold: 1) obtain better insight into the relationship between ownership of a vessel and inspection data by visualizing this relationship. 2) confirm that differences across port state control inspections exist, visualize and interpret these differences.

The port state control data used in this analysis is a unique combination of 183,819 port state control inspections within the time frame 1999 to 2004 from the following regimes: Paris Memorandum of Understanding (MoU), Indian Ocean MoU, Caribbean MoU, United States Coast Guard (USCG), Viña del Mar Agreement on Port State Control and the Australian Maritime Safety Authority (AMSA). For the purpose of this study, AMSA was treated as an individual country and apart from the Indian Ocean MoU although it is normally a member thereof. This is due to the fact that AMSA provided more detailed data of a longer time frame than the Indian Ocean MoU and it was felt that it would be better to keep both sets separate in order not to bias the results of the Indian Ocean MoU.

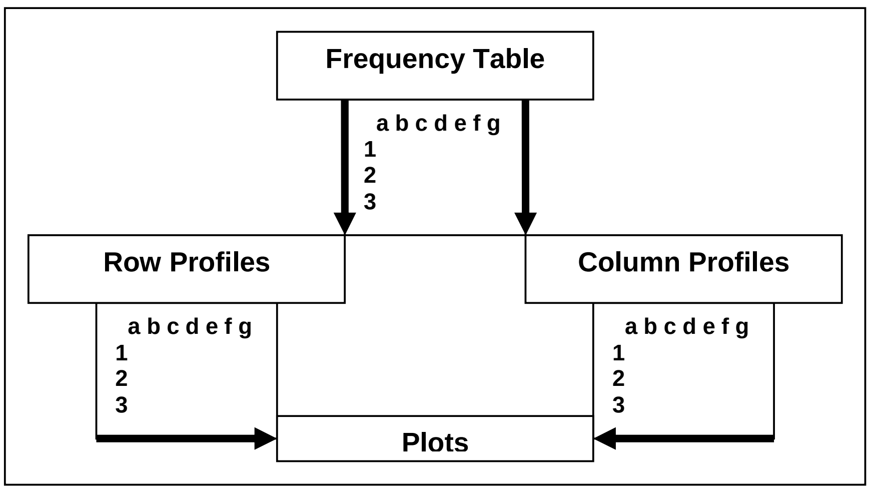
The paper is organized as follows. In Section 2, we present a brief overview of correspondence analysis and give an explanation of the variables that were used for the analysis. In Section 3, we look into the differences of port state control regimes with respect to the distribution over deficiencies found during inspections for each ownership group. We conclude the paper by summarizing our results in Section 4.

2. Short Overview of Correspondence Analysis and Variables Used

2.1. Description of Correspondence Analysis

Correspondence analysis (CA) is a powerful tool that allows a graphical representation of categorical variables. Typically, CA is applied to a contingency table. That is, a table with as its elements the counts of co-occurrences between categories of two variables. In CA, both rows and columns of such a contingency table are considered as points in high-dimensional space. A least-squares approximation is then obtained to approximate these high-dimensional points in a space of low (preferably two) dimensional space. This makes it possible to plot the data in such a way that relationships between categories and variables are exposed. In this paper we will not treat the mathematical properties and derivation of CA and refer the reader to excellent expositions available in, for example, Greenacre (1984) Jobson (1992) or Clausen (1998). The process of calculation contains various stages for each set of variable as illustrated in Figure 1.

Figure 1: Analytical Process of Correspondence Analysis



Source: Clausen, S. (1998), Applied Correspondence Analysis, page 4

The rows represent categories of the row-variable, and the columns represent categories of the column variable. Then, the entries of the table are the frequencies of co-occurrences of the two variables. If we divide the row elements of the contingency table by the corresponding row totals, we obtain so-called *relative frequencies* (or conditional proportions). These relative frequencies give the distributions of the categories of the row variable over the categories of the column variable.

In a similar way, we can obtain relative frequencies for the column variables. That is, instead of dividing the elements by the row totals, we divide the elements by column totals. We shall refer to the relative frequencies for the row and column variables as row- and column profiles respectively. The aim of CA is to depict both the row and column profiles in a low (typically two) dimensional space in such a way that this low-dimensional approximation "best" represents the original data. To achieve this "best" approximation, the row and column points are obtained by minimizing the so-called Chi-squared distances between the original profiles and their low-dimensional approximations.

There are several options for plotting the CA results. For example, separate plots can be made for the row and columns of the contingency table. In such plots, the distances between the points (representing the categories of one of the variables) are approximated Chi-squared distances. If these distances are large, it means that the profiles (i.e. the distribution over the categories of the other variable) differ considerably. On the other hand, if the points are close, the profiles are similar.

A nice feature of CA is the fact that the solution for the row and column profiles can be obtained simultaneously. Moreover, the row and column points are related to each other and plots can be constructed that represents both the rows and columns jointly. Such plots are called symmetrical plots in the literature. Although these joint plots are appealing and frequently used, some caution must be taken when interpreting them. In particular, it should be noted that distances between row and column categories are not defined. Hence, proximity of such points in the symmetrical map should not be interpreted as a distance. However, these proximities do give an indication of the correlations between the categories. That is, if row- and column points are located in the same direction, the correlation between the two categories is high. In other words, a relatively high number of observations in that row category, suggests that there was also a relatively high number of observations in the column category.

An important feature of CA is its ability to incorporate additional profiles into the map. Such profiles are called supplementary profiles. They are projected into the CA map as supplementary points in such a way that their relative position with respect to the other profiles is revealed. However, the points play no role in the determination of the map. They are therefore also sometimes referred to as *passive* points.

To assess the quality of a CA solution, we consider the percentage of total variance, or, as one calls it in CA, inertia, that is accounted for by the low dimensional solution. This explained inertia measure is similar to the explained variance in principal component analysis. Note that, the total inertia in a table is equal to the total number of observations times Pearson's Chi-squared statistic for testing independence in a contingency table.

The quality of individual points in CA can also be assessed by studying the so-called absolute and relative contributions. The absolute contributions of the points indicate how much a certain point contributed to the inertia of a dimension. If the absolute contribution is high, it means that the point played an important role in determining the direction of the axes. By looking at points with a higher than average (that is, a contribution greater than 1 divided by the number of points) absolute contribution, we can simplify our interpretation. On the other hand, relative contributions, give an indication of the quality of the representation of individual points. They indicate how well points are represented along a certain dimension. Adding the relative contributions of points over the displayed dimensions gives a measure for the quality of the point's representation in the low-dimensional CA map. For a more detailed discussion of the absolute and relative contributions discussions see Greenacre (1984) or Clausen (1998).

2.2. Variables Used in the Analysis

In this paper we use correspondence analysis to study the distribution of observed deficiencies categorized according to the main deficiency codes (C0100 to C2500) as listed in Table 1. The port state control regimes are as follows: Paris MoU (paris), Caribbean MoU (carib), Indian Ocean MoU (india), United States Coast Guard (usacg), Viña del Mar Agreement on Port State Control (vina) and the Australian Maritime Safety Authority (amsa). The data is presented with port state control regimes in the rows and the deficiency codes in columns.

The ships are further classified according to their ownership so that a separate plot can be obtained for each ownership groups. The reason for this splitting up according to ownership is to account for the different legal frameworks of the countries of locations of an owner and the different economic conditions within each group. This creates a playing field for the safety culture of an owner towards the operation of his/her vessels. For this analysis we determine ship ownership by considering the respective country/location of the beneficial owner[2]. Then, in order to decrease the amount of variables to a manageable size, the countries are grouped according to a classification for registries as developed by Alderton and Winchester (1999). In this analysis, the same grouping is applied for countries of location of the owner versus the registry and all open registries are further combined to one group.

Table 1: Description of Main Deficiency Codes

Code	Deficiency Code Description	Code	Deficiency Code Description
100	Ship's certificates and documents	1300	Mooring arrangements (ILO 147)
200	Crew certificates	1400	Propulsion & auxiliary engine
300	Accommodation	1500	Safety of navigation
400	Food and catering	1600	Radio communications
500	Working spaces and accident prev.	1700	MARPOL Annex I (Oil Pollution)
600	Life saving appliances	1800	Gas and chemical carriers
700	Fire Safety measures	1900	MARPOL Annex II (Noxious Liquids)
800	Accident prevention (ILO147)	2000	SOLAS Operational deficiencies
900	Structural Safety	2100	MARPOL related oper. deficiencies
1000	Alarm signals	2200	MARPOL Annex III (Pack.Harmf.Sub.)
1100	Cargoes	2300	MARPOL Annex V (Garbage)
1200	Load lines	2500	ISM related deficiencies

A separate analysis was performed for each ownership group because while we would expect to see differences between the various plots of the ownership groups, we would, at least ideally, not expect to see differences of the treatments of vessels by each regime – the treatment of the vessels with respect to deficiency codes. In our analysis, ownership is chosen for the split up of plots over the registry because it is felt that ownership is the most important variable to determine the safety quality of a vessel. It is further felt that given the same owner with a fleet registered under various flags, the safety quality of all vessels is expected to be the same across the fleet regardless of the flag.

The final grouping that is used in our analysis is as follows: traditional maritime nations (TMN), emerging maritime nations (EMN), open registries[3] (a combination of old open registries, international open registries and new open registries) and the category "unknown" for unknown ownership of a vessel.

The entries of the contingency matrix are the frequencies of occurrences of each deficiency code for each port state control regime. If we divide the observations by the row totals, we see that each row represents the distribution of the deficiency codes for a port state control regime. We are interested in these distributions of deficiency codes for the various regimes and we will visualize them by using correspondence analysis for each ownership group.

3. Differences of port state control inspections across regimes

As discussed previously, the plots present the characteristics of ships with respect to the distribution over deficiencies found during port state control inspections across regimes for each ownership group separately. Our hypothesis is that while differences between the plots of the

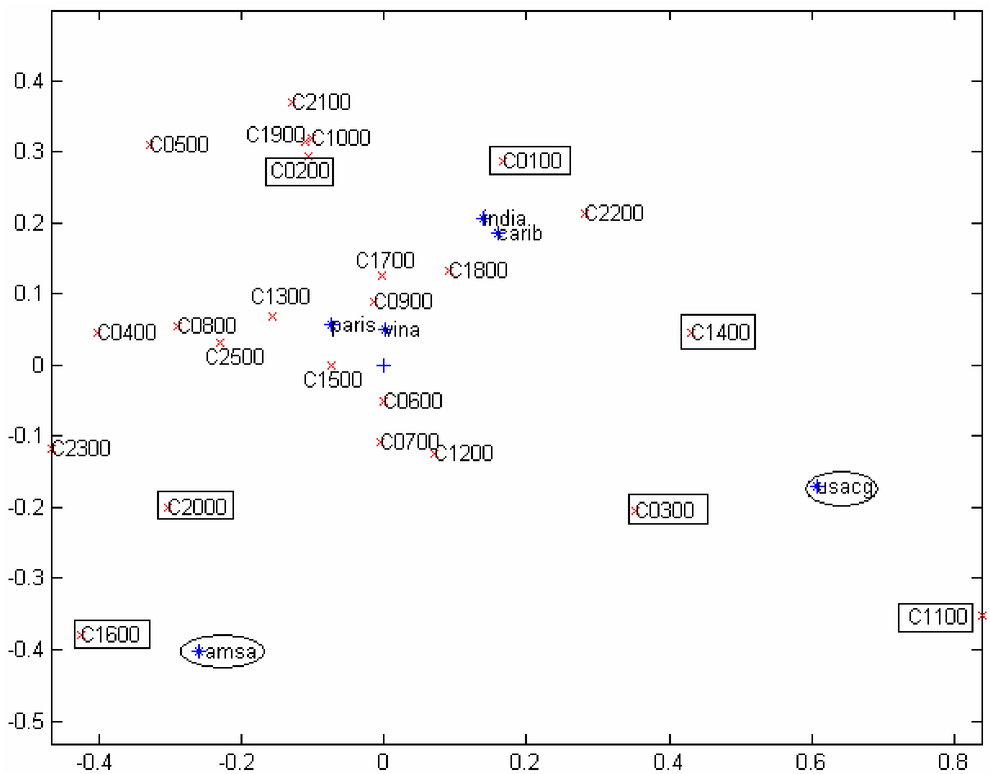
[2] Based on Lloyd's Register Fairplay data of the "World Shipping Encyclopedia CD" and Lloyd's "Maritime Database CD" plus custom made queries on ownership history

[3] Alderton and Winchester further splits up this group into new open registry, international open registries and old open registries but for the sake of simplicity and available observations per ownership group, this was combined into one group.

ownership groups are to be expected, the location of the regimes within each plot should not vary considerable with respect to the deficiency codes. The closer the regimes are located to each other, the more similar their treatment of vessels.

The plots are presented in Figure 2 to Figure 5[4]. To interpret the plots, we focus mainly on the codes that played an important role in the determination of the axes, that is, the points with relatively high absolute contributions. An absolute contribution is called high if it is larger than what could be expected if all points contribute equally to the inertia. The deficiency codes and port state control regimes with relatively high absolute contributions are highlighted in the plot by boxes (deficiency codes) and circles (regimes). Moreover, a summary of the findings of all plots can be found in Appendix A.

Figure 2: Traditional Maritime Nations, explained inertia 79.2%



Comparing each plot individually, we note the spread for the regimes of all plots. Ideally, there should not be such a pattern. That is, the distribution of deficiency codes for the regimes should not differ in a structural manner. Apparently, and in accordance with earlier findings of Knapp and Franses (2007), this equality does not exist. In addition, especially for the first three plots, the spread of the regimes and codes appears to be somewhat similar. This indicates that differences for the different ownerships are limited.

[4] Figure 5 is enlarged in order to facilitate interpretation. As a consequence, C2200 is no longer in the picture but its direction and coordinates are indicated in the plot

Figure 3: Emerging Maritime Nations, explained inertia 78.1%

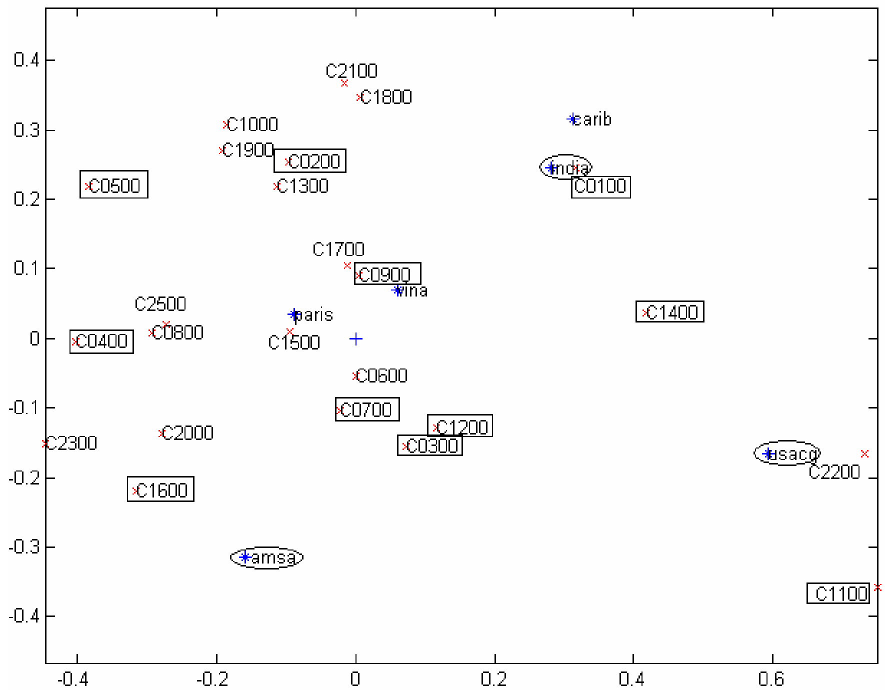


Figure 4: Open Registry Countries, explained inertia 75.6%

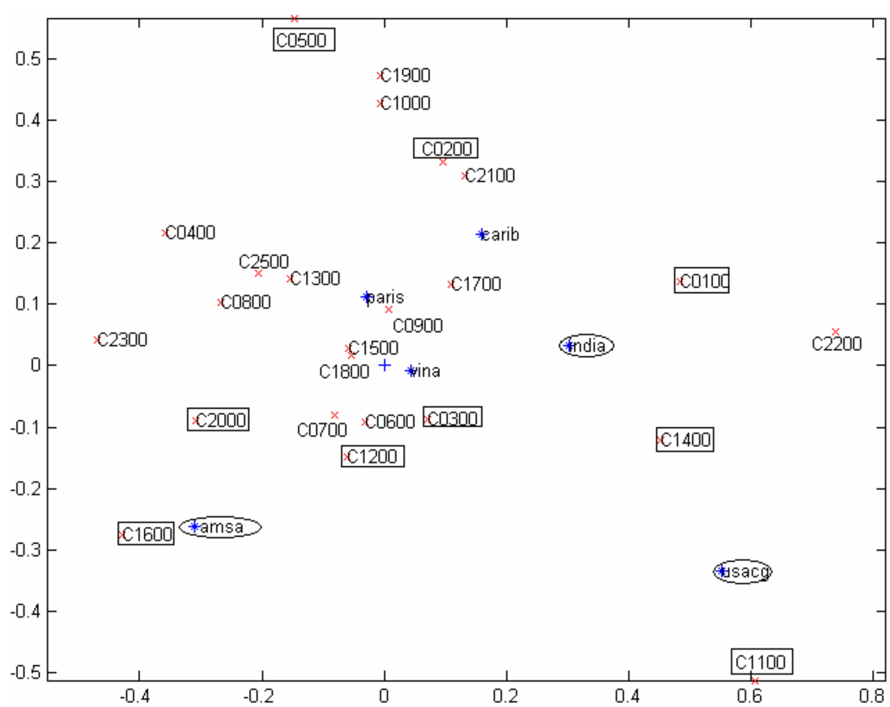
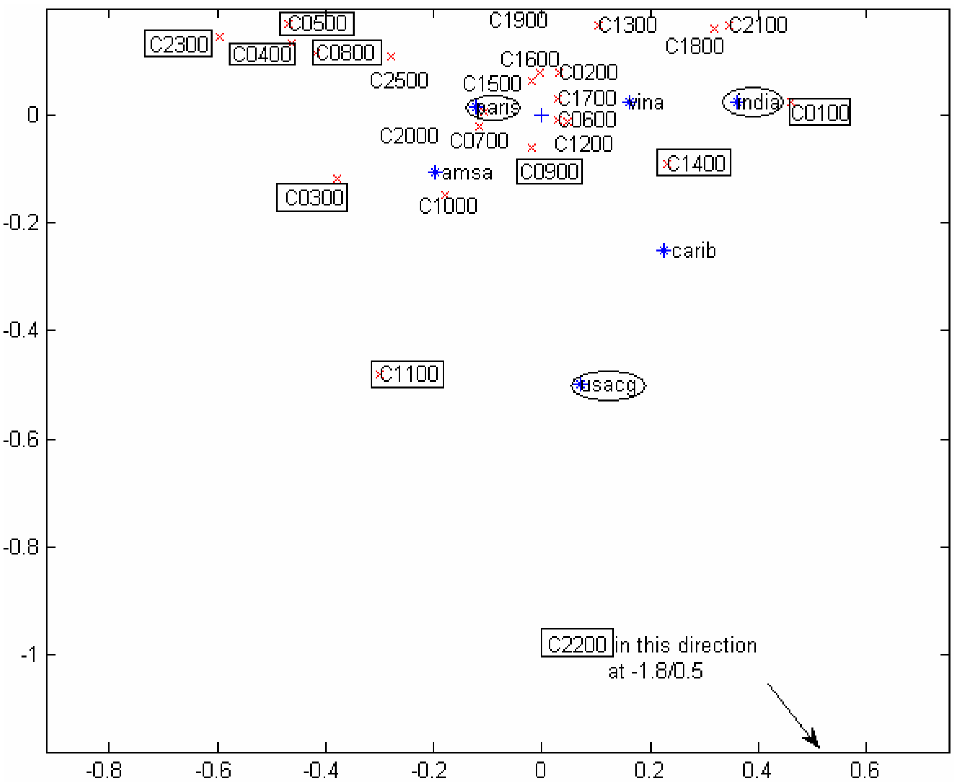


Figure 5: Unknown Ownership, explained inertia 81.6%



Concerning the similar patterns of the regimes over the different ownership groups, we note the closeness of the Paris MoU with the Viña del Mar Agreement, and the Caribbean MoU with the Indian Ocean MoU. On the other hand, AMSA and the USCG are separated from the other regimes in all plots. In the last plot, corresponding to the group of unknown owners, the regimes appear to be spread out more evenly.

The similarity in the grouping of regimes reflects the similarity with respect to treatments of vessels for the various owner groups. Correspondingly, we see that there is a similar grouping of certain deficiencies across the plots as well. This indicates that the differences between regimes may be caused by a structurally different use of the deficiency codes in the inspections. This reflects the different emphasis that is given during an inspection by a particular regime. There are also small differences in the plots corresponding to the different ownerships. However, the deficiencies show more consistency across the ownership groups within the same MoU.

Concerning the distribution of the deficiency codes in relationship to the different port state regimes, we note that certain codes appear to have a specific and strong emphasis in certain regimes only. This applies to code 900 (structural safety) for the Paris MoU and the Viña del Mar Agreement or code 100 (ship certificates) for the Caribbean MoU and the Indian Ocean MoU. AMSA is mostly associated with code 1600 (radio communication) and code 2000 (SOLAS operational deficiencies) and the USCG with code 1100 (cargoes) and 1400 (propulsion & aux. engine). However, all other codes show different locations across the plots. A surprising result is to see the location of the ISM codes (2500) and its closeness with the Paris MoU across all plots.

As mentioned earlier, Appendix A provides a compact summary and lists the deficiency codes, their importance in the plots and association to a regime for each ownership plot. If a deficiency code shows a higher contribution than the average towards the dimensions of a plot, it has an *"x"* in the table. The importance of a regime towards the dimensions of the plots is indicated in *"bold and italic"* for easier identification.

It is understandable to find the relative closeness of the Indian Ocean MoU with the Caribbean MoU as both regimes are seen as emerging regimes and have been in operation since 1996 and 1998 respectively. This can further explain their importance for the inspection of certificates versus other areas since it is easy to inspect. The USCG and AMSA are both separate countries and are expected to be more homogeneous than the other regimes. This could be an explanation for both regimes to be separate from the rest of the group in all plots and with their importance in determining the dimensions of the plots. Both regimes use checklists in their inspections and are in operation since considerable time (early 1990's).

The most interesting grouping is to see the closeness of the Paris MoU with the Viña del Mar Agreement despite their large difference in years of operation (1982 versus 1992). One similarity between the regimes is that they consist of various countries with different nationalities where the mentality and culture of the South of Europe is very similar to Latin America. It is an unexpected finding and not in line with industry perception.

In addition to the distribution of the deficiency codes with respect to the regimes per ownership group, it may also be interesting to consider other ship particulars such as ship types, flags and classification societies. For this purpose, *Flag states* are categorized as black, white or grey listed flag states as per the Paris MoU Annual Reports plus a separate category for flags that are not on the list and therefore undefined. Classification societies are classified as either members of IACS[5] or not IACS and ship types are grouped into six main categories such as: general cargo, dry bulk, container vessels, tankers (including oil, chemical and gas tankers), passenger vessels and other ship types. We can add all these ship particulars as supplementary points in the plots. However, the results of such a procedure showed that all supplementary points were drawn towards the average. This means that for a particular owner ship group, the distribution over the deficiencies for a certain flag group, ship type or classification society group is more or less equal to the average distribution of all regimes. Therefore, these supplementary points are left out in the final plots.

4. Conclusions

This article looked at providing a better insight into the relationship between ownership of a vessel and inspection data by visualizing this relationship and to confirm that differences across port state control inspections exist, to visualize and interpret these differences.

The results of this analysis confirm that treatment of vessels across port state control regimes differ although some groupings of regimes across the ownership groups can be found. With respect to the ownership groups, one can see that within certain regimes, deficiencies found within the area of structural safety, Marpol Annex I (oil pollution), ship certificates, radio communication, SOLAS operational deficiencies, cargo operations and propulsion and aux. engines do not vary across ownership groups. This can mean two things. First, there is no

[5] the International Association of Classification Society

difference between the ownership groups with regard to these deficiencies and second, there is a difference in the emphasis that is placed by the regimes during their inspections. This emphasis is consistent across ownership groups.

Association can be found between the Paris MoU and the Viña del Mar Agreement and between the Caribbean MoU and the Indian Ocean MoU while the USCG and AMSA are always apart from the rest which can be explained by the fact that those regimes are more homogeneous. The association between the Caribbean MoU and the Indian Ocean MoU is easier to explain since both regimes are emerging regimes. The association between the Paris MoU and the Vina del Mar Agreement is a surprising result of this analysis. Another surprising result is to see the location of the ISM codes (2500) and its closeness with the Paris MoU across all plots rather than similar association with regimes like AMSA or the USCG.

The results visualize the differences in treatments and show that there is room for harmonization in all inspection areas which can best be achieved by increasing the cooperation amongst regimes, combined training and the use of a centralized database such as GISIS which would allow a more accurate ship history and dissemination of results to flags, classification societies and port state control officers across all regimes.

References

Alderton T. and Winchester N (2002). "Flag States and Safety: 1997-1999". *Maritime Policy and Management*, Vol 29, No. 2, pp 151-162

Clausen, S.E. (1998). *Applied Correspondence Analysis* - An Introduction (Sage University Paper Series, Nr. 07-121)

Greenacre, M.J. (1984). Theory and Applications of Correspondence Analysis, Academic Press, London.

Jobson, J.D. (1992). Applied Multivariate Data Analysis, Volume II: Categorical and Multivariate Methods, Springer-Verlag, New York.

Knapp S, Franses PH (2007), A global view on port state control - econometric analysis of the differences across port state control regimes, *Maritime Policy and Management*, forthcoming

Paris MoU Annual Reports for 2000 to 2004

Appendix A: Summary of findings across PSC regime per ownership group

Code	Code Description	TMN		EMN		Unknown		Open Reg.	
100	Ship's certificates	**x**	india/carib	**x**	india/carib	**x**	***india***	**x**	***india***
200	Crew certificates	**x**	india/carib	**x**			vina	**x**	carib
300	Accommodation	**x**	***usacg***	**x**	***amsa***	**x**		**x**	vina
400	Food and catering			**x**		**x**			
500	Working spaces			**x**		**x**		**x**	
600	Life saving appliances				paris		vina		
700	Fire Safety measures			**x**	***amsa***		amsa		
800	Accident prevention				paris	**x**			
900	Structural Safety		vina/paris	**x**	vina/paris	**x**	amsa		vina/paris
1000	Alarm signals						amsa		
1100	Cargoes	**x**	***usacg***	**x**	***usacg***	**x**	***usacg***	**x**	
1200	Load lines			**x**			vina	**x**	
1300	Mooring arrangements		vina/paris						paris
1400	Propulsion & engine	**x**	***usacg***	**x**	***usacg***	**x**	***india***/carib	**x**	***india/usacg***
1500	Safety of navigation		vina/paris		vina/paris		***paris***		vina
1600	Radio communications	**x**	***amsa***	**x**	***amsa***			**x**	***amsa***
1700	MARPOL Annex I		india/carib				vina		paris/carib
1800	Gas and chemical carriers		india/carib				***india***		vina
1900	MARPOL Annex II								
2000	SOLAS Oper. deficiencies	**x**	***amsa***				amsa	**x**	***amsa***
2100	MARPOL oper. deficiencies								carib
2200	MARPOL Annex III		india/carib		***usacg***	**x**			
2300	MARPOL Annex V					**x**			
2500	ISM related deficiencies		vina/paris				paris		paris

Note: ***regimes in bold and italic*** *highlight strong contribution towards the plot, abbreviations are the same as used in the plots while an x would indicate if a particular deficiency code is important*