ERASMUS UNIVERSITY OF ROTTERDAM - ERASMUS SCHOOL OF ECONOMICS
MSC ECONOMICS AND BUSINESS
MASTER SPECIALIZATION: URBAN, PORT AND TRANSPORT ECONOMICS

Port-related socio-economic impact studies: a common framework

Master Thesis

Supervisor: Dr. M. Nijdam
Student name: Tiliakos, Alexandros / 356862

Table of Contents

Abstract 3

Introduction 4

Overview 5

1. Theoretical Background 6
   - 1.1. Historical Arguments concerning PEIS Validity 6
2. Analysis 15
   - 2.1. Objectives 15
   - 2.2. Variables and Methodology 15
   - 2.3. Global Data 17
   - 2.4. Region-specific Data: Europe 22
   - 2.5. Region-specific Data: USA 26
   - 2.6. Region-specific Data: Canada 29
   - 2.7. Region-specific Data: Australia 33
   - 2.8. Chronological Considerations 36
3. Summary and Conclusions 38
   - 3.1. General Observations 38
   - 3.2. Region-specific Observations 40
   - 3.3. Optimization towards a Common Framework 49

Acknowledgements 52

References 53

Appendix I: List of PEIS original publication titles and maps 56

Appendix II: List of PEIS codification and reference years 60

Appendix III: List of variables and definitions 61

Abstract

Port economic impact studies have been often accused for lack of transparency and non-meritocratic management of information. In the present paper, a meta-analysis is conducted to formulate a common framework for port socio-economic impact studies (PEIS) by using case studies from Europe, the USA, Canada and Australia. A wide range of variables have been analyzed, based on the frequency they are encountered and their contribution to reliability and objectivity. In conclusion, a common framework is proposed, aiming towards the optimization of future port-related socio-economic impact studies.

Keywords: *port economic impact studies (PEIS), meta-analysis, optimization*

Introduction

It is commonly accepted that ports have always been a source of wealth for their local societies. However nowadays, this former "axiom" has been highly doubted for a number of reasons: the port industry is no longer labor intensive since the majority of processes taking place are automated; on the contrary, ports may be considered as capital intensive, but due to the international capital markets and to horizontal and vertical integration strategies of the global actors, ports' benefits are not only confined to port regions. Moreover, port development and expansion require ever more space, in general land from the local society, but the land hiring cost is usually lower than its opportunity cost. Apart from that, port negative externalities, including traffic congestion, land consumption and environmental problems, affect mainly local societies, while positive gains from port functioning, mainly the generation of taxes and duties, have country-wide characteristics.

For the aforementioned reasons, port economic impact studies (PEIS) are decidedly useful tools for the evaluation of ports impacts. However, their results are sometimes confusing due to the different methodologies used. Furthermore, depending on the methodology, port impact studies have been highly criticized for promoting specific political and business ends.

Although varying PEIS methodologies exhibit different advantages and disadvantages, a common optimal framework has not been as yet proposed. Such variations cover not only practical considerations, such as which variables are used or what methodology is followed, but also conceptual ones. Mismatches include differences in the definition of a port, the spatial level of a port's impact, its area of effect, firms that are active in a port and are taken into consideration in the research. On the other hand, port impact studies have been severely criticized for omitting, either on purpose or not, variables of critical importance, usually with negative results, and for not taking into account firms that are not directly active in the ports but indirectly involved with it. It has to be noted that such firms are usually omitted due to high cost and time involved in data collection and management. Thus, port impact studies have been often accused for lack of transparency and for non meritocratic information management.

Under the prism of meritocratic management and exchange of information, and also of the transparency as the principal component for evaluating investments, the existence of a common framework for port related socio-economic impact studies is considered of high importance. A common methodology, using common variables, may guarantee clear and accurate results that will be indifferent towards political and non-political ends. Furthermore, a common methodology will elevate PEIS to a primary tool for evaluating and comparing investments with mixed results towards societies, such as ports.

The present study aims at creating a common framework for port related socio-economic impact studies by comparing a significant number of port impact studies from all over the world and by consequently identifying best practices and the most appropriate variables. The studies cover a wide range of ports of various specializations and sizes.

Overview

The first chapter of this study is dedicated to the presentation of the theoretical background concerning port economic impact studies, by outlining the historical arguments concerning PEIS validity. The theoretical mathematical model used as the main engine in constructing the PEIS is presented, along with arguments and counter-arguments that have historically arisen concerning the model's reliability. Furthermore, we examine a list of proposals and various methodologies that have been offered as alternatives. The various stages of an impact analysis assessment are also presented, along with the conceptual differentiations between direct and indirect impacts. Lastly, we examine the key mathematical methodologies for tackling the algebraic challenges of Input-Output Tables that are used extensively in the model.

The second chapter of this study covers the analytical methodology followed in our endeavor. Thus, the main objectives of the project are defined and the key variables are selected and codified. Aiming firstly at a general understanding, we treat the data gathered from all case studies collectively, in order to deduce global observations regarding the scope and validity of PEIS when used in an international setting. To tackle the problems of treating an overwhelming quantity of data stemming from various sources, regions and eras, we focus our efforts towards identifying the dynamics and importance of common variables used in PEIS globally. Thus, we introduce a binary/Boolean (True/False) scheme to treat the variable matrices and we apply selective methods of classifying statistical relationships: our analyses target Variable Threshold and Study Adequacy levels.

Continuing the meta-analysis, we extend our methodology in regional-specific sets of data, by examining consequently case studies from ports in the regions of Europe, the USA, Canada and Australia. For continuity reasons, we follow the same methodology of identifying Variable Threshold and Study Adequacy levels, in order to compare obtained results with the general observations. Lastly, we cross-examine case studies that are chronologically distanced from the main body of PEIS, which is generally located in the first decade of the 21st century.

The final chapter of this study serves as an overview of all observations and conclusions. We re-examine our general observations regarding the importance of common variables used in PEIS globally and then we cross-examine these observations with the ones regarding region-specific observations. Thus, we outline general systems of variables used preferentially in individual regions and we compare them with the global trends. We also examine the effects of statistical limitations due to restricted sample data availability. Finally, we attempt to comment on model optimization by drawing on our observations based on this meta-analysis.

1. Theoretical Background

Throughout history, ports have been closely linked to the cultural and economic growth of the local societies and they have been considered a source of wealth and prosperity for the local population. Ports have always managed to attract businesses, mainly due to the minimization of the transportation costs for resources and commodities, as well as large numbers of job seekers due to high labor force demand. However, in our days the aforementioned historical characteristics of the ports have changed radically and the importance of the ports themselves for the local societies stands ambiguous. On one hand the economic significance of the port stands under question, while on the other hand negative externalities burdening local societies seem out of proportion.

According to recent studies[1] the port industry is no longer characterized by the labor intensity of the past, since the majority of the processes (cargo loading/unloading) have been automated, resulting in the negation of past positive impacts on labor demands. Instead, the port industry of today, is much more capital intensive, although the functioning of international capital markets, as well as horizontal and vertical integration strategies of the global actors, disperse most positive effects of the port industry beyond the confinement of regional borders, extending to a global scale. On other hand, negative externalities, such as traffic congestion and environmental pressures, manifest as a general problem of land consumption; this becomes aggravated through the continuous expansion and development, thus characterizing port facilities as land intensive, mainly due to containerized high traffic. As a side note, however, it has been observed that purchasing costs of new land for port expansion are usually comparatively lower than opportunity costs.

Earlier studies have also hinted on the changing relationship between ports and societies.[2] This change has been attributed to two fundamental factors: the ever increasing range of the international trade, resulting in an altered distribution of trading destinations, and the port's industrialization originating from structural changes, as a direct result of specialization and rapid mechanization/automation of the production line.

1.1. Historical Arguments concerning PEIS Validity

The most widely used tool for evaluating a port's impact is the Port-related socio-Economic Impact Study. The first PEIS was conducted in the US during the second half of the 1960's, addressing the ports of New York and New Jersey. It must be noted that from the very beginning of the employment of PEIS, their validity has been the topic of controversy, instigating great debates by a number of scholars and economists. The initial arguments against the validity of PEIS[3] can be summarized in the following points:

[1] *Marco et al (2001)*
[2] *Suykens (1989)*
[3] *Waters (1977)*

- PEIS fail to provide useful guidelines for port development planning.
- PEIS fail to assess any changes in public investments in port facilities' development.
- PEIS are traditionally modeled under the macroeconomic steady-state assumption that technology remains stable.
- PEIS are founded on an expenditure approach, based on the questionable assumptions that a region's export result in a 100% increase of that region's income, that all prices remain rigidly stable and that all port-related expenditures result in an induced impact on that region's income levels.
- The economic model is based on the assumption of a simplistic multiplier, which is unrealistic since the multiplier is different for different classes of commodities.
- PEIS operate under an initial assumption that all transportation expenses are exogenous to the region's economy, something that often leads to unjustifiable implications.

Although the arguments against the validity of the PEIS seem correct, the fundamental counter-argument posed by PEIS supporters poses that the PEIS objectives themselves need to be taken under consideration. PEIS are static models on purpose, in other words they depict the port's situation only for the time that the data were collected. However, the argument concerning the multiplier is mostly accurate, but since there is no better alternative, its use is recommended only if its limitations and its problems are taken into consideration, and if the proper additions are implemented. Cost-benefit analyses, along with input-output models, seem as much more preferable options; however, PEIS and cost-benefit analyses should not be regarded as supplementary methods but as complementary. Thus, the combination of the two may produce the fullest and most accurate picture of the port's situation possible.[4]

1.1.1. Mathematical Model

The combined model may be used satisfactorily to estimate whether a port's expansion should be funded. According to case studies on data procured from the port of Mobile in 1967, a port should be expanded only if three conditions are met:[5]

- The port enjoys a profitable operation of a sustainable nature.
- The port enjoys high rates of actual to preferred capacity utilization in recent years.
- There is an expected increase in demand for services of the port in the foreseeable future.

[4] *Chang (1978)*
[5] *Chang (1978)*

The first model condition is addressed by maximizing the following equation:

$$max\ Q = wL + rK$$

subject to:

$$Q = AL^a K^b e^{c(T/L)}$$

Where: Q = annual gross earnings of the Port of Mobile in 1967 prices,

w = wages per worker employed directly by the port,

L = average of monthly number of port's employees,

r= rate of return of capital, *i.e.*$(Q - wL)/K$,

K = value of net assets in 1967 prices,

a = labor coefficient,

b = capital coefficient,

$e^{c(T/L)}$ = proxy for technological improvement, where T/L is the tonnage per unit of labor, loaded and unloaded at the port.

The port capacity function:

$$T = f(time, H, M)$$

estimates the rates of actual to preferred capacity utilization, where:

T = actual tonnage handled at the port facilities,

H = dummy variable for the peak tonnage of each tonnage cycle,

M= dummy variable representing non-peak and non-bottom tonnages for each tonnage cycle.

Using time data series we estimate the rate of actual to preferred capacity utilization.

In the original case study, the author considers *a priori* that, as far as the future demand of the port's services is concerned, it will increase due to the Tennessee-Tombigbee waterway project.

1.1.2. Arguments and Counter-arguments

After critical consideration of the PEIS debate, we may reach the conclusion that the debate did not aim to weaken the PEIS, but quite the opposite: PEIS end up acquiring a fixed structure and consistency after highlighting their limitations, thus setting the foundations of a solution.

Beyond the methodological part, which will be further analyzed in subsequent section, PEIS have been severely criticized for promoting individual political/business interests and for seeking to influence the public opinion. This has even been mentioned in the original debate concerning the validity of PEIS, where it was placed inside the framework of PEIS' objectives, one of which was to enhance the port's image in the public opinion. One characteristic example of PEIS employed for political manipulation concerns the West Coast port lockout in the 1980s, when a suitably constructed PEIS had been used to trigger governmental mediation.[6] Similar examples may be drawn from other industries as well, *e.g.* economic impact studies being used as a political tool in the professional sports industry.[7] More precisely, there are examples where the employed studies highlighted huge economic benefits for the concerned cities at the time when the authorities were seeking to attract a certain professional team.

The counter-argument, employed by PEIS supporters, is that one of the main objectives of such studies is to bridge the gap between the port industry and the concerned local society.[8] More recent studies extend the number of objectives, by focusing on the relationship between the port industry and the concerned regional economy as a more thorough extension of the local society. Thus, the PEIS can enhance the understanding of the economic relationships between the port industry and the regional economy, they can estimate the total regional economic impact and they can function as simulation models to evaluate future investments.[9]

1.1.3. Proposals and Methodologies

Concerning the methodological processes of the PEIS, a matter which has steered great controversy, there have been a number of proposals seeking to alleviate the negative implications of the original model. These proposals may be summarized to the following fundamental methodologies:[10]

- Methodology 1 is founded on the cost aggregation of the different economic agents to develop the transport of goods and services related to the port studied.
- Methodology 2 is founded on the added value aggregation of the two big groups of economic agents (Port Industry and Port Authority) to study the direct economic impact of ports.
- Methodology 3 is founded upon the added value of three big groups of economic agents (Port Industry, Port Authority and Port Users) to study the total economic impact, both direct and indirect.

[6] *Hall (2004)*
[7] *Hudson (2001)*
[8] *Chang (1978)*
[9] *Marco et al (2001)*
[10] *Cotto-Millan et al (2010), p152*

The main disadvantage of the first methodology is that it is predominantly based on questionnaire surveys, thus its results may be untrustworthy. On the other hand, its advantage is that it can depict the impact per commodity and per tonnage of commodity, assuming that the analysis of the impacts per commodity is more useful than the analysis of the total impact.

As far as the second methodology is concerned, its main advantage is that it can depict short term results and thus the studies can be easily updated. On the other hand, it is mainly based on subjective opinion, thus its results cannot be empirically verified. Moreover, its value is more limited, since it may only be used to estimate direct impacts, thus providing only a partial picture of the total impact.

The third methodology has many similarities to the second one, with the added feature of measuring direct, indirect and induced impacts, thus managing to provide the total picture of the port's situation. This methodology gives some freedom to the user, as far as data collection and data management are concerned, and for that reason its total time cost and level of confidence may vary.

1.1.4. Impact Analysis Stages: Direct *vs.* Indirect Impacts

The process of the port's impact analysis covers four discrete stages:[11]

- Estimation of the economic activity that is directly linked to the port.
- Estimation of the relationship of the aforementioned activity with other sectors of the economic system.
- Estimation of the multipliers for direct and indirect impacts on the economic system.
- Estimation of the induced impacts on the economic system.

There seems to be a disagreement in the literature concerning the definition, and thus the estimation, of direct, indirect and induced impact. However, in order to adopt a single concrete foundation for the needs of this study, one definition has been selected in view of its capacity to encompass the majority of the critical factors concerning impact analysis:

"*Direct impact refers to the volume of employment, wages and salaries, sales, Gross Operating Surplus (GOS), taxes and Gross Value Added (GVA), generated by the Port Community and the Port Users Community. Indirect impact refers to the economic activity deriving from the purchasing and investment relationships that the Port Community and Port Users Community maintain with the rest of the economy. Induced impact refers to the effects produced as a result of private consumption deriving from the wages and salaries received by the workers who (in the Port Community or*

[11] *Garcia, Lopez (2004)*

Port Users Community) undertake their activity due to the existence of the port".[12]

At this point it is considered appropriate to revisit the original arguments concerning the stable prices assumption followed in the PEIS. The negative implication of such an assumption is that the changes in the production resulting from the changes in the prices are not taken into consideration, thus the estimation of the direct impact in the PEIS is flawed[13]. In a theoretical level, we can estimate direct impacts by accounting for changes in production, proving that the negative implications of the assumption can be tackled[14]. However, the stable prices assumption is nevertheless valid, since we have accepted that PEIS are static and not dynamic models in the first place[15].

Further research on direct impacts and on methods to estimate them has pointed out that not all of the economic activities taking place in a port should be exclusively classified under the category of direct impacts[16]. A case study in the port of Hampton Roads has showed that out of the multitude of activities taking place in the port, only stevedoring, steamship owner/agent, freight forwarder consolidator, launch service, pilots, diving centre, naval architects, container service, coal inspection, weighers and samplers, seaman service and towing and transportation can be classified exclusively as direct impacts.[17]

TABLE 1. PERCENTAGE OF TOTAL EMPLOYMENT THAT IS PORT-RELATED

Industry	% of Total Employment That is Port Related
Stevedoring	100
Steamship Owner/Agent	100
Freight Forwarder Consolidator	100
Motor Freight	30
Railroads	29
Ship Repair	5
Launch Service	100
Ship Watching	41
Pilots	100
Diving Service	100
Ship Chandler	96
Naval Architects	100
Marine Surveyors	99
Marine Supplies	67
Container Service	100
Insurance	12
Crane Service	34
Coal Inspection	100
Weighers & Samplers	100
Seaman Service	100
Towing & Transportation	100
Terminals & Warehouses	67

Table I: Total port-related employment, Port of Hampton Roads, *Yochum et al (1987)*

[12] *Cotto-Millan et al (2010) p.169*
[13] *Waters (1977)*
[14] *DeSalvo (1994)*
[15] *Chang (1978)*
[16] *Yochum, Agarwal (1987)*
[17] see *Table I*

Of course each case study being different, the results of the above table cannot be replicated in every case. Each port has different characteristics and different traffic mix, therefore the percentage of the total employment that is port related is unique for each port. Thus, for each port under study, one needs to reevaluate the matrix corresponding to the direct impact activities.

1.1.5. Input-Output Tables

As far as indirect and induced impacts are concerned, a wide variety of methods has been used for their estimation, although the most accepted one is the Input-Output analysis. The Input-Output tables used as the analytical tool actually represent the total economy of a region or a country, given the fact that the output of an industry is used as the input in another industry. More precisely, the tables give a set of measurable relationships that reflect the interdependencies in a given economic region. The Input-Output tables are structured into three blocks: the intermediate consumption block, the primary inputs block and the final demand block.[18]

BRANCHES	1	2	3	...	N	ΣIC	C	G	I	D	TE
1				...							
2		INTERMEDIATE CONSUMPTION						FINAL DEMAND			
3											
...	...	...	...	...	...	...	...	...	...	...	...
N				...							
ΣCI				...							
W											
GOS		PRIMARY INPUTS									
GVA				...							
P				...							

Graph 2 IOT. Adding table

Table II: Input-Output Matrix, *Cotto-Millan et al (2010), p179*

The table presents the inputs and outputs generated in the economy. Assuming that we include *N* branches of activity in the tables, then, if the table is read in columns, each cell represents the amount of inputs generated by the activity in the corresponding row. If the table is read in rows, then each cell represents the amount of outputs that the activity in the corresponding row sells to the activity of the corresponding column. The last row represents the total amount of inputs acquired by each branch of activity, while the last

[18] see *Table II*

column indicates the total amount of outputs that each activity sells to the rest branches of the economy. The sums of the totals in all rows and columns must tally.

The final demand section describes the exact destinations of the activities' outputs and it is divided into five columns. *C* stands for households' consumption, *G* for public consumption, *I* for gross capital formation, *D* for the sum of final destinations - excluding intermediate consumption - and *TE* for the total sum of the final destinations.

The final section represents the primary inputs of the economy. *W* stands for the remuneration of the employment factor, i.e. wages and salaries, *GOS* for the remuneration of the capital factor, *GVA* for the *Gross Value Added* of each branch, i.e. in general terms *W+GOS*, and *P* stands for the value of the production of each branch of activity as the sum of the value of the intermediate consumption plus the *Value Added*, i.e. *IC* (*Intermediate Consumption*) + *GVA*. The values in the *P* row tally with the values in the *TE* column.[19]

Therefore the tables provide us with the complete picture of the economy, since the columns illustrate the total costs of each branch of activity, by including the intermediate consumption and the remuneration of primary inputs; the rows illustrate both intermediate destinations and final ones.

The mathematical methodology for solving analytically the original input-output tables (generally known as *Walserian equilibrium* tables)[20] has been based on a series of suppositions, the most important of which is the elimination of all price effects in the replacement of inputs.[21] The assumption states that there are no limited factors and the technology remains constant, something which subsequently had led to the critique concerning the weaknesses of the PEIS. The analytical model introduced three main hypotheses:

- *Homogeneity*: each product is supplied for just one branch of activity
- *Proportionality*: the amount of input used depends on the level of production of each branch; this presupposes the existence of constant returns to scale.
- *Additivity*: there are no external economies or diseconomies; in other words, the total effect of various types of production constitutes the sum of individual effects.[22]

[19] *Cotto-Millan et al (2010)*
[20] *Leontief et al (1936)*
[21] *Cotto-Millan et al (2010)*
[22] *Cotto-Millan et al (2010), p180*

The following matrix equation represents the demand model:

$$(I - A)P = D$$

Where: I is the identity matrix,

A is the technical coefficients matrix,

P is the production column vector,

D is the final demand column vector,

$(I - A)$ is Leontief's matrix.

Despite the limitations of the PEIS, the criticism and the debate around their applications, they are considered an important tool for the evaluation of a port's impact on the regional economy. Still, a common framework for the PEIS may provide with consistency, transparency and reliability needed to reestablish them as widely used analytical tools.

2. Analysis

To be more precise, we are actually conducting a meta-analysis of the original PEIS data, in order to formulate a common framework. Thirty seven studies out of forty two have been used in this analysis. Five surveys have been omitted due to irrelevancy: they were focusing only on the environmental externalities of ports. The initial goal was to include at least one PEIS from each continent, but due to language restrictions and unavailability of data this was not achieved. The sample includes studies from the USA, Canada, Australia and Europe. The PEIS used have been codified based on their country of origin.[23]

2.1. Objectives

The objectives of this meta-analysis are primarily qualitative; the sheer volume of arithmetic data, as well as the discontinuity in variable usage between individual studies and between regions, constitutes an attempt at quantitative analysis as inappropriate. Thus, we focus our efforts in identifying sets of variables that are commonly used in PEIS globally, examining at the same time the frequency of use of such variables, the conceptual value and the degree of objectivity and validity they offer in the studies themselves. By doing so, we seek to uncover general trends in the conceptualization of PEIS, which underline the scope of such studies in the first place.

On a second level, the question of PEIS objectivity is also addressed by examining whether or not a given study is addressing the port's relationships to the broader economic region; regional differences in adopting systematic frameworks and guidelines for PEIS may point out degrees of introspectiveness characterizing the scope of such studies. The range of variables employed in a given study is also revealing as to the broadness of the questions that it seeks to answer, something that may also be linked either to its aforementioned introspectiveness or to the systematic limitations set in each region. Thus, an objective of the meta-analysis focuses on uncovering such dynamics.

Lastly, the meta-analysis seeks to underline common conceptual trends between regional frameworks and their flexibility, the frequency of use of key concepts and the possibility of establishing fixed sets of core concepts that may constitute a reliable and necessary common framework for future PEIS, with the possible option of including secondary concepts as a means to provide additional in-depth insight at an auxiliary level.

2.2. Variables and Methodology

Through careful examination of all case studies, eighteen variables have been identified as fundamental elements of PEIS, based on their contribution to reliability and objectivity. The variables most commonly used are:

[23] for a complete list of PEIS codification, see Appendices I & II

- *Availability of definitions*
- *Availability of direct and indirect impacts*
- *Availability of induced impacts*
- *Availability of results for development of business establishments*
- *Availability of results for employment*
- *Availability of results for employment expressed in person-years or full-time equivalent*
- *Availability of results for private investments*
- *Availability of results for taxes*
- *Availability of results for value added*
- *Detailed information on sectoral boundaries*
- *Distinction of traffic mix*
- *Growth expressed in TEUs or other units*
- *Impact per commodity group*
- *Impact per industry*
- *Inclusion of hinterland*
- *Productivity*
- *Use of input-output methods*
- *Use of surveys*[24]

A binary/Boolean system has been used to allocate an arithmetic score value to each individual study in terms of frequency of each and every variable employed. These score values allow for the classification of the totality of PEIS used in terms of *Study Adequacy* levels, a measure which provides insight as to the extent of a study's thoroughness, as it uncovers the fraction of variables used in a single study over the total set of variables usually employed. Moreover, this binary system of evaluation allows for the calculation of single variable use frequency over the total number of studies, leading to the classification of the totality of variables used in terms of *Variable Threshold* levels, which uncovers the frequency of a single variable use over the total set of PEIS examined.[25]

Further classification of PEIS adequacy levels and variable frequencies utilizes both descriptive statistics and observations of the frequency distributions. By allocating zones of occurrence based on units of *standard deviations* around *the statistical mean* of each and every distribution, we can categorize individual variable frequency and study adequacy levels by their placement in the respective zones:

- Zone of negligible occurrence: $A: x \in [0\%, \bar{x} - s]$
- Zone of common occurrence: $B: x \in (\bar{x} - s, \bar{x} + s)$
- Zone of absolute certainty of occurrence: $C: x \in [\bar{x} + s, 100\%]$

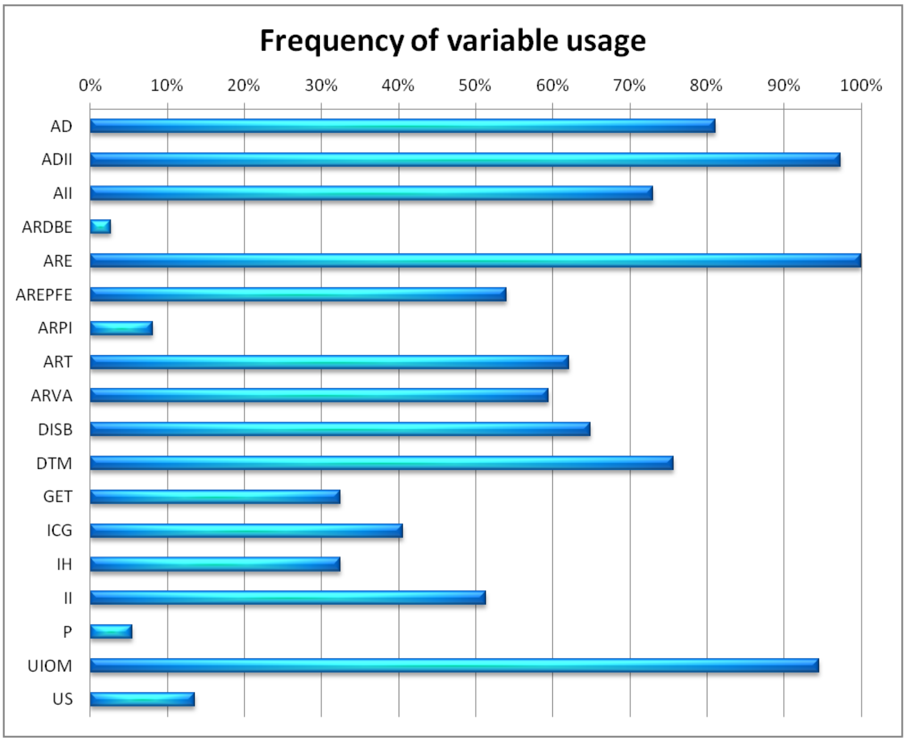
The above operation is critical for drawing conclusions on common framework recommendations for future PEIS.

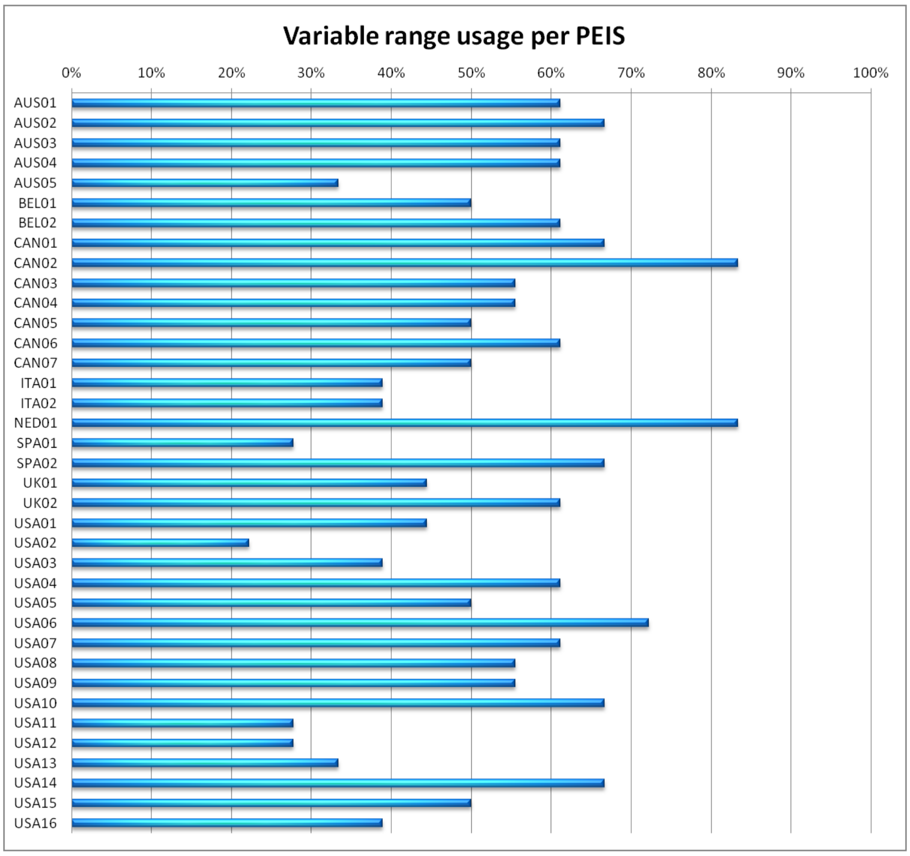
[24] The employment of the variable *use of surveys* (***US***) may be regarded as a questionable method as far as objectivity is concerned; thus, its employment has been valued with a zero (FALSE).
[25] for a complete list of variables codification, see Appendix III

2.3. Global Data

The employment of the binary evaluation system over the totality of PEIS examined leads to the comprisal of a collective template, which calculates the variable frequency levels (bottom row) and individual study adequacy levels (far right column) useful for further analysis.

PEIS/Var.	AD	ADII	AII	ARDBE	ARE	AREPFE	ARPI	ART	ARVA	DISB	DTM	GET	ICG	IH	II	P	UIOM	US	*Total*	%
AUS01	1	1	1	0	1	1	0	0	1	1	1	0	1	0	1	0	1	0	*11*	**61%**
AUS02	1	1	1	0	1	1	0	1	1	1	1	0	1	0	1	0	1	0	*12*	**67%**
AUS03	1	1	1	0	1	1	0	0	1	1	1	0	1	0	1	0	1	0	*11*	**61%**
AUS04	1	1	1	0	1	1	0	0	1	1	1	0	1	0	1	0	1	0	*11*	**61%**
AUS05	1	1	0	0	1	1	0	0	1	0	0	0	0	0	0	0	1	0	*6*	**33%**
BEL01	1	1	0	0	1	1	0	0	1	1	0	0	0	1	1	0	1	0	*9*	**50%**
BEL02	1	1	0	0	1	1	1	0	1	1	1	1	0	0	1	0	1	0	*11*	**61%**
CAN01	1	1	1	0	1	1	0	1	1	1	1	1	1	0	0	0	1	0	*12*	**67%**
CAN02	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	0	*15*	**83%**
CAN03	1	1	1	0	1	1	0	0	1	1	1	0	1	0	0	0	1	0	*10*	**56%**
CAN04	0	1	1	0	1	1	0	1	1	1	1	1	0	0	0	0	0	1	*10*	**56%**
CAN05	1	1	1	0	1	1	0	1	1	1	0	0	0	0	0	0	1	0	*9*	**50%**
CAN06	1	1	1	0	1	1	0	1	1	1	0	1	0	0	0	1	1	0	*11*	**61%**
CAN07	1	1	1	0	1	1	0	0	1	1	1	0	0	0	0	0	1	0	*9*	**50%**
ITA01	1	1	0	0	1	0	0	0	1	1	1	0	0	0	0	0	1	0	*7*	**39%**
ITA02	1	1	1	0	1	0	0	1	1	0	0	0	0	0	0	0	1	0	*7*	**39%**
NED01	1	1	0	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	*15*	**83%**
SPA01	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	1	*5*	**28%**
SPA02	1	1	1	0	1	1	0	1	1	1	1	0	0	1	1	0	1	0	*12*	**67%**
UK01	1	1	0	0	1	1	0	0	0	0	1	0	1	0	1	0	1	0	*8*	**44%**
UK02	1	1	1	0	1	0	0	0	1	1	1	1	0	1	1	0	1	0	*11*	**61%**
USA01	1	1	1	0	1	0	0	1	1	0	1	0	0	0	0	0	1	0	*8*	**44%**
USA02	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	*4*	**22%**
USA03	1	1	1	0	1	1	0	1	0	0	0	0	0	0	0	0	1	0	*7*	**39%**
USA04	1	1	1	0	1	0	0	1	1	1	1	0	1	0	1	0	1	0	*11*	**61%**
USA05	1	1	1	0	1	0	0	1	1	0	1	0	0	0	1	0	1	0	*9*	**50%**
USA06	1	1	1	0	1	1	0	1	0	1	1	1	1	1	1	0	1	0	*13*	**72%**
USA07	1	1	1	0	1	0	0	1	0	1	1	1	1	0	1	0	1	0	*11*	**61%**
USA08	1	1	1	0	1	0	0	1	0	1	1	0	1	0	1	0	1	0	*10*	**56%**
USA09	0	1	1	0	1	1	0	1	0	0	1	1	0	1	0	0	1	1	*10*	**56%**
USA10	1	1	1	0	1	0	0	1	0	1	1	1	1	1	1	0	1	0	*12*	**67%**
USA11	0	1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0	*5*	**28%**
USA12	0	1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0	*5*	**28%**
USA13	1	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	*6*	**33%**
USA14	1	1	1	0	1	1	0	1	0	1	1	0	1	1	1	0	1	0	*12*	**67%**
USA15	1	1	1	0	1	0	0	1	0	1	1	0	0	0	1	0	1	0	*9*	**50%**
USA16	0	1	1	0	1	0	0	1	0	0	0	1	0	1	0	0	1	0	*7*	**39%**
Total	*30*	*36*	*27*	*1*	*37*	*20*	*3*	*23*	*22*	*24*	*28*	*12*	*15*	*12*	*19*	*2*	*35*	*5*		
%	**81%**	**97%**	**73%**	**3%**	**100%**	**54%**	**8%**	**62%**	**59%**	**65%**	**76%**	**32%**	**41%**	**32%**	**51%**	**5%**	**95%**	**14%**		



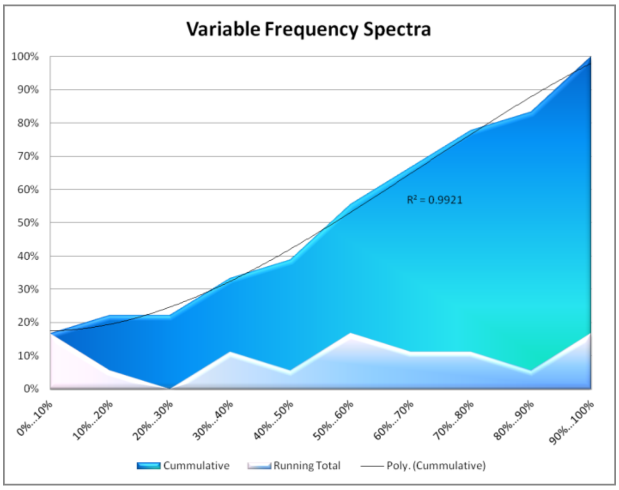


2.3.1. Variable Threshold Analysis

In order to expand the initial evaluation variable usage, we employ a detailed scheme of analysis, termed *Variable Threshold Analysis*, which examines the distribution of variables across 10% bins (thresholds), thus allowing the calculation of descriptive statistics, histograms and cumulative distributions of the frequency spectra for all variables.

		Variable usage threshold analysis: Frequency Spectra									
Variable	*Frequency*	*0%...10%*	*10%...20%*	*20%...30%*	*30%...40%*	*40%...50%*	*50%...60%*	*60%...70%*	*70%...80%*	*80%...90%*	*90%...100%*
AD	*81%*	0	0	0	0	0	0	0	0	1	0
ADII	*97%*	0	0	0	0	0	0	0	0	0	1
AII	*73%*	0	0	0	0	0	0	0	1	0	0
ARDBE	*3%*	1	0	0	0	0	0	0	0	0	0
ARE	*100%*	0	0	0	0	0	0	0	0	0	1
AREPFE	*54%*	0	0	0	0	0	1	0	0	0	0
ARPI	*8%*	1	0	0	0	0	0	0	0	0	0
ART	*62%*	0	0	0	0	0	0	1	0	0	0
ARVA	*59%*	0	0	0	0	0	1	0	0	0	0
DISB	*65%*	0	0	0	0	0	0	1	0	0	0
DTM	*76%*	0	0	0	0	0	0	0	1	0	0
GET	*32%*	0	0	0	1	0	0	0	0	0	0
ICG	*41%*	0	0	0	0	1	0	0	0	0	0
IH	*32%*	0	0	0	1	0	0	0	0	0	0
II	*51%*	0	0	0	0	0	1	0	0	0	0
P	*5%*	1	0	0	0	0	0	0	0	0	0
UIOM	*95%*	0	0	0	0	0	0	0	0	0	1
US	*14%*	0	1	0	0	0	0	0	0	0	0
	Running Total	*3*	*1*	*0*	*2*	*1*	*3*	*2*	*2*	*1*	*3*
	%	**17%**	**6%**	**0%**	**11%**	**6%**	**17%**	**11%**	**11%**	**6%**	**17%**
	Cummulative	*3*	*4*	*4*	*6*	*7*	*10*	*12*	*14*	*15*	*18*
	%	**17%**	**22%**	**22%**	**33%**	**39%**	**56%**	**67%**	**78%**	**83%**	**100%**

Variable Statistical Analysis	
Mean	52.67%
Standard Error	7.51%
Median	56.50%
Mode	32.00%
Standard Deviation	31.87%
Sample Variance	10.15%
Kurtosis	-108.53%
Skewness	-16.60%
Range	97.00%
Minimum	3.00%
Maximum	100.00%
Count	18
Confidence Level(95.0%)	15.85%



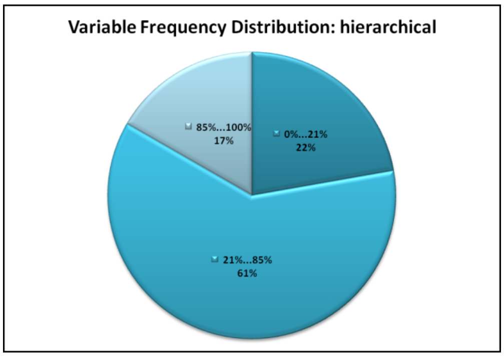
From the above statistics and cumulative distribution, we may discern hints of a normal Gaussian distribution, which can be justified by the global perspective. However, since the physical significance of such an occurrence may be debatable and the statistical sample is not wide enough, we will limit ourselves in observing the calculated statistics of arithmetic average and standard deviation; these help us map the distribution in three distinct zones:

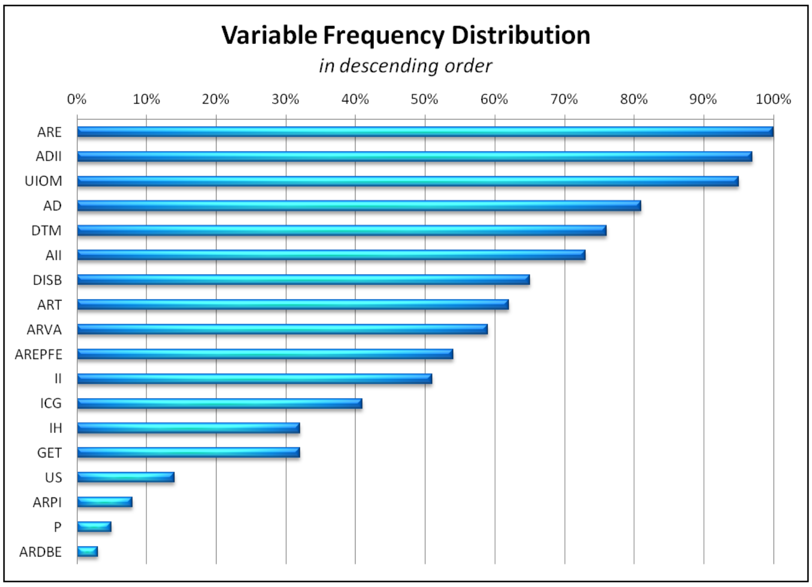
- Zone of negligible occurrence: $A: x \in [0\%, \bar{x} - s]$
- Zone of mean occurrence: $B: x \in (\bar{x} - s, \bar{x} + s)$
- Zone of absolute certainty of occurrence: $C: x \in [\bar{x} + s, 100\%]$

These zones correspond to the following sets:

- $A: x \in [0\%, 20.8\%]$
- $B: x \in (20.8\%, 84.5\%)$
- $C: x \in [84.5\%, 100\%]$

Variable	*Frequency* *in descending order*	*Frequency Range: hierarchical* *0%...21%*	*21%...85%*	*85%...100%*
ARE	*100%*	0	0	1
ADII	*97%*	0	0	1
UIOM	*95%*	0	0	1
AD	*81%*	0	1	0
DTM	*76%*	0	1	0
AII	*73%*	0	1	0
DISB	*65%*	0	1	0
ART	*62%*	0	1	0
ARVA	*59%*	0	1	0
AREPFE	*54%*	0	1	0
II	*51%*	0	1	0
ICG	*41%*	0	1	0
IH	*32%*	0	1	0
GET	*32%*	0	1	0
US	*14%*	1	0	0
ARPI	*8%*	1	0	0
P	*5%*	1	0	0
ARDBE	*3%*	1	0	0
	Running Total	*4*	*11*	*3*
	%	**22%**	**61%**	**17%**





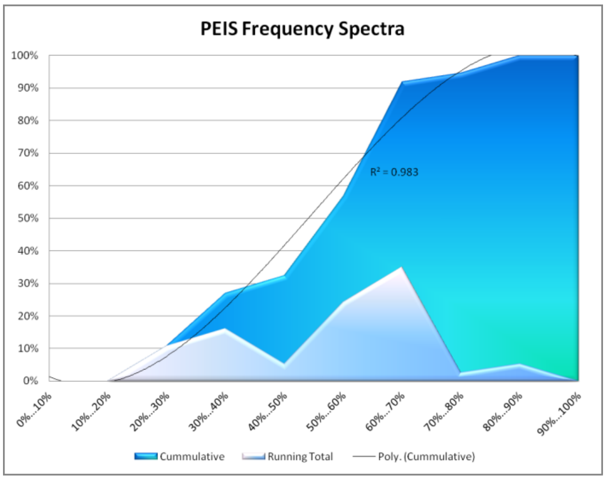
The above graphs illustrate the division of variable employment in areas of certain, mean and negligible occurrence, which will be addressed in the final chapter of this paper.

2.3.2. Study Adequacy Analysis

		PEIS range of analysis: Frequency Spectra									
PEIS	*Frequency*	*0%...10%*	*10%...20%*	*20%...30%*	*30%...40%*	*40%...50%*	*50%...60%*	*60%...70%*	*70%...80%*	*80%...90%*	*90%...100%*
AUS01	**61%**	0	0	0	0	0	0	1	0	0	0
AUS02	**67%**	0	0	0	0	0	0	1	0	0	0
AUS03	**61%**	0	0	0	0	0	0	1	0	0	0
AUS04	**61%**	0	0	0	0	0	0	1	0	0	0
AUS05	**33%**	0	0	0	1	0	0	0	0	0	0
BEL01	**50%**	0	0	0	0	0	1	0	0	0	0
BEL02	**61%**	0	0	0	0	0	0	1	0	0	0
CAN01	***67%***	0	0	0	0	0	0	1	0	0	0
CAN02	***83%***	0	0	0	0	0	0	0	0	1	0
CAN03	**56%**	0	0	0	0	0	1	0	0	0	0
CAN04	**56%**	0	0	0	0	0	1	0	0	0	0
CAN05	**50%**	0	0	0	0	0	1	0	0	0	0
CAN06	**61%**	0	0	0	0	0	0	1	0	0	0
CAN07	**50%**	0	0	0	0	0	1	0	0	0	0
ITA01	**39%**	0	0	0	1	0	0	0	0	0	0
ITA02	**39%**	0	0	0	1	0	0	0	0	0	0
NED01	**83%**	0	0	0	0	0	0	0	0	1	0
SPA01	**28%**	0	0	1	0	0	0	0	0	0	0
SPA02	**67%**	0	0	0	0	0	0	1	0	0	0
UK01	**44%**	0	0	0	0	1	0	0	0	0	0
UK02	**61%**	0	0	0	0	0	0	1	0	0	0
USA01	***44%***	0	0	0	0	1	0	0	0	0	0
USA02	***22%***	0	0	1	0	0	0	0	0	0	0
USA03	***39%***	0	0	0	1	0	0	0	0	0	0
USA04	***61%***	0	0	0	0	0	0	1	0	0	0
USA05	***50%***	0	0	0	0	0	1	0	0	0	0
USA06	***72%***	0	0	0	0	0	0	0	1	0	0
USA07	***61%***	0	0	0	0	0	0	1	0	0	0
USA08	***56%***	0	0	0	0	0	1	0	0	0	0
USA09	***56%***	0	0	0	0	0	1	0	0	0	0
USA10	***67%***	0	0	0	0	0	0	1	0	0	0
USA11	***28%***	0	0	1	0	0	0	0	0	0	0
USA12	***28%***	0	0	1	0	0	0	0	0	0	0
USA13	***33%***	0	0	0	1	0	0	0	0	0	0
USA14	***67%***	0	0	0	0	0	0	1	0	0	0
USA15	***50%***	0	0	0	0	0	1	0	0	0	0
USA16	***39%***	0	0	0	1	0	0	0	0	0	0
	Running Total	*0*	*0*	*4*	*6*	*2*	*9*	*13*	*1*	*2*	*0*
	%	***0%***	***0%***	***11%***	***16%***	***5%***	***24%***	***35%***	***3%***	***5%***	***0%***
	Cummulative	0	0	4	10	12	21	34	35	37	37
	%	**0%**	**0%**	**11%**	**27%**	**32%**	**57%**	**92%**	**95%**	**100%**	**100%**

A similar detailed scheme of analysis, termed *Study Adequacy Analysis*, examines the distribution of variable range levels in individual studies across 10% bins (thresholds).

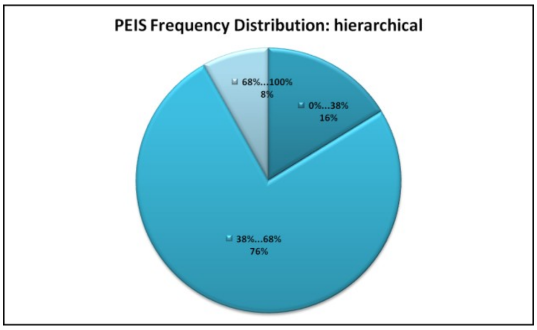
PEIS Statistical Analysis	
Mean	52.70%
Standard Error	2.50%
Median	55.56%
Mode	61.11%
Standard Deviation	15.19%
Sample Variance	2.31%
Kurtosis	-43.77%
Skewness	-14.82%
Range	61.11%
Minimum	22.22%
Maximum	83.33%
Count	37
Confidence Level(95.0%)	5.07%



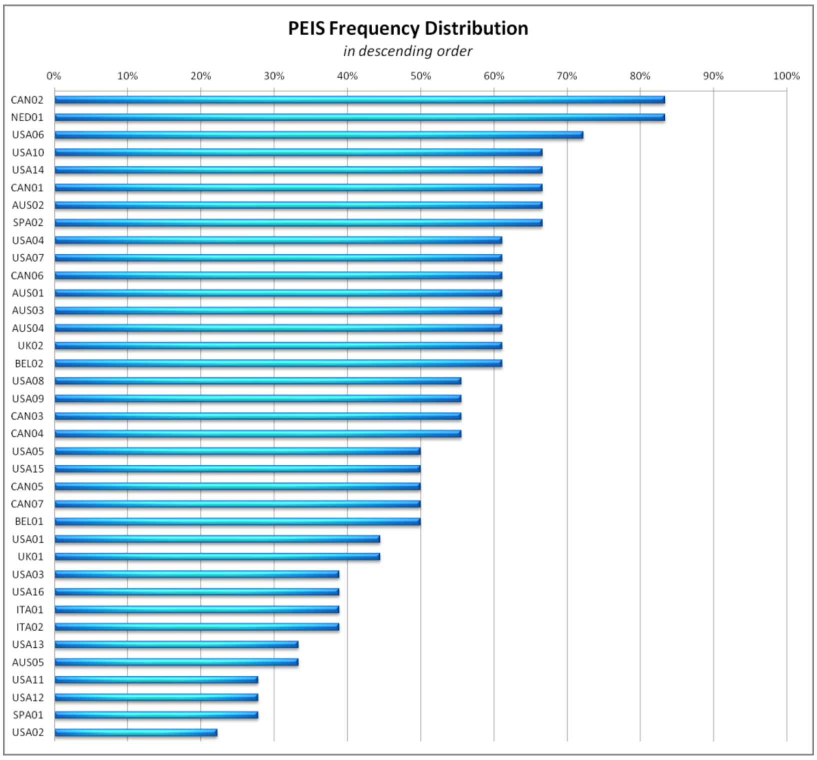
The three zones of significance formulate in the following areas:

- $A: x \in [0\%, 37.5\%]$
- $B: x \in (37.5\%, 67.9\%)$
- $C: x \in [67.9\%, 100\%]$

PEIS in descending order	Frequency	Frequency Range: hierarchical 0%...38%	38%...68%	68%...100%
CAN02	83%	0	0	1
NED01	83%	0	0	1
USA06	72%	0	0	1
USA10	67%	0	1	0
USA14	67%	0	1	0
CAN01	67%	0	1	0
AUS02	67%	0	1	0
SPA02	67%	0	1	0
USA04	61%	0	1	0
USA07	61%	0	1	0
CAN06	61%	0	1	0
AUS01	61%	0	1	0
AUS03	61%	0	1	0
AUS04	61%	0	1	0
UK02	61%	0	1	0
BEL02	61%	0	1	0
USA08	56%	0	1	0
USA09	56%	0	1	0
CAN03	56%	0	1	0
CAN04	56%	0	1	0
USA05	50%	0	1	0
USA15	50%	0	1	0
CAN05	50%	0	1	0
CAN07	50%	0	1	0
BEL01	50%	0	1	0
USA01	44%	0	1	0
UK01	44%	0	1	0
USA03	39%	0	1	0
USA16	39%	0	1	0
ITA01	39%	0	1	0
ITA02	39%	0	1	0
USA13	33%	1	0	0
AUS05	33%	1	0	0
USA11	28%	1	0	0
USA12	28%	1	0	0
SPA01	28%	1	0	0
USA02	22%	1	0	0
	Running Total	6	28	3
	%	16%	76%	8%



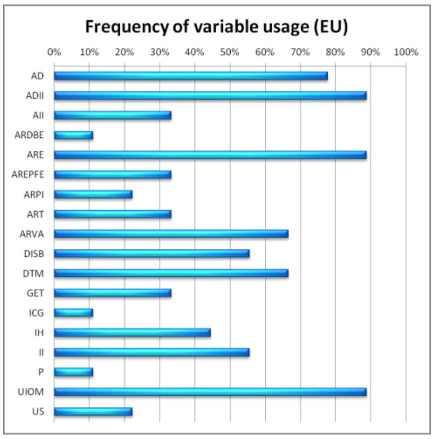
The distribution analysis and its graphical representation are consistent with the overall picture concerning study adequacy levels; the greatest volume of studies employs 39% to 67% of the common variables, while a comparatively small number of studies manage to incorporate a greater variable range and a similarly small number may be labeled as inadequate. The last may be due to the fact that the respective PEIS serve as targeted tools of a given perspective because of regional guidelines or needs (as half the sample corresponding to inadequacy originates from a single region). However, this remains a hypothesis.

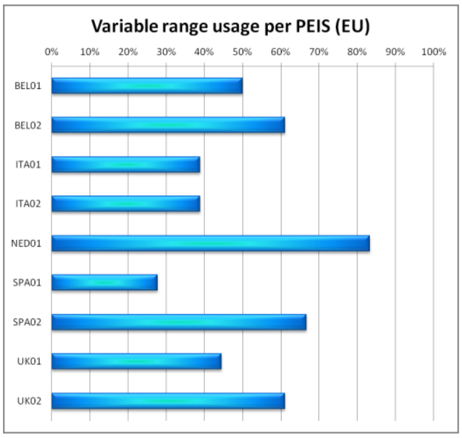


2.4. Region-specific Data: Europe

We follow the same procedure for analyzing sets of PEIS having the same regional origin in order to locate possible regional patterns. The initial template for Europe follows:

EU PEIS	AD	ADII	AII	ARDBE	ARE	AREPFE	ARPI	ART	ARVA	DISB	DTM	GET	ICG	IH	II	P	UIOM	US	*Total*	%
BEL01	1	1	0	0	1	1	0	0	1	1	0	0	0	1	1	0	1	0	*9*	**50%**
BEL02	1	1	0	0	1	1	1	0	1	1	1	1	0	0	1	0	1	0	*11*	**61%**
ITA01	1	1	0	0	1	0	0	0	1	1	1	0	0	0	0	0	1	0	*7*	**39%**
ITA02	1	1	1	0	1	0	0	1	1	0	0	0	0	0	0	0	1	0	*7*	**39%**
NED01	1	1	0	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	*15*	**83%**
SPA01	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	1	*5*	**28%**
SPA02	1	1	1	0	1	1	0	1	1	1	1	0	0	1	1	0	1	0	*12*	**67%**
UK01	1	1	0	0	1	1	0	0	0	0	1	0	1	0	1	0	1	0	*8*	**44%**
UK02	1	1	1	0	1	0	0	0	1	1	1	1	0	1	1	0	1	0	*11*	**61%**
Total	*7*	*8*	*3*	*1*	*8*	*3*	*2*	*3*	*6*	*5*	*6*	*3*	*1*	*4*	*5*	*1*	*8*	*2*		
%	**78%**	**89%**	**33%**	**11%**	**89%**	**33%**	**22%**	**33%**	**67%**	**56%**	**67%**	**33%**	**11%**	**44%**	**56%**	**11%**	**89%**	**22%**		

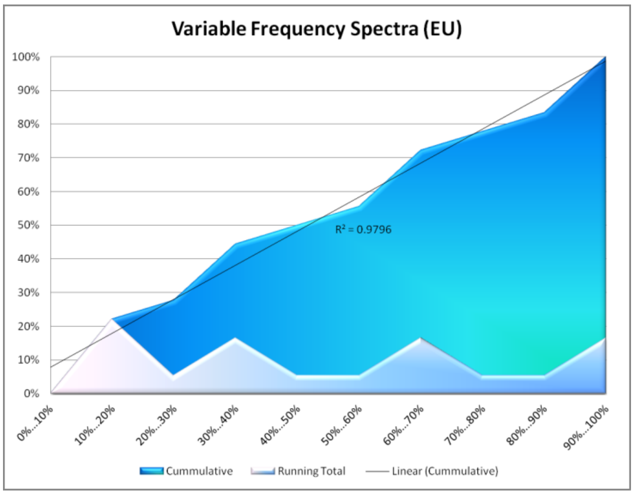




2.4.1. Variable Threshold Analysis

EU		Variable usage threshold analysis: Frequency Spectra									
Variable	*Frequency*	*0%...10%*	*10%...20%*	*20%...30%*	*30%...40%*	*40%...50%*	*50%...60%*	*60%...70%*	*70%...80%*	*80%...90%*	*90%...100%*
AD	*89%*	0	0	0	0	0	0	0	0	1	0
ADII	*100%*	0	0	0	0	0	0	0	0	0	1
AII	*33%*	0	0	0	1	0	0	0	0	0	0
ARDBE	*11%*	0	1	0	0	0	0	0	0	0	0
ARE	*100%*	0	0	0	0	0	0	0	0	0	1
AREPFE	*44%*	0	0	0	0	1	0	0	0	0	0
ARPI	*22%*	0	0	1	0	0	0	0	0	0	0
ART	*33%*	0	0	0	1	0	0	0	0	0	0
ARVA	*78%*	0	0	0	0	0	0	0	1	0	0
DISB	*67%*	0	0	0	0	0	0	1	0	0	0
DTM	*67%*	0	0	0	0	0	0	1	0	0	0
GET	*33%*	0	0	0	1	0	0	0	0	0	0
ICG	*11%*	0	1	0	0	0	0	0	0	0	0
IH	*56%*	0	0	0	0	0	1	0	0	0	0
II	*67%*	0	0	0	0	0	0	1	0	0	0
P	*11%*	0	1	0	0	0	0	0	0	0	0
UIOM	*100%*	0	0	0	0	0	0	0	0	0	1
US	*22%*	0	1	0	0	0	0	0	0	0	0
	Running Total	*0*	*4*	*1*	*3*	*1*	*1*	*3*	*1*	*1*	*3*
	%	**0%**	**22%**	**6%**	**17%**	**6%**	**6%**	**17%**	**6%**	**6%**	**17%**
	Cummulative	*0*	*4*	*5*	*8*	*9*	*10*	*13*	*14*	*15*	*18*
	%	**0%**	**22%**	**28%**	**44%**	**50%**	**56%**	**72%**	**78%**	**83%**	**100%**

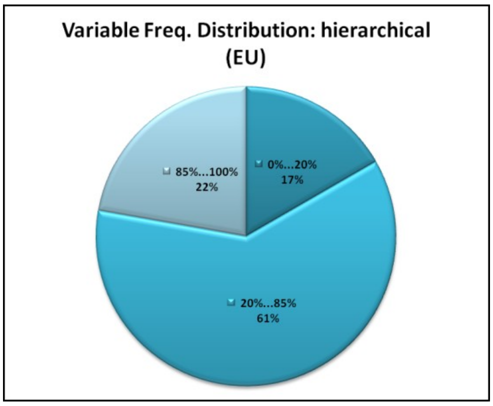
Statistical Analysis	
EU	
Mean	52.44%
St. Error	7.59%
Median	50.00%
Mode	100.00%
St. Deviation	32.21%
Sample Var.	10.37%
Kurtosis	-137.73%
Skewness	20.76%
Range	89.00%
Minimum	11.00%
Maximum	100.00%
Count	18
Conf.(95.0%)	16.02%



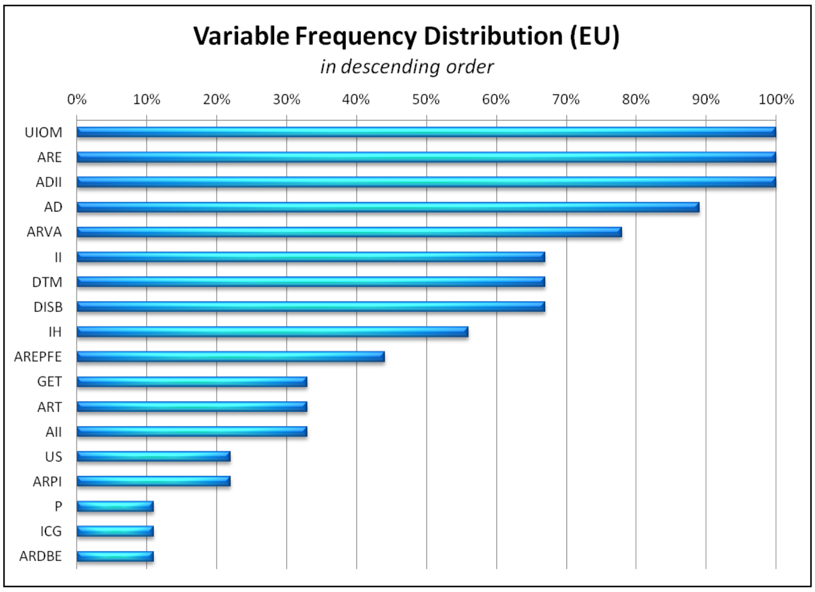
The three zones of significance formulate in the following areas:

- $A: x \in [0\%, 20.2\%]$
- $B: x \in (20.2\%, 84.7\%)$
- $C: x \in [84.7\%, 100\%]$

Variable	*Frequency*	*Frequency Range: hierarchical*		
in descending order		*0%...20%*	*20%...85%*	*85%...100%*
UIOM	*100%*	0	0	1
ARE	*100%*	0	0	1
ADII	*100%*	0	0	1
AD	*89%*	0	0	1
ARVA	*78%*	0	1	0
II	*67%*	0	1	0
DTM	*67%*	0	1	0
DISB	*67%*	0	1	0
IH	*56%*	0	1	0
AREPFE	*44%*	0	1	0
GET	*33%*	0	1	0
ART	*33%*	0	1	0
AII	*33%*	0	1	0
US	*22%*	0	1	0
ARPI	*22%*	0	1	0
P	*11%*	1	0	0
ICG	*11%*	1	0	0
ARDBE	*11%*	1	0	0
	Running Total	*3*	*11*	*4*
EU	**%**	**17%**	**61%**	**22%**



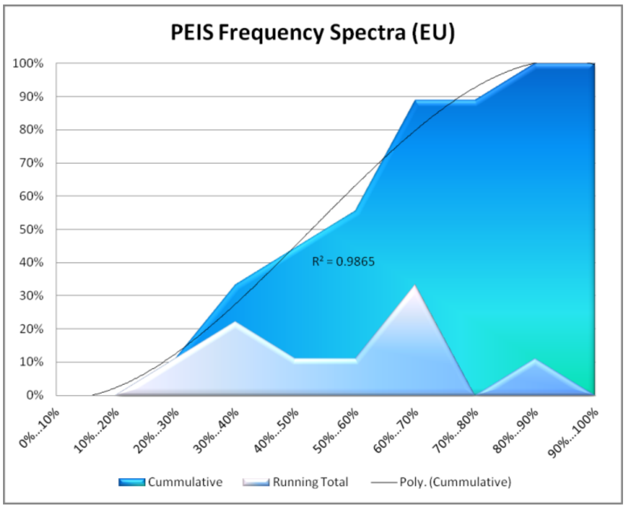
As a first remark, the variable statistics for Europe seem to retain the same trends as the global set.



2.4.2. Study Adequacy Analysis

PEIS	EU Frequency	PEIS range of analysis: Frequency Spectra 0%...10%	10%...20%	20%...30%	30%...40%	40%...50%	50%...60%	60%...70%	70%...80%	80%...90%	90%...100%
BEL01	50%	0	0	0	0	0	1	0	0	0	0
BEL02	61%	0	0	0	0	0	0	1	0	0	0
ITA01	39%	0	0	0	1	0	0	0	0	0	0
ITA02	39%	0	0	0	1	0	0	0	0	0	0
NED01	83%	0	0	0	0	0	0	0	0	1	0
SPA01	28%	0	0	1	0	0	0	0	0	0	0
SPA02	67%	0	0	0	0	0	0	1	0	0	0
UK01	44%	0	0	0	0	1	0	0	0	0	0
UK02	61%	0	0	0	0	0	0	1	0	0	0
	Running Total	*0*	*0*	*1*	*2*	*1*	*1*	*3*	*0*	*1*	*0*
	%	0%	0%	11%	22%	11%	11%	33%	0%	11%	0%
	Cummulative	*0*	*0*	*1*	*3*	*4*	*5*	*8*	*8*	*9*	*9*
	%	0%	0%	11%	33%	44%	56%	89%	89%	100%	100%

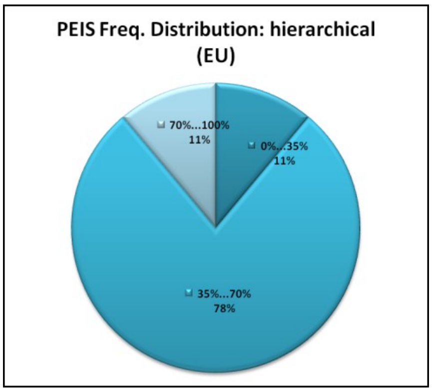
PEIS Statistical Analysis EU	
Mean	52.47%
St. Error	5.72%
Median	50.00%
Mode	38.89%
St. Deviation	17.15%
Sample Var.	2.94%
Kurtosis	-22.53%
Skewness	41.30%
Range	55.56%
Minimum	27.78%
Maximum	83.33%
Count	9
Conf.(95.0%)	13.18%



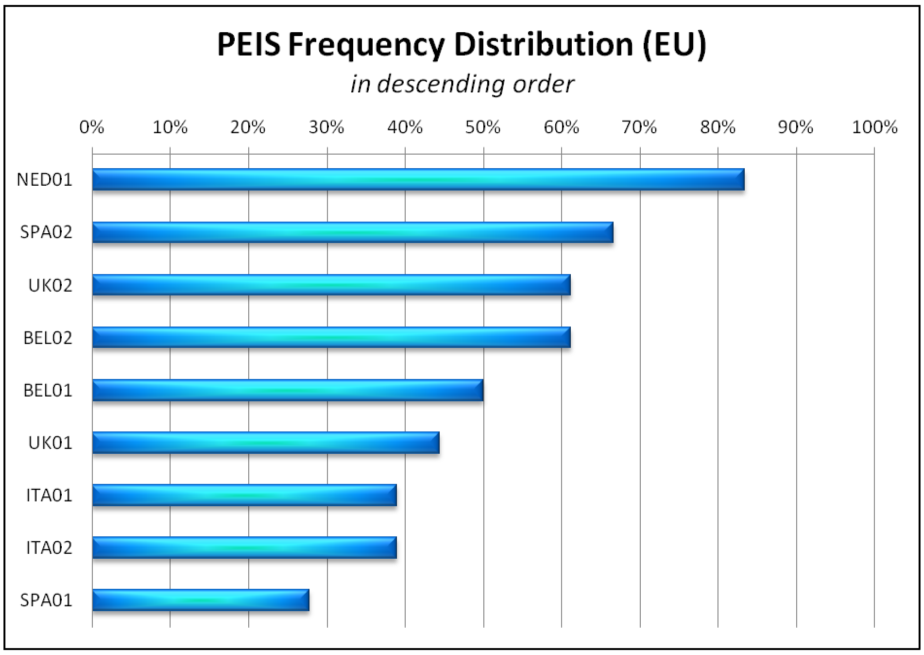
The three zones of significance formulate in the following areas:

- $A: x \in [0\%, 35.3\%]$
- $B: x \in (35.3\%, 69.6\%)$
- $C: x \in [69.6\%, 100\%]$

PEIS	Frequency	Frequency Range: hierarchical		
in descending order		*0%...35%*	*35%...70%*	*70%...100%*
NED01	83%	0	0	1
SPA02	67%	0	1	0
UK02	61%	0	1	0
BEL02	61%	0	1	0
BEL01	50%	0	1	0
UK01	44%	0	1	0
ITA01	39%	0	1	0
ITA02	39%	0	1	0
SPA01	28%	1	0	0
	Running Total	*1*	*7*	*1*
EU	%	11%	78%	11%



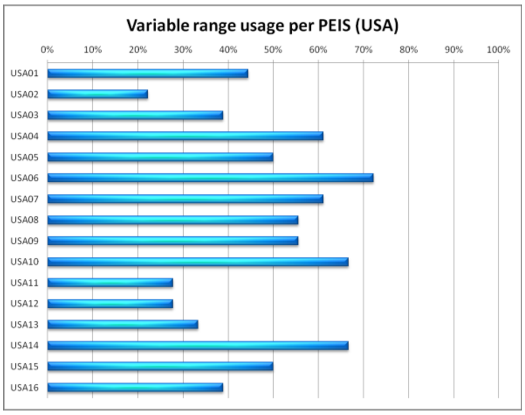
The statistics show clearly the greatest number of studies being within the adequate level, with the Dutch studies exceeding usual quality marks, probably due to a high level of standards in the Dutch port industry. Of course, this is only a hypothesis, as we do not have more samples for comparison.

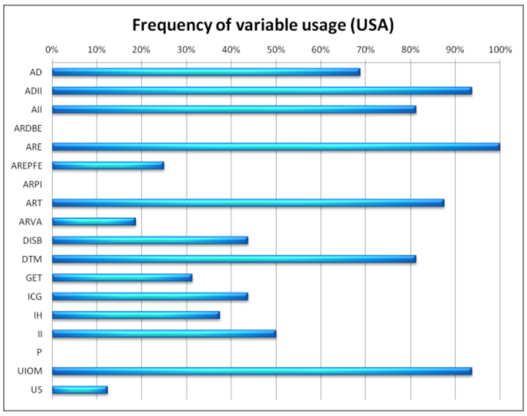


Study adequacy levels range from a low 28% to a high 83%, with the majority of cases in the intermediate region distributed in a relatively close proximity to the mean. The small number of sample cases does not allow us to discern any significant patterns concerning country-specific practices, apart from the single case in the Netherlands. However, the small deviation from the mean signifies a general tendency towards studies utilizing a satisfactory range of descriptive variables.

2.5. Region-specific Data: USA

USA PEIS	AD	ADII	AII	ARDBE	ARE	AREPFE	ARPI	ART	ARVA	DISB	DTM	GET	ICG	IH	II	P	UIOM	US	*Total*	%
USA01	1	1	1	0	1	0	0	1	1	0	1	0	0	0	0	0	1	0	*8*	**44%**
USA02	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	*4*	**22%**
USA03	1	1	1	0	1	1	0	1	0	0	0	0	0	0	0	0	1	0	*7*	**39%**
USA04	1	1	1	0	1	0	0	1	1	1	1	0	1	0	1	0	1	0	*11*	**61%**
USA05	1	1	1	0	1	0	0	1	1	0	1	0	0	0	1	0	1	0	*9*	**50%**
USA06	1	1	1	0	1	1	0	1	0	1	1	1	1	1	1	0	1	0	*13*	**72%**
USA07	1	1	1	0	1	0	0	1	0	1	1	1	1	0	1	0	1	0	*11*	**61%**
USA08	1	1	1	0	1	0	0	1	0	1	1	0	1	0	1	0	1	0	*10*	**56%**
USA09	0	1	1	0	1	1	0	1	0	0	1	1	0	1	0	0	1	1	*10*	**56%**
USA10	1	1	1	0	1	0	0	1	0	1	1	1	1	1	1	0	1	0	*12*	**67%**
USA11	0	1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0	*5*	**28%**
USA12	0	1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0	*5*	**28%**
USA13	1	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	1	*6*	**33%**
USA14	1	1	1	0	1	1	0	1	0	1	1	0	1	1	1	0	1	0	*12*	**67%**
USA15	1	1	1	0	1	0	0	1	0	1	1	0	0	0	1	0	1	0	*9*	**50%**
USA16	0	1	1	0	1	0	0	1	0	0	0	1	0	1	0	0	1	0	*7*	**39%**
Total	*11*	*15*	*13*	*0*	*16*	*4*	*0*	*14*	*3*	*7*	*13*	*5*	*7*	*6*	*8*	*0*	*15*	*2*		
%	**69%**	**94%**	**81%**	**0%**	**100%**	**25%**	**0%**	**88%**	**19%**	**44%**	**81%**	**31%**	**44%**	**38%**	**50%**	**0%**	**94%**	**13%**		





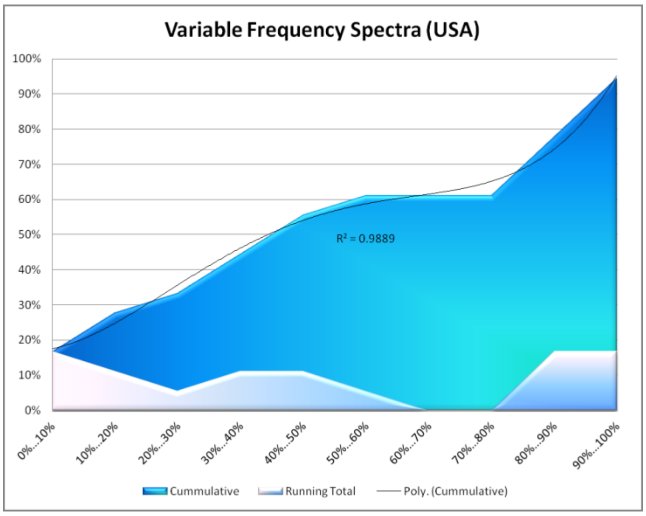
The USA regional set benefits from a greater number of sample studies and an initial image of more pronounced contradictions between preferred variables.

2.5.1. Variable Threshold Analysis

USA		*Variable usage threshold analysis: Frequency Spectra*									
Variable	*Frequency*	*0%...10%*	*10%...20%*	*20%...30%*	*30%...40%*	*40%...50%*	*50%...60%*	*60%...70%*	*70%...80%*	*80%...90%*	*90%...100%*
AD	*69%*	0	0	0	0	0	0	1	0	0	0
ADII	*94%*	0	0	0	0	0	0	0	0	0	1
AII	*81%*	0	0	0	0	0	0	0	0	1	0
ARDBE	*0%*	1	0	0	0	0	0	0	0	0	0
ARE	*100%*	0	0	0	0	0	0	0	0	0	1
AREPFE	*25%*	0	0	1	0	0	0	0	0	0	0
ARPI	*0%*	1	0	0	0	0	0	0	0	0	0
ART	*88%*	0	0	0	0	0	0	0	0	1	0
ARVA	*19%*	0	1	0	0	0	0	0	0	0	0
DISB	*44%*	0	0	0	0	1	0	0	0	0	0
DTM	*81%*	0	0	0	0	0	0	0	0	1	0
GET	*31%*	0	0	0	1	0	0	0	0	0	0
ICG	*44%*	0	0	0	0	1	0	0	0	0	0
IH	*38%*	0	0	0	1	0	0	0	0	0	0
II	*50%*	0	0	0	0	0	1	0	0	0	0
P	*0%*	1	0	0	0	0	0	0	0	0	0
UIOM	*94%*	0	0	0	0	0	0	0	0	0	1
US	*13%*	0	1	0	0	0	0	0	0	0	0
	Running Total	*3*	*2*	*1*	*2*	*2*	*1*	*0*	*0*	*3*	*3*
	%	**17%**	**11%**	**6%**	**11%**	**11%**	**6%**	**0%**	**0%**	**17%**	**17%**
	Cummulative	*3*	*5*	*6*	*8*	*10*	*11*	*11*	*11*	*14*	*17*
	%	**17%**	**28%**	**33%**	**44%**	**56%**	**61%**	**61%**	**61%**	**78%**	**94%**

The contradiction between preferred variables can be clearly discerned from the greater accumulation of frequencies in the lower and the higher ends of the spectrum, while the middle region is dominated by a much less pronounced plateau.

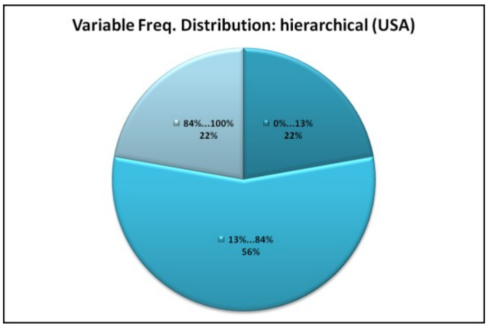
Statistical Analysis	
Mean	48.39%
St. Error	8.30%
Median	44.00%
Mode	0.00%
St. Deviation	35.22%
Sample Var.	12.40%
Kurtosis	-144.25%
Skewness	5.84%
Range	100.00%
Minimum	0.00%
Maximum	100.00%
Count	18
Conf.(95.0%)	17.51%

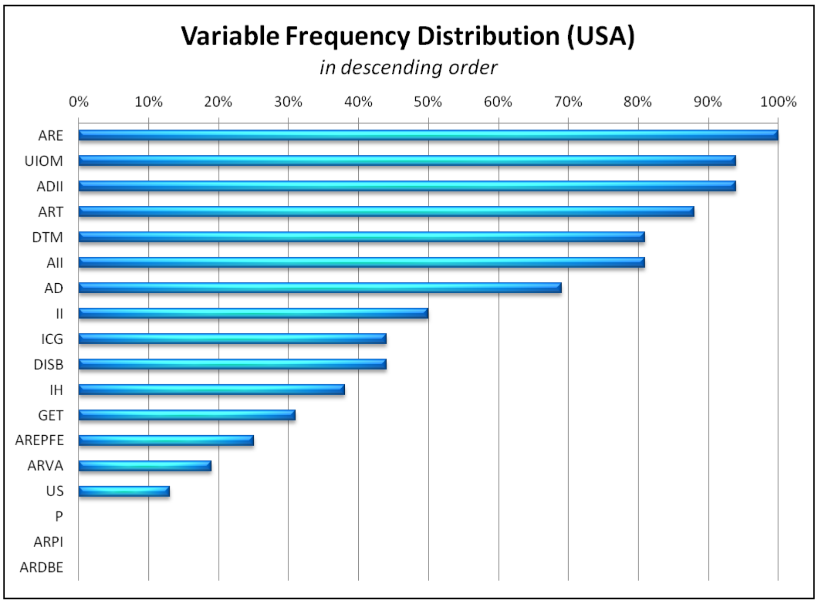


The three zones of significance formulate in the following areas:

- $A: x \in [0\%, 13.2\%]$
- $B: x \in (13.2\%, 83.6\%)$
- $C: x \in [83.6\%, 100\%]$

Variable *in descending order*	***Frequency***	***Frequency Range: hierarchical*** ***0%...13%***	***13%...84%***	***84%...100%***
ARE	*100%*	0	0	1
UIOM	*94%*	0	0	1
ADII	*94%*	0	0	1
ART	*88%*	0	0	1
DTM	*81%*	0	1	0
AII	*81%*	0	1	0
AD	*69%*	0	1	0
II	*50%*	0	1	0
ICG	*44%*	0	1	0
DISB	*44%*	0	1	0
IH	*38%*	0	1	0
GET	*31%*	0	1	0
AREPFE	*25%*	0	1	0
ARVA	*19%*	0	1	0
US	*13%*	1	0	0
P	*0%*	1	0	0
ARPI	*0%*	1	0	0
ARDBE	*0%*	1	0	0
	Running Total	*4*	*10*	*4*
USA	**%**	**22%**	**56%**	**22%**



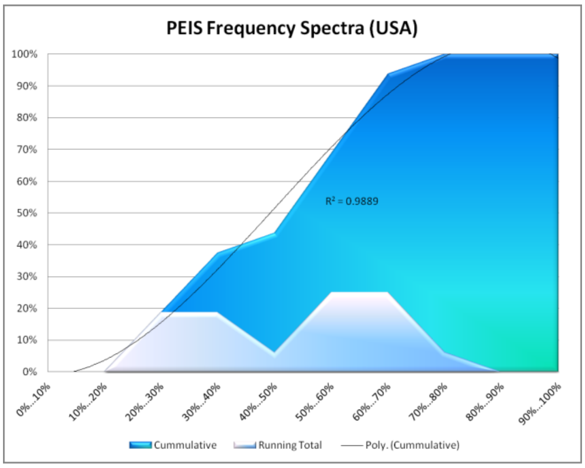


The distribution analysis shows a set of variables that are nearly always present in USA-specific PEIS, while a respective variable set is almost totally ignored. One point that needs to be noted is the lower preference of the *ARVA* value, being located in the lower end of the commonly used variables spectrum, which we would normally regard as a highly important variable for describing the economic status of a port and is highly used in European PEIS.

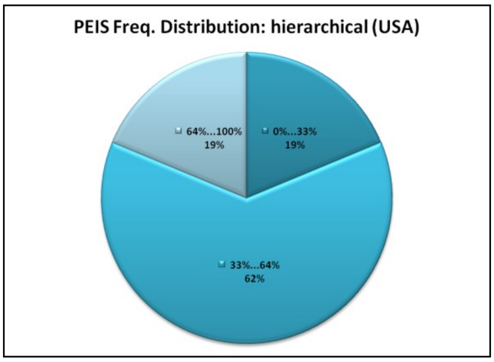
2.5.2. Study Adequacy Analysis

USA		*PEIS range of analysis: Frequency Spectra*									
PEIS	*Frequency*	*0%...10%*	*10%...20%*	*20%...30%*	*30%...40%*	*40%...50%*	*50%...60%*	*60%...70%*	*70%...80%*	*80%...90%*	*90%...100%*
USA01	***44%***	0	0	0	0	1	0	0	0	0	0
USA02	***22%***	0	0	1	0	0	0	0	0	0	0
USA03	***39%***	0	0	0	1	0	0	0	0	0	0
USA04	***61%***	0	0	0	0	0	0	1	0	0	0
USA05	***50%***	0	0	0	0	0	1	0	0	0	0
USA06	***72%***	0	0	0	0	0	0	0	1	0	0
USA07	***61%***	0	0	0	0	0	0	1	0	0	0
USA08	***56%***	0	0	0	0	0	1	0	0	0	0
USA09	***56%***	0	0	0	0	0	1	0	0	0	0
USA10	***67%***	0	0	0	0	0	0	1	0	0	0
USA11	***28%***	0	0	1	0	0	0	0	0	0	0
USA12	***28%***	0	0	1	0	0	0	0	0	0	0
USA13	***33%***	0	0	0	1	0	0	0	0	0	0
USA14	***67%***	0	0	0	0	0	0	1	0	0	0
USA15	***50%***	0	0	0	0	0	1	0	0	0	0
USA16	***39%***	0	0	0	1	0	0	0	0	0	0
	Running Total	*0*	*0*	*3*	*3*	*1*	*4*	*4*	*1*	*0*	*0*
	%	**0%**	**0%**	**19%**	**19%**	**6%**	**25%**	**25%**	**6%**	**0%**	**0%**
	Cummulative	*0*	*0*	*3*	*6*	*7*	*11*	*15*	*16*	*16*	*16*
	%	**0%**	**0%**	**19%**	**38%**	**44%**	**69%**	**94%**	**100%**	**100%**	**100%**

PEIS Statistical Analysis **USA**	
Mean	48.26%
St. Error	3.89%
Median	50.00%
Mode	38.89%
St. Deviation	15.54%
Sample Var.	2.42%
Kurtosis	-115.52%
Skewness	-17.53%
Range	50.00%
Minimum	22.22%
Maximum	72.22%
Count	16
Conf.(95.0%)	8.28%



PEIS	*Frequency*	*Frequency Range: hierarchical*		
in descending order		*0%...33%*	*33%...64%*	*64%...100%*
USA06	***72%***	0	0	1
USA10	***67%***	0	0	1
USA14	***67%***	0	0	1
USA04	***61%***	0	1	0
USA07	***61%***	0	1	0
USA08	***56%***	0	1	0
USA09	***56%***	0	1	0
USA05	***50%***	0	1	0
USA15	***50%***	0	1	0
USA01	***44%***	0	1	0
USA03	***39%***	0	1	0
USA16	***39%***	0	1	0
USA13	***33%***	0	1	0
USA11	***28%***	1	0	0
USA12	***28%***	1	0	0
USA02	***22%***	1	0	0
	Running Total	*3*	*10*	*3*
USA	**%**	**19%**	**63%**	**19%**



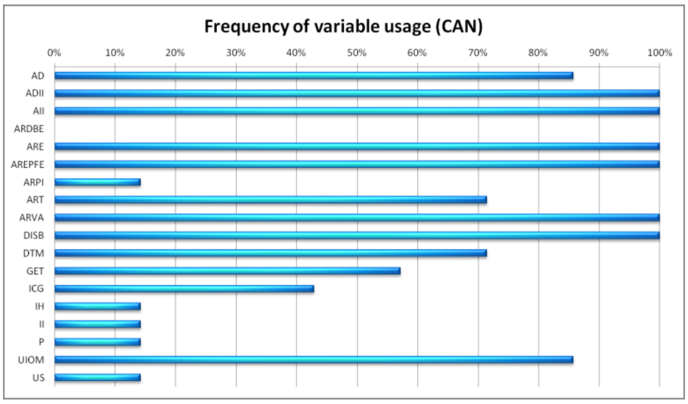
The three zones of significance formulate in the following areas:

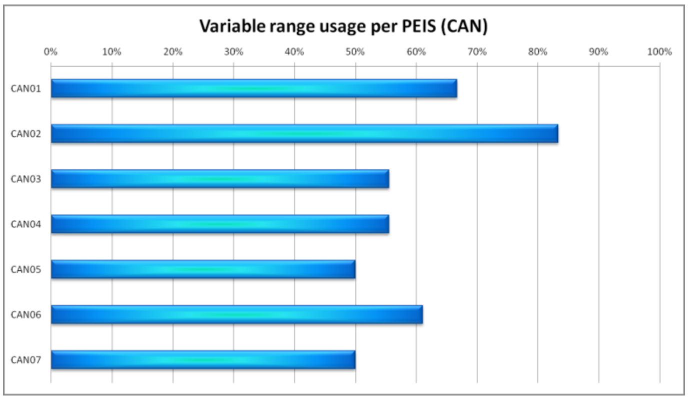
- $A: x \in [0\%, 32.7\%]$
- $B: x \in (32.7\%, 63.8\%)$
- $C: x \in [63.8\%, 100\%]$

The American tendency of relying predominantly in a certain set of distributed variables while completely or nearly omitting another set, in contrast to the European practices, leads to a distribution of study adequacy levels that exhibits a significant number of studies with relatively high adequacy levels - but not as high as the European studies – and the trailing majority of studies of relative middle-level adequacy, while an equally significant portion of studies can be summarily characterized as laconic. This is also evident from the frequency distribution that, when compared to the European one, seems shifted a certain distance to the left and the corresponding cumulative distribution reaches its maximum plateau sooner.

2.6. Region-specific Data: Canada

	AD	ADII	AII	ARDBE	ARE	AREPFE	ARPI	ART	ARVA	DISB	DTM	GET	ICG	IH	II	P	UIOM	US	*Total*	%
CAN01	1	1	1	0	1	1	0	1	1	1	1	1	1	0	0	0	1	0	*12*	**67%**
CAN02	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	0	*15*	**83%**
CAN03	1	1	1	0	1	1	0	0	1	1	1	0	1	0	0	0	1	0	*10*	**56%**
CAN04	0	1	1	0	1	1	0	1	1	1	1	1	0	0	0	0	0	1	*10*	**56%**
CAN05	1	1	1	0	1	1	0	1	1	1	0	0	0	0	0	0	1	0	*9*	**50%**
CAN06	1	1	1	0	1	1	0	1	1	1	0	1	0	0	0	1	1	0	*11*	**61%**
CAN07	1	1	1	0	1	1	0	0	1	1	1	0	0	0	0	0	1	0	*9*	**50%**
Total	*6*	*7*	*7*	*0*	*7*	*7*	*1*	*5*	*7*	*7*	*5*	*4*	*3*	*1*	*1*	*1*	*6*	*1*		
%	**86%**	**100%**	**100%**	**0%**	**100%**	**100%**	**14%**	**71%**	**100%**	**100%**	**71%**	**57%**	**43%**	**14%**	**14%**	**14%**	**86%**	**14%**		



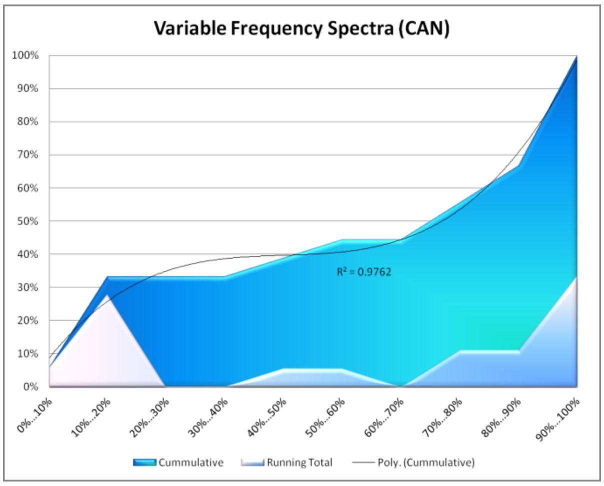


The Canadian practices seemingly follow the American tendency of emphasizing certain variable sets, albeit different ones, while their study adequacy levels seem uniformly greater than average. This logically stems from the fact that the Canadian PEIS are based on concrete guidelines issued by the country's Bureau of Transport.

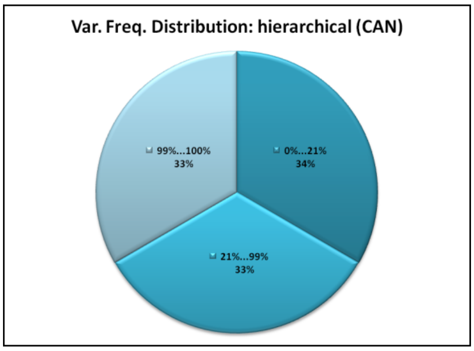
2.6.1. Variable Threshold Analysis

CAN		Variable usage threshold analysis: Frequency Spectra									
Variable	*Frequency*	*0%...10%*	*10%...20%*	*20%...30%*	*30%...40%*	*40%...50%*	*50%...60%*	*60%...70%*	*70%...80%*	*80%...90%*	*90%...100%*
AD	*86%*	0	0	0	0	0	0	0	0	1	0
ADII	*100%*	0	0	0	0	0	0	0	0	0	1
AII	*100%*	0	0	0	0	0	0	0	0	0	1
ARDBE	*0%*	1	0	0	0	0	0	0	0	0	0
ARE	*100%*	0	0	0	0	0	0	0	0	0	1
AREPFE	*100%*	0	0	0	0	0	0	0	0	0	1
ARPI	*14%*	0	1	0	0	0	0	0	0	0	0
ART	*71%*	0	0	0	0	0	0	0	1	0	0
ARVA	*100%*	0	0	0	0	0	0	0	0	0	1
DISB	*100%*	0	0	0	0	0	0	0	0	0	1
DTM	*71%*	0	0	0	0	0	0	0	1	0	0
GET	*57%*	0	0	0	0	0	1	0	0	0	0
ICG	*43%*	0	0	0	0	1	0	0	0	0	0
IH	*14%*	0	1	0	0	0	0	0	0	0	0
II	*14%*	0	1	0	0	0	0	0	0	0	0
P	*14%*	0	1	0	0	0	0	0	0	0	0
UIOM	*86%*	0	0	0	0	0	0	0	0	1	0
US	*14%*	0	1	0	0	0	0	0	0	0	0
	Running Total	*1*	*5*	*0*	*0*	*1*	*1*	*0*	*2*	*2*	*6*
	%	**6%**	**28%**	**0%**	**0%**	**6%**	**6%**	**0%**	**11%**	**11%**	**33%**
	Cummulative	*1*	*6*	*6*	*6*	*7*	*8*	*8*	*10*	*12*	*18*
	%	**6%**	**33%**	**33%**	**33%**	**39%**	**44%**	**44%**	**56%**	**67%**	**100%**

Statistical Analysis **CAN**	
Mean	60.22%
St. Error	9.16%
Median	71.00%
Mode	100.00%
St. Deviation	38.86%
Sample Var.	15.10%
Kurtosis	-170.11%
Skewness	-34.16%
Range	100.00%
Minimum	0.00%
Maximum	100.00%
Count	18
Conf.(95.0%)	19.32%



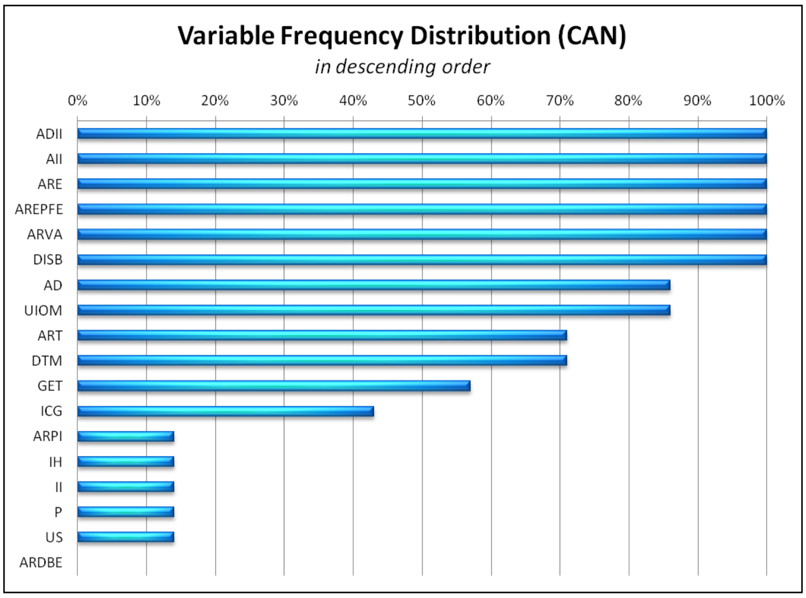
Variable	*Frequency*	*Frequency Range: hierarchical*		
in descending order		*0%...21%*	*21%...99%*	*99%...100%*
ADII	*100%*	0	0	1
AII	*100%*	0	0	1
ARE	*100%*	0	0	1
AREPFE	*100%*	0	0	1
ARVA	*100%*	0	0	1
DISB	*100%*	0	0	1
AD	*86%*	0	1	0
UIOM	*86%*	0	1	0
ART	*71%*	0	1	0
DTM	*71%*	0	1	0
GET	*57%*	0	1	0
ICG	*43%*	0	1	0
ARPI	*14%*	1	0	0
IH	*14%*	1	0	0
II	*14%*	1	0	0
P	*14%*	1	0	0
US	*14%*	1	0	0
ARDBE	*0%*	1	0	0
	Running Total	*6*	*6*	*6*
CAN	**%**	**33%**	**33%**	**33%**



The three zones of significance formulate in the following areas:

- $A: x \in [0\%, 21.4\%]$
- $B: x \in (21.4\%, 99.1\%)$
- $C: x \in [99.1\%, 100\%]$

This clearly underlines the Canadian preset variables of preference, corresponding to a set of six variables that are always employed. This set of variables deals predominantly with direct, indirect and induced impacts, employment levels, value added and detailed information on sectoral boundaries. A second, equal in number, set of variables is used auxiliary to broaden the descriptive range of the study, while the last set of variables clearly falls in disfavor. This can also be discerned from the gentle slope of the cumulative distribution, as well as from the location of peaks in the frequency distribution.

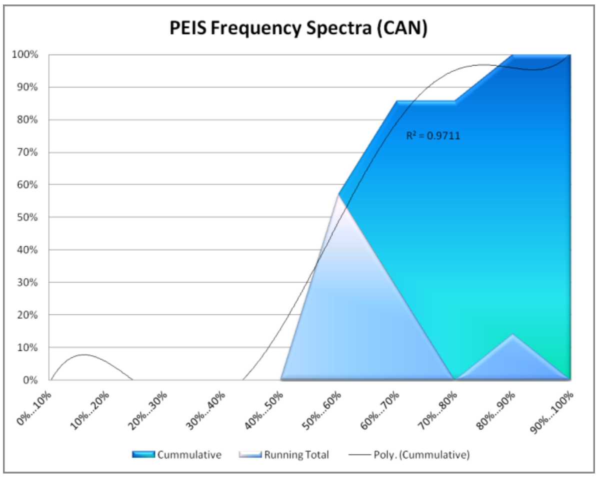


2.6.2. Study Adequacy Analysis

CAN		PEIS range of analysis: Frequency Spectra									
PEIS	*Frequency*	*0%...10%*	*10%...20%*	*20%...30%*	*30%...40%*	*40%...50%*	*50%...60%*	*60%...70%*	*70%...80%*	*80%...90%*	*90%...100%*
CAN01	67%	0	0	0	0	0	0	1	0	0	0
CAN02	83%	0	0	0	0	0	0	0	0	1	0
CAN03	56%	0	0	0	0	0	1	0	0	0	0
CAN04	56%	0	0	0	0	0	1	0	0	0	0
CAN05	50%	0	0	0	0	0	1	0	0	0	0
CAN06	61%	0	0	0	0	0	0	1	0	0	0
CAN07	50%	0	0	0	0	0	1	0	0	0	0
	Running Total	*0*	*0*	*0*	*0*	*0*	*4*	*2*	*0*	*1*	*0*
	%	**0%**	**0%**	**0%**	**0%**	**0%**	**57%**	**29%**	**0%**	**14%**	**0%**
	Cummulative	*0*	*0*	*0*	*0*	*0*	*4*	*6*	*6*	*7*	*7*
	%	**0%**	**0%**	**0%**	**0%**	**0%**	**57%**	**86%**	**86%**	**100%**	**100%**

The initial inspection in the frequency spectra shows a significant shift to the right end of the spectrum, which can also be witnessed in the cumulative distribution, marking the high adequacy levels of Canadian practices.

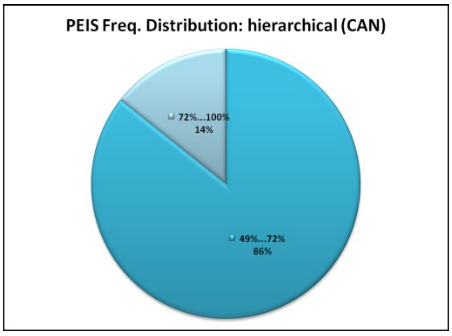
PEIS Statistical Analysis CAN	
Mean	60.32%
St. Error	4.44%
Median	55.56%
Mode	55.56%
St. Deviation	11.75%
Sample Var.	1.38%
Kurtosis	208.00%
Skewness	144.21%
Range	33.33%
Minimum	50.00%
Maximum	83.33%
Count	7
Conf.(95.0%)	10.87%

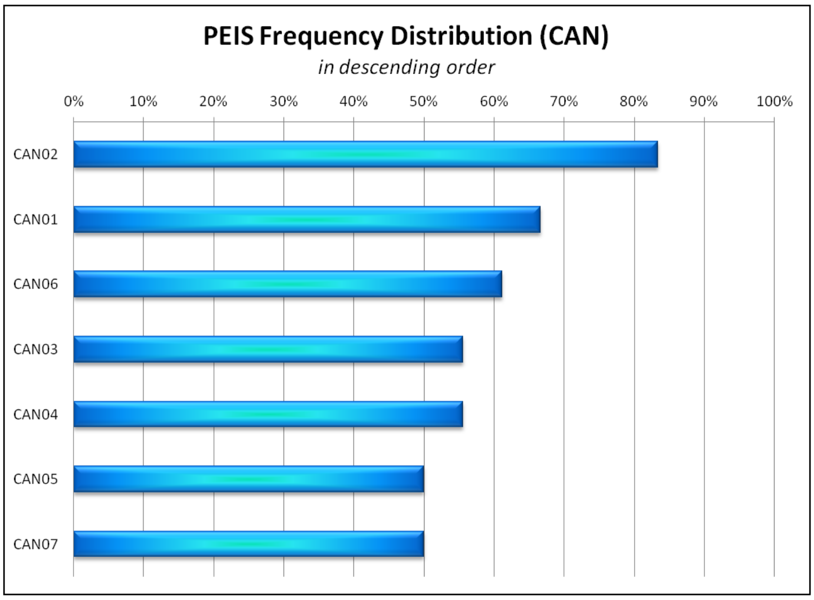


The three zones of significance formulate in the following areas:

- $A: x \in [0\%, 48.6\%]$
- $B: x \in (48.6\%, 72.1\%)$
- $C: x \in [72.1\%, 100\%]$

PEIS	*Frequency*	*Frequency Range: hierarchical*		
in descending order		*0%...49%*	*49%...72%*	*72%...100%*
CAN02	83%	0	0	1
CAN01	67%	0	1	0
CAN06	61%	0	1	0
CAN03	56%	0	1	0
CAN04	56%	0	1	0
CAN05	50%	0	1	0
CAN07	50%	0	1	0
	Running Total	*0*	*6*	*1*
CAN	**%**	**0%**	**86%**	**14%**

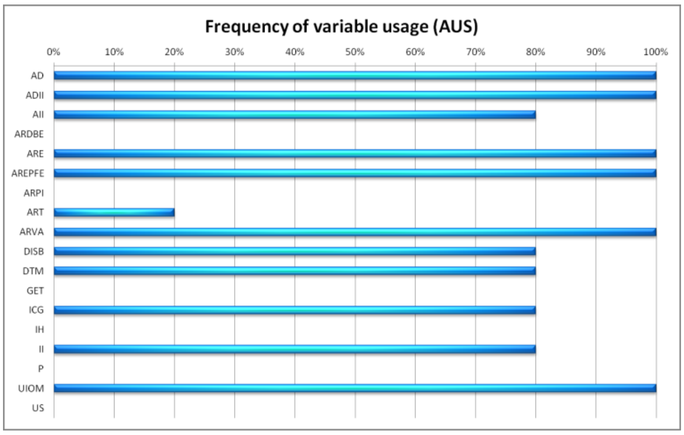


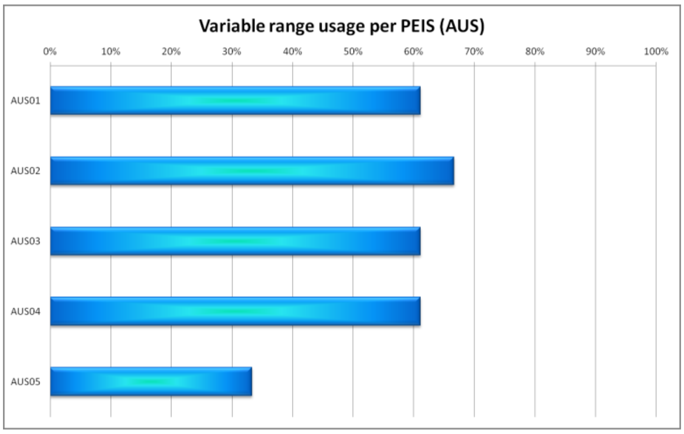


The resulting distribution marks the totality of Canadian PEIS as above average adequacy standards and underlines a continuity of standardized practices in the region. It is also very interesting that there are no sample studies bearing a below average evaluation, signifying an emphasis towards descriptive cohesiveness.

2.7. Region-specific Data: Australia

	AD	ADII	AII	ARDBE	ARE	AREPFE	ARPI	ART	ARVA	DISB	DTM	GET	ICG	IH	II	P	UIOM	US	*Total*	%
AUS01	1	1	1	0	1	1	0	0	1	1	1	0	1	0	1	0	1	0	*11*	**61%**
AUS02	1	1	1	0	1	1	0	1	1	1	1	0	1	0	1	0	1	0	*12*	**67%**
AUS03	1	1	1	0	1	1	0	0	1	1	1	0	1	0	1	0	1	0	*11*	**61%**
AUS04	1	1	1	0	1	1	0	0	1	1	1	0	1	0	1	0	1	0	*11*	**61%**
AUS05	1	1	0	0	1	1	0	0	1	0	0	0	0	0	0	0	1	0	*6*	**33%**
Total	*5*	*5*	*4*	*0*	*5*	*5*	*0*	*1*	*5*	*4*	*4*	*0*	*4*	*0*	*4*	*0*	*5*	*0*		
%	100%	100%	80%	0%	100%	100%	0%	20%	100%	80%	80%	0%	80%	0%	80%	0%	100%	0%		



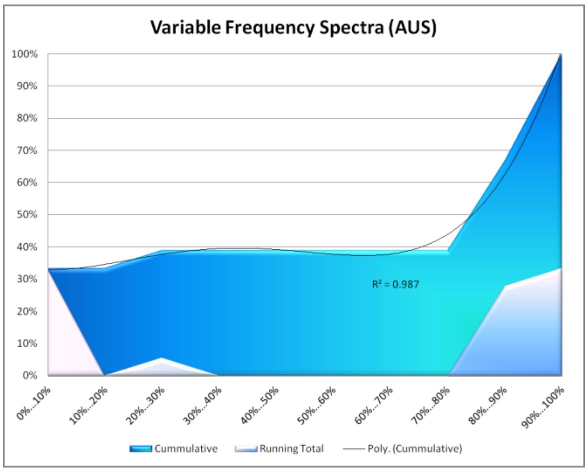


A first inspection of the Australian case studies reveals patterns that hint of a merging of the Canadian and American practices; this may actually be a development of the aforementioned systems, as it follows the policy of exclusive predominance of a closed set of variables, which is almost identical to the Canadian set, with the difference being that the Australian set focuses on input-output methods (classic theoretical model), while the Canadian set shifts the focus on impact assessment. Moreover, the Australian studies seem even more compact and laconic than the American ones, by omitting more variables from their common framework.

2.7.1. Variable Threshold Analysis

AUS		Variable usage threshold analysis: Frequency Spectra									
Variable	***Frequency***	***0%...10%***	***10%...20%***	***20%...30%***	***30%...40%***	***40%...50%***	***50%...60%***	***60%...70%***	***70%...80%***	***80%...90%***	***90%...100%***
AD	100%	0	0	0	0	0	0	0	0	0	1
ADII	100%	0	0	0	0	0	0	0	0	0	1
AII	80%	0	0	0	0	0	0	0	0	1	0
ARDBE	0%	1	0	0	0	0	0	0	0	0	0
ARE	100%	0	0	0	0	0	0	0	0	0	1
AREPFE	100%	0	0	0	0	0	0	0	0	0	1
ARPI	0%	1	0	0	0	0	0	0	0	0	0
ART	20%	0	0	1	0	0	0	0	0	0	0
ARVA	100%	0	0	0	0	0	0	0	0	0	1
DISB	80%	0	0	0	0	0	0	0	0	1	0
DTM	80%	0	0	0	0	0	0	0	0	1	0
GET	0%	1	0	0	0	0	0	0	0	0	0
ICG	80%	0	0	0	0	0	0	0	0	1	0
IH	0%	1	0	0	0	0	0	0	0	0	0
II	80%	0	0	0	0	0	0	0	0	1	0
P	0%	1	0	0	0	0	0	0	0	0	0
UIOM	100%	0	0	0	0	0	0	0	0	0	1
US	0%	1	0	0	0	0	0	0	0	0	0
	Running Total	*6*	*0*	*1*	*0*	*0*	*0*	*0*	*0*	*5*	*6*
	%	**33%**	**0%**	**6%**	**0%**	**0%**	**0%**	**0%**	**0%**	**28%**	**33%**
	Cummulative	*6*	*6*	*7*	*7*	*7*	*7*	*7*	*7*	*12*	*18*
	%	**33%**	**33%**	**39%**	**39%**	**39%**	**39%**	**39%**	**39%**	**67%**	**100%**

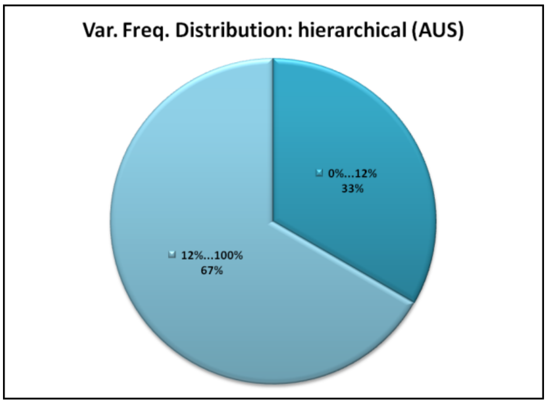
Statistical Analysis **AUS**	
Mean	56.67%
St. Error	10.63%
Median	80.00%
Mode	100.00%
St. Deviation	45.11%
Sample Var.	20.35%
Kurtosis	-186.59%
Skewness	-42.47%
Range	100.00%
Minimum	0.00%
Maximum	100.00%
Count	18
Conf.(95.0%)	22.43%



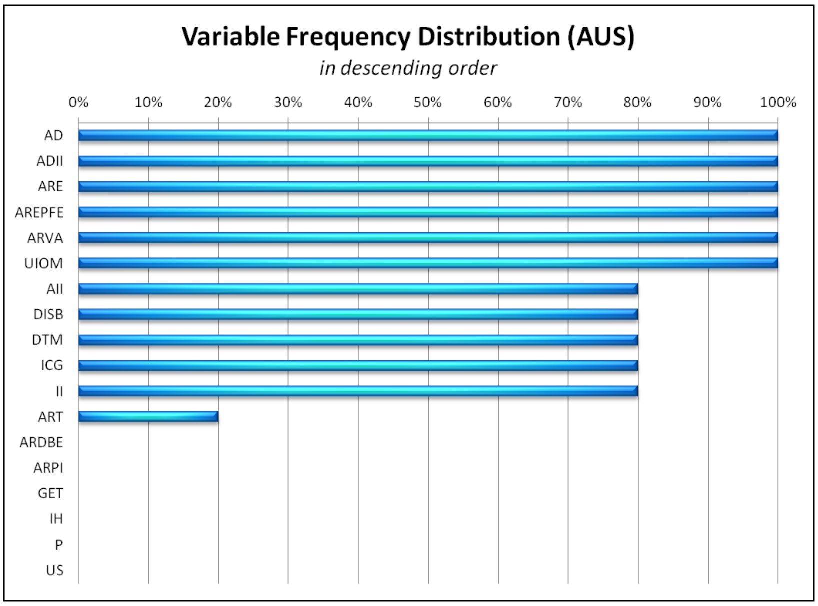
As $\bar{x} + s = 101.8\%$, well over the upper limit, the zones of significance formulate in the following areas:

- $A: x \in [0\%, 11.6\%]$
- $B: x \in [11.6\%, 100\%]$

Variable	***Frequency*** *in descending order*	***Freq. Range: hierarchical*** ***0%...12%***	***12%...100%***
AD	100%	0	1
ADII	100%	0	1
ARE	100%	0	1
AREPFE	100%	0	1
ARVA	100%	0	1
UIOM	100%	0	1
AII	80%	0	1
DISB	80%	0	1
DTM	80%	0	1
ICG	80%	0	1
II	80%	0	1
ART	20%	0	1
ARDBE	0%	1	0
ARPI	0%	1	0
GET	0%	1	0
IH	0%	1	0
P	0%	1	0
US	0%	1	0
	Running Total	*6*	*12*
AUS	**%**	**33%**	**67%**



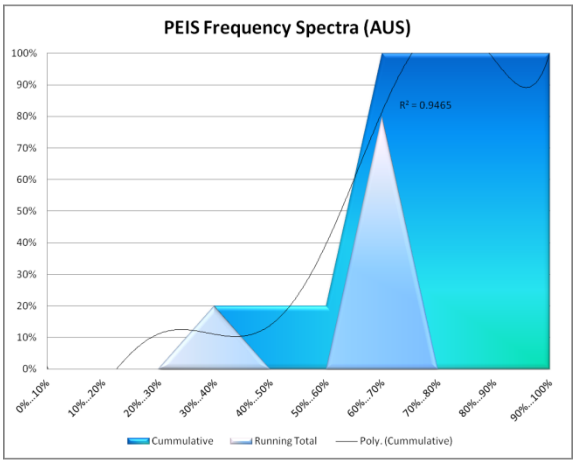
This collapse of the original three significance zones into two indicates the polarization in variable usage, where a certain set of variables is predominant while the rest are mostly omitted.



2.7.2. Study Adequacy Analysis

AUS		PEIS range of analysis: Frequency Spectra									
PEIS	*Frequency*	*0%...10%*	*10%...20%*	*20%...30%*	*30%...40%*	*40%...50%*	*50%...60%*	*60%...70%*	*70%...80%*	*80%...90%*	*90%...100%*
AUS01	61%	0	0	0	0	0	0	1	0	0	0
AUS02	67%	0	0	0	0	0	0	1	0	0	0
AUS03	61%	0	0	0	0	0	0	1	0	0	0
AUS04	61%	0	0	0	0	0	0	1	0	0	0
AUS05	33%	0	0	0	1	0	0	0	0	0	0
	Running Total	*0*	*0*	*0*	*1*	*0*	*0*	*4*	*0*	*0*	*0*
	%	**0%**	**0%**	**0%**	**20%**	**0%**	**0%**	**80%**	**0%**	**0%**	**0%**
	Cummulative	*0*	*0*	*0*	*1*	*1*	*1*	*5*	*5*	*5*	*5*
	%	**0%**	**0%**	**0%**	**20%**	**20%**	**20%**	**100%**	**100%**	**100%**	**100%**

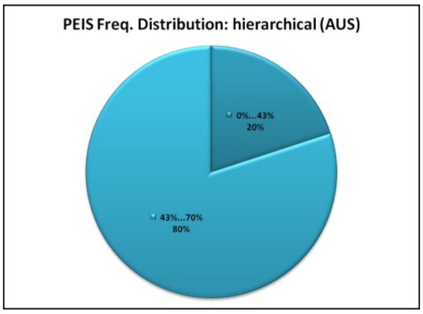
PEIS Statistical Analysis **AUS**	
Mean	56.67%
St. Error	5.93%
Median	61.11%
Mode	61.11%
St. Deviation	13.26%
Sample Var.	1.76%
Kurtosis	442.29%
Skewness	-204.28%
Range	33.33%
Minimum	33.33%
Maximum	66.67%
Count	5
Conf.(95.0%)	16.47%

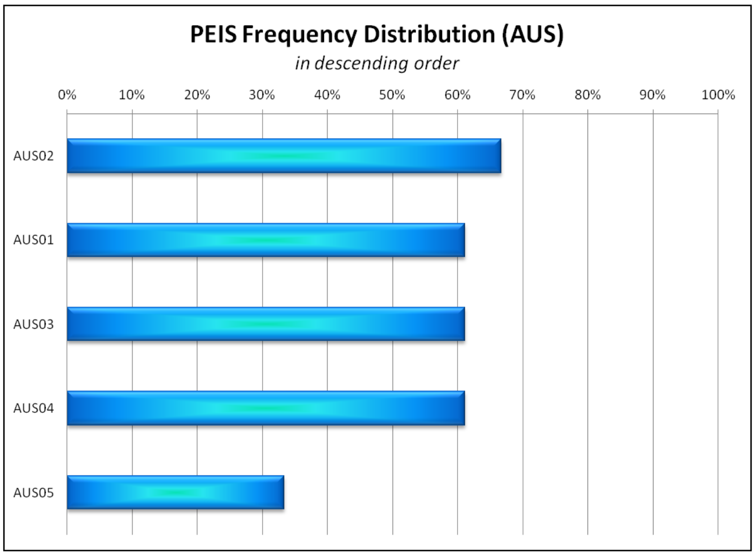


The three zones of significance formulate in the following spaces:

- $A: x \in [0\%, 43.4\%]$
- $B: x \in (43.4\%, 69.9\%)$
- $C: x \in [69.9\%, 100\%]$

PEIS	Frequency	Frequency Range: hierarchical		
in descending order		*0%...43%*	*43%...70%*	*70%...100%*
AUS02	67%	0	1	0
AUS01	61%	0	1	0
AUS03	61%	0	1	0
AUS04	61%	0	1	0
AUS05	33%	1	0	0
	Running Total	*1*	*4*	*0*
AUS	**%**	**20%**	**80%**	**0%**





The shift to the left of the Australian PEIS adequacy distribution is most probably a result of the greater number of omitted variables, when compared to the global set. It follows that a different picture would have emerged if the analysis had been conducted strictly through the set of variables employed in Australian studies, removing from the distribution variables that are never employed in this region. Thus, the distribution of PEIS adequacy level has to be readjusted and the zone of common occurrence needs to be considered as a zone of high occurrence instead. In this perspective the zone of negligible occurrence is also rebranded as the zone of common occurrence, thus outlining a most satisfactory regional score.

2.8. Chronological Considerations

The overwhelming majority of the case studies examined belong to a chronological period corresponding to the first decade of the 21st century. However, there were two case studies included in our set that were significantly distanced from this time frame: specifically, the case study of the port of Liverpool (***UK01***) was conducted in 1979 and the case study of the port of Fraser (***CAN03***) in 1992. This chronological distance forces us to re-examine separately these cases in order to evaluate whether there are significant differences in the observed statistics and variable usage, when compared to the more general regional patterns corresponding to each one.

First, we examine the study adequacy levels of each case and its placement in the corresponding regional statistics. Adequacy levels of the Canadian study ***CAN03*** of 1992 rise up to 55.56%, which places the study in the middle range of its regional statistics, namely the zone of common occurrence ranging from 48.6% to 72.1% for the Canadian region. The British study ***UK01*** of 1979 rises up to an adequacy level of 44.44%, which likewise belongs in the middle range of the European statistics, ranging from 35.3% to 69.6%. Conclusively, both of the studies do not pose significant variations of the overall statistics corresponding to their respective regions, albeit they are placed on the lower end of the spectrum of their common occurrence zones. The same holds true when compared to the global statistics, which exhibit a middle range from 37.5% to 67.9%.

However, chronological differences are more evident when considering variable usage preferences in the corresponding regions. The Canadian study follows the general pattern of the region, with the subtle difference of omitting two commonly used variables, namely *availability of results for taxes* (***ART***) and *growth expressed in TEUs or other units* (***GET***), having frequencies of 71% and 57% respectively in the regional statistics, 62% and 32% in the global ones.

The British study exhibits the same tendencies in a more pronounced scale, omitting common variables of the middle zone of the regional frequency distribution, namely ***AII*** (33%), ***ARPI*** (22%), ***ART*** (33%), ***ARVA*** (67%), ***DISB*** (56%), ***GET*** (33%) and ***IH*** (44%). This is also evident from the global statistics, where the above common variables obtain scores of 73%, 8%, 62%, 59%, 65% and 32% respectively, marking them as commonly used variables of an auxiliary descriptive character, except in the case of the almost negligible ***ARPI*** (14% and 8% globally).

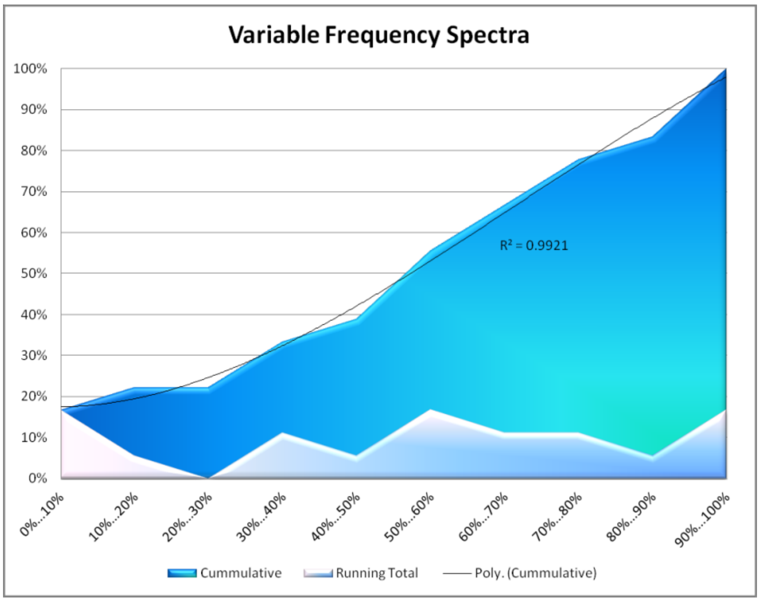
The overall conclusion regarding chronological considerations is that earlier studies probably did not feel the need to provide extensive descriptive scenarios, relying mostly on easily accessible sets of data.

3. Summary and Conclusions

The preceding analysis on global and regional scales has allowed us to formulate a number of conclusions based on our original objectives. Such conclusions stem from the identification of variable sets commonly used in PEIS, marked by frequency patterns that retain a level of statistical significance; these patterns also retain a cohesive conceptual value, thus furnishing individual studies with degrees of objectivity and validity. Through the identification of such variable sets and the determination of their statistical significance, insights towards the scope of such studies and usually the policies or guidelines prevalent in a region can be acquired.

3.1. General Observations

Global patterns of variable employment do not necessarily coincide with individual regional patterns, but still retain most of the individual dynamics. However, due to the extended range of data and the unequal number of case studies per region of origin, there are hints of obscuring or smoothing out of statistically discrete variable sets, as identified in the consecutive regional analysis. This can be further verified by observing the global statistics, which center on a *mean* of 52.67% and disperse with a *standard deviation* of 31.87%, both values being very close to the theoretical standards of a normal distribution. Furthermore, frequency histograms appear spread out uniformly through the spectrum and the cumulative distribution appears smoothed out and of a gentler inclination: *kurtosis* value of -108.53%, *skewness* value of -16.60%, its best fit curve provided by a spline with $R^2 \cong 99\%$.



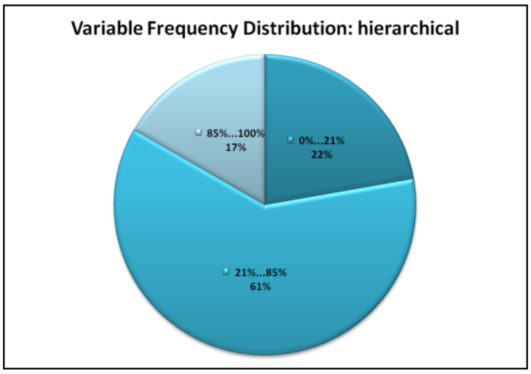
By utilizing the primary statistics, we have defined zones of high, medium and low significance concerning variable frequency, according to the concepts outlined in our methodology. Thus, in the zone of high significance we encounter three variables:

- Availability of results for employment (***ARE***): 100% frequency of occurrence
- Availability of results for direct and indirect impacts (***ADII***): 97%
- Use of Input-Output methods (***UIOM***): 95%

Similarly, in the low zone of negligible significance we encounter variables that are seldom used:

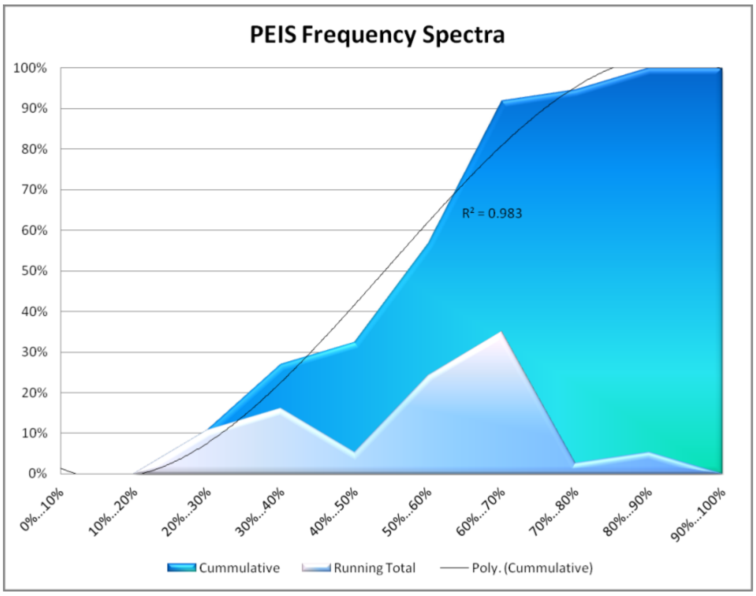
- Use of surveys (***US***): 14%
- Availability of results for private investments (ARPI): 8%
- Productivity (***P***): 5%
- Availability of results for development of business establishments(ARDBE): 3%

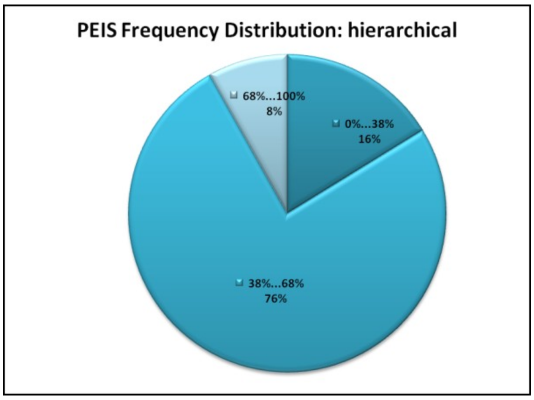
As the variable *Use of surveys* (***US***) has already been allocated with an inverse effect, due to debatable objectivity questions, meaning that a positive value actually signifies the use of surveys and questionnaires, its presence in the low zone corresponds to an actual survey use by the majority of 86% of our case studies. Still, we consider the practice of survey use as detrimental towards the objectivity of PEIS and one of the reasons that such studies have met with considerable criticism in the past; possible alternatives include fully automated systems of data acquisition that can be solely employed within the port boundaries themselves and only for the express requirements of PEIS, as dictated by the port authorities themselves and probably enjoying the legal cover of anonymity.



The rest eleven variables, out of a total of eighteen, fall in the middle zone of significance, ranging from frequencies of 32% to 81%. Whether or not they are used for auxiliary descriptive purposes is quite difficult to discern, due to the effects of regional patterns overlapping explained above. However, it might be safe to assume that variables of a reasonably high frequency also have statistical significance, the inverse holding true for variables at the low end of the medium zone. Pattern overlapping may also be justified by noticing the distribution patterns in the global variable frequency hierarchical chart, where the middle region consumes roughly ⅔ of the distribution, again reminiscent of the normal distribution.

Moving on to the studies' adequacy level analysis, which examines the range of variable coverage employed in PEIS, we find out that the statistics form a different picture than the one in variable frequency analysis. With the *mean* centering on 52.70% and a *standard deviation* defining a narrow band of 15.19%, the greatest fraction of studies belongs in the middle zone of statistical significance, defining PEIS with adequacy levels ranging from 37.5% to 67.9%. This can also be discerned from the frequency histogram, which exhibits peak areas around the middle part of the spectrum, and the cumulative distribution, which rises quite steeply to reach an early plateau around the 68% mark, with a *kurtosis* of -43.77% and -14.82% *skewness*.





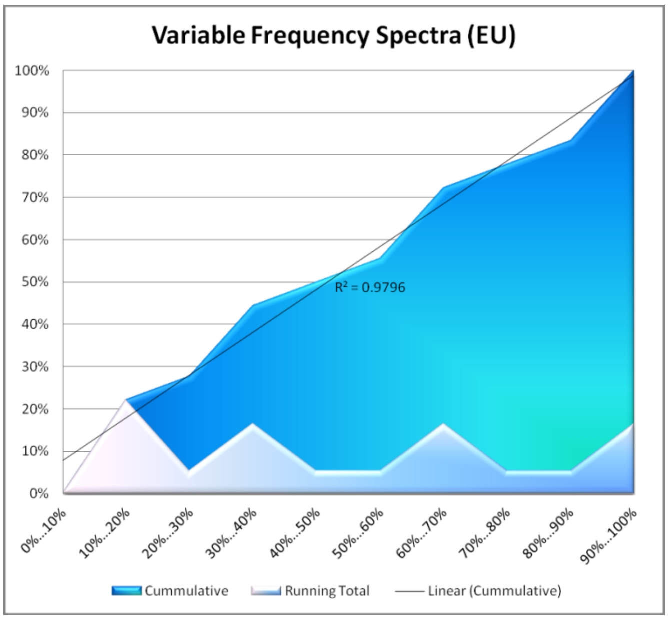
Continuing on the statistical distribution of PEIS adequacy levels in a global setting, we verify the predominance of middle-adequacy studies by attributing the greatest majority of 28 studies in the zone of medium significance, contributing to 76% of the sample size, while the zone of negligible significance totals 6 cases and 16% of the sample size, and the zone of high significance corresponds to only 3 cases and 8% of the sample size. PEIS that appear in the high end of the adequacy spectrum originate from the regions of Canada, the USA and the Netherlands of Europe; conversely, the low end of the adequacy spectrum is dominated by USA studies, as well as single cases from Australia and Spain of Europe. This will be discussed in the following paragraphs, as it is dependent on the discrete patterns encountered in each region.

3.2. Region-specific Observations

Moving on to regional settings, we come across pronounced differences between individual regions and also between regions and the global pattern itself. These are differences in variable set preferences, underlying philosophies or official guidelines set by each region separately; consequently they also affect study adequacy levels, by imposing limitations on specific variable usage, which limit the available set when compared to global perspectives. Such decisive differences between regional settings account for the special considerations concerning the global setting and the limited insight that it provides.

3.2.1. Europe

High frequencies of variable usage acts as a general trend in the European setting, exhibiting clear distinctions between variables that are almost always used and variables that are, as a rule, avoided, but at the same time the situation with middle-range variables cannot be clearly defined. Variable frequency statistics place the *mean* at 52.44% with a comparable *standard deviation* of 32.21%, raising and lowering the mark for high and negligible occurrence respectively at 84.7% and 20.2%. This leaves an extended "grey" area in the middle ranges, where variable frequencies disperse between the respective high and low limits. The definitive allocation of high-end and low-end variables can also be discerned in the frequency histogram, where early peaks appear at the far left side of the spectrum, mirrored by late peaks at the right side. However, peaks appearing in the middle range are also polarized towards the respective spectral limits, consistent with the high *standard deviation* value of the statistics. This also affects the cumulative distribution, which rises in a relatively linear progression: *kurtosis* -137.73% and *skewness* 20.76% and a linear fit with $R^2 \cong 98\%$.



In the zone of high frequency we encounter the following variables:

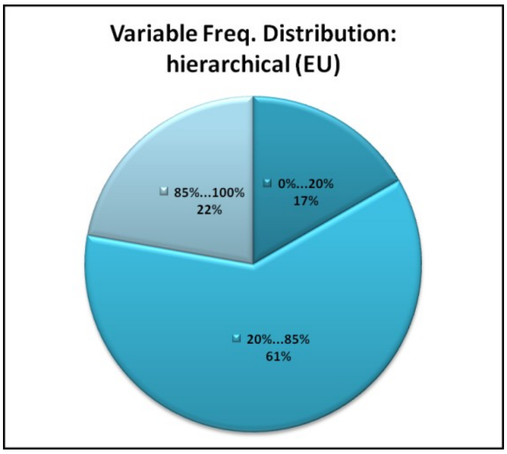
- Availability of results for direct and indirect impacts (***ADII***): 100%
- Availability of results for employment (***ARE***): 100%
- Use of Input-Output methods (***UIOM***): 100%
- Availability of definitions (***AD***): 89%

The above variables coincide with the high frequency variables identified in the global distribution, with *AD* following as the top middle-level variable globally.

In the zone of negligible frequency we encounter the following variables:

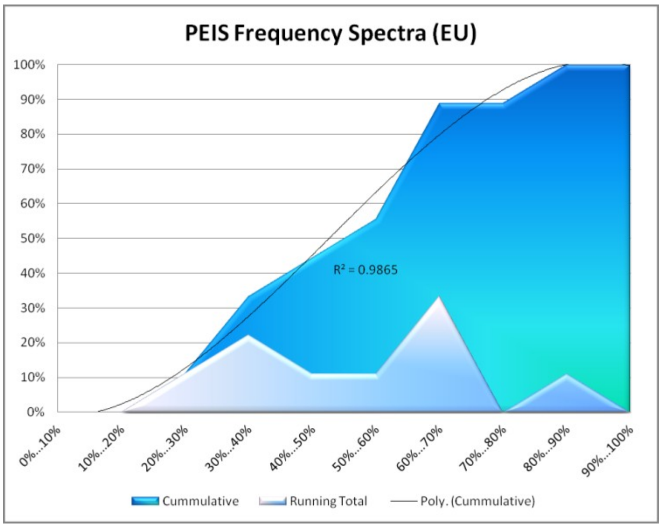
- Availability of results for development of business: (***ARDBE***): 11%
- Impact per commodity group (***ICG***): 11%
- Productivity (***P***): 11%

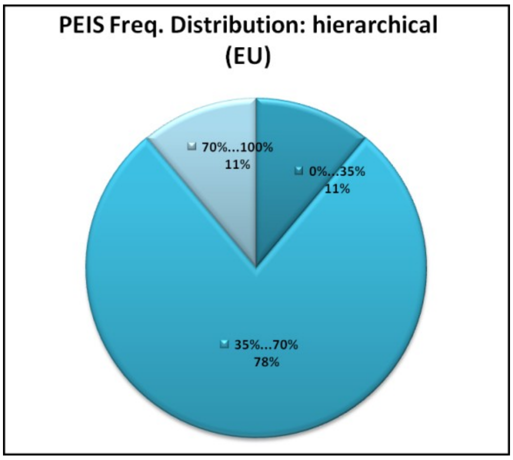
Two of the above variables, ***ARDBE*** and ***P***, also appear in the negligible frequency variables list, identified in the global distribution.



In the middle range of variables, spanning from 20.2% to 84.7% we find eleven variables that also appear in the global middle range, albeit in different order and bearing different frequencies of occurrence. Even in the middle range though, these variables exhibit comparatively high frequencies, which carry over as a tendency in the lower-frequency zone as well, where no variables are omitted at the lowest score of 11%, defining European PEIS as rather pluralistic in character and scope. The reasons for this most probably lie in the nature of the EU as a collection of member states with separate policies and practices, albeit the central EU Council provides for common guidelines.

This pluralistic character also reflects in the study adequacy scores, where the greatest majority of studies lie in the middle range, albeit with relative high scores. With a *statistical mean* of 52.47% and a narrow *standard deviation* of 17.15%, peaks start accumulating from the low frequency zone to the lower end of the middle zone, carrying over progressively to the final peak in the high zone. Consecutively, the cumulative distribution exhibits its first plateau as soon as it enters into the high zone.

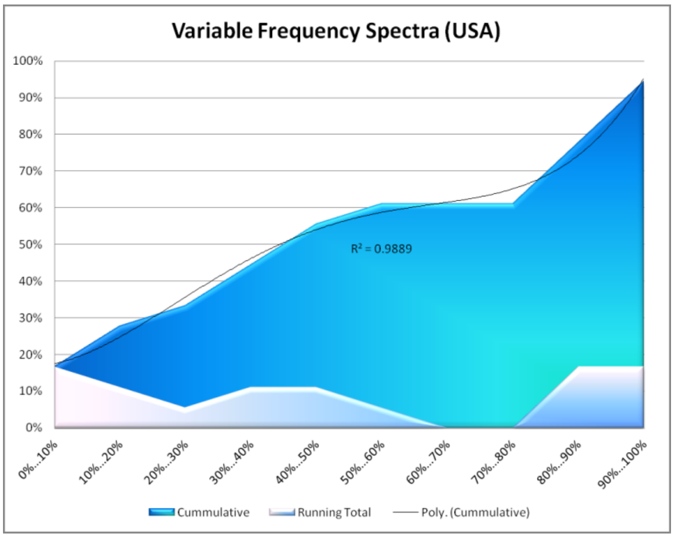




The statistical distribution of European study adequacy levels shows that the greatest majority of 9 studies fall in the zone of medium significance contributing to 78% of the sample size, while both the zone of negligible significance and the zone of high significance correspond to 1 case and 11% of the sample size respectively. The zone of high significance is covered by the Dutch study, while the zone of negligible significance is covered by a Spanish study.

3.2.2. USA

High frequencies of variable usage act as a general trend in the USA setting as well, exhibiting clear distinctions between variables that are predominant and variables that are omitted, in contrast to the European setting. Variable frequency statistics place the *mean* at 48.39% with a significant *standard deviation* of 35.22%, raising and lowering the mark for high and negligible occurrence respectively at 83.61% and 13.17%. This results in an accumulation of variables of very high frequency at the high zone and a set of variables always omitted at the low zone. The definitive allocation of high-end and low-end variables can also be discerned in the frequency histogram, where early pronounced peaks appear at the far left side of the spectrum, mirrored by late similarly pronounced peaks at the right side. Peaks appearing in the middle range form a low plateau. The cumulative distribution rises from an early stage, forming a corresponding plateau towards the right end of the middle zone: *kurtosis* -144.25% and *skewness* 5.84% and a spline fit with $R^2 \cong 99\%$.



In the zone of high frequency we encounter the following variables:

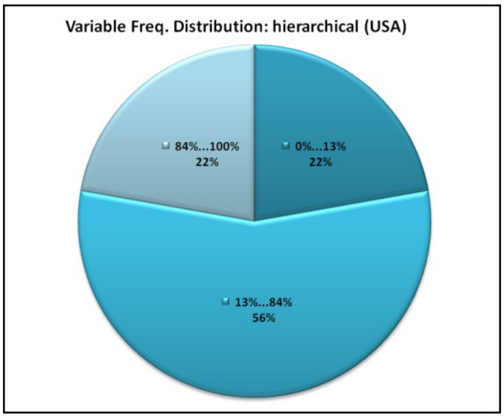
- Availability of results for employment (***ARE***): 100%
- Availability of results for direct and indirect impacts (***ADII***): 94%
- Use of Input-Output methods (***UIOM***): 94%
- Availability of results for taxes (***ART***): 88%

The above variable set coincides with the European and global patterns, apart from the variable ***ART***, which has replaced ***AD*** from the USA pattern.

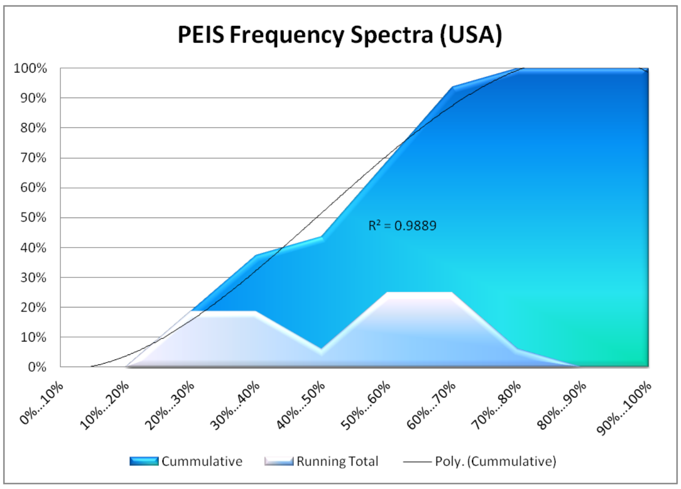
In the zone of negligible frequency we encounter the following variables:

- Use of surveys (***US***): 13%
- Availability of results for development of business: (***ARDBE***): 0%
- Availability of results for private investments (***ARPI***): 0%
- Productivity (***P***): 0%

These variables are also encountered in the global patterns – and the European ones with the exception of **ARPI** in place of ***ICG*** – albeit they are completely omitted in the USA setting, instead retaining very low scores at the others. The middle range is covered by 10 variables, also appearing in the corresponding global range with different frequencies.



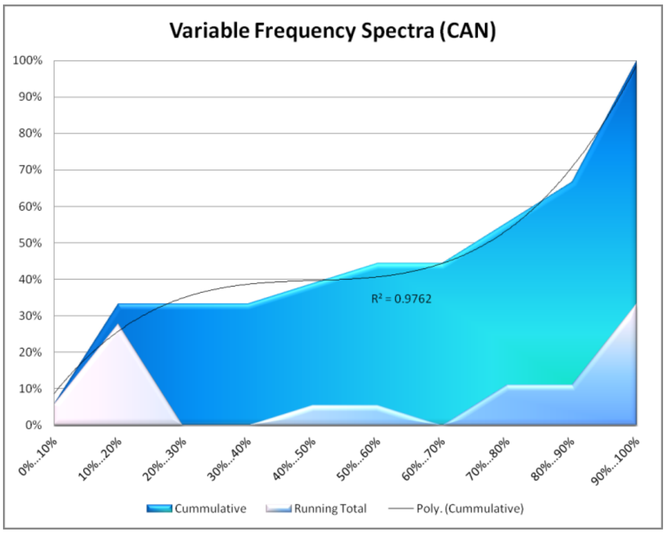
Concerning study adequacy levels, the greatest majority of studies lie in the middle range, albeit with relative high scores. With a *statistical mean* of 48.26% and a narrow *standard deviation* of 15.54%, peaks start accumulating at the low frequency zone, with the second peak accumulation appearing at the middle range. Consequently, the cumulative distribution rises steeply.

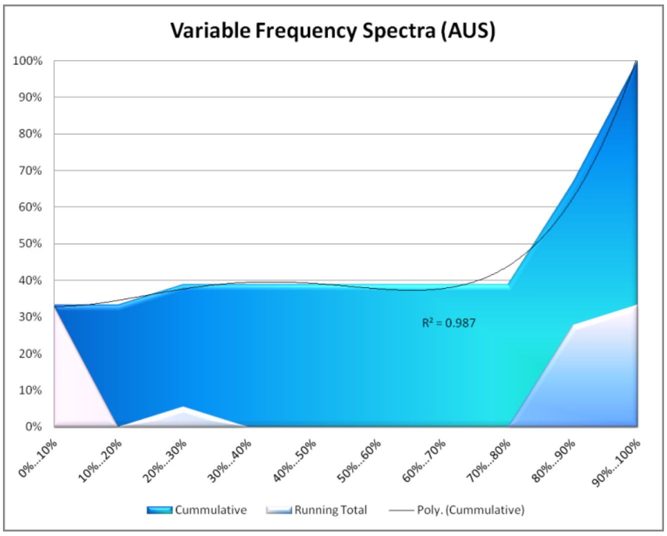


This adequacy level pattern reflects the USA tendency of relying predominantly on a trusted set of variables, while completely disregarding variables that seem problematic or irrelevant. If such variables had not been incorporated in the adequacy level calculations, the distribution would have shifted to the right, awarding the USA studies with higher scores.

3.2.3. Canada and Australia

The two regional sets exhibit extensive pattern similarities, mostly because they tend to follow up on the American philosophy of certain employment of high frequency variables and omitting undesirable ones; indeed, both regions appear as the logical evolution of the USA practices.





Variable frequency statistics place the *mean* at 60.22% for Canada and at 56.67% for Australia, with a significant *standard deviation* of 38.86% and 45.11% respectively; thus, high and negligible occurrence limits form around the marks of 99.08% and 21.36% for Canada, 101.78% and 11.56% for Australia. In the Canadian setting the high mark nears to and almost coincides with the upper outer limit of the distribution, thus allowing the allocation of variables of strictly certain employment – namely, having frequencies of 100% - into its high frequency zone; the Australian setting seems even more extreme, as the high mark exceeds the upper outer limit of the distribution, thus collapsing the zones of high and medium occurrence into one and allowing the allocation of omitted variables into the lower occurrence zone. Consequently, frequency histograms of both settings appear nearly identical, with peaks accumulating initially at the far left end of the spectrum, followed by peaks accumulating at the far right. Cumulative distributions appear smooth with gentle slopes rising up early into extended plateaus commencing from the low frequency zone to the thresholds of high frequency zone, where they rise again steeply to account for the peak accumulation of high frequencies: *kurtosis* -170.11%, *skewness* -34.16% and a spline fit with $R^2 \cong 98\%$ for the Canadian cumulative distribution; *kurtosis* -186.59%, *skewness* -42.47% and a spline fit with $R^2 \cong 99\%$ for the Australian one.

In the Canadian zone of high frequency we encounter the following variables:

- Availability of results for direct and indirect impacts (***ADII***): 100%
- Availability of results for employment (***ARE***): 100%
- Availability of results for employment in person-years or full-time equivalent (***AREPFE***): 100%
- Availability of results for induced impacts (***AII***): 100%
- Availability of results for value added (***ARVA***): 100%
- Detailed information on sectoral boundaries (***DISB***): 100%
- Availability of definitions (***AD***): 86%
- Use of Input-Output methods (***UIOM***): 86%

The Australian zone of high frequency is even more densely populated:

- Availability of definitions (***AD***): 100%
- Availability of results for direct and indirect impacts (***ADII***): 100%
- Availability of results for employment (***ARE***): 100%
- Availability of results for employment in person-years or full-time equivalent (***AREPFE***): 100%
- Availability of results for value added (***ARVA***): 100%
- Use of Input-Output methods (***UIOM***): 100%
- Availability of results for induced impacts (***AII***): 80%
- Detailed information on sectoral boundaries (***DISB***): 80%
- Distinction of traffic mix (***DTM***): 80%
- Impact per commodity group (***ICG***): 80%
- Impact per industry (***II***): 80%
- Availability of results for taxes (***ART***): 20%

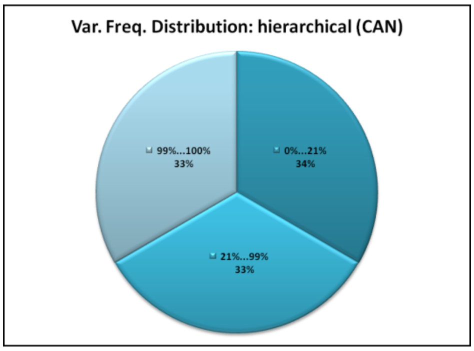
In the Canadian zone of negligible frequency we encounter the following variables:

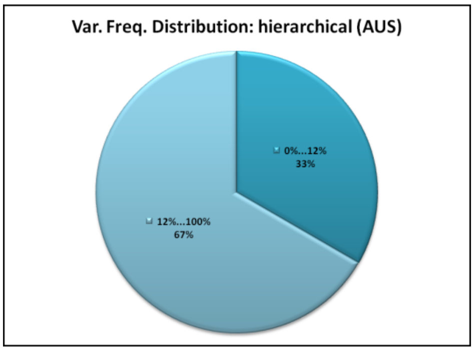
- Availability of results for private investments (***ARPI***): 14%
- Impact per industry (***II***): 14%
- Inclusion of hinterland (***IH***): 14%
- Productivity (***P***): 14%
- Use of surveys (***US***): 14%
- Availability of results for development of business: (***ARDBE***): 0%

The Australian zone of negligible frequency is formulated in a nearly identical manner, with the variable ***GET*** in place of ***II***:

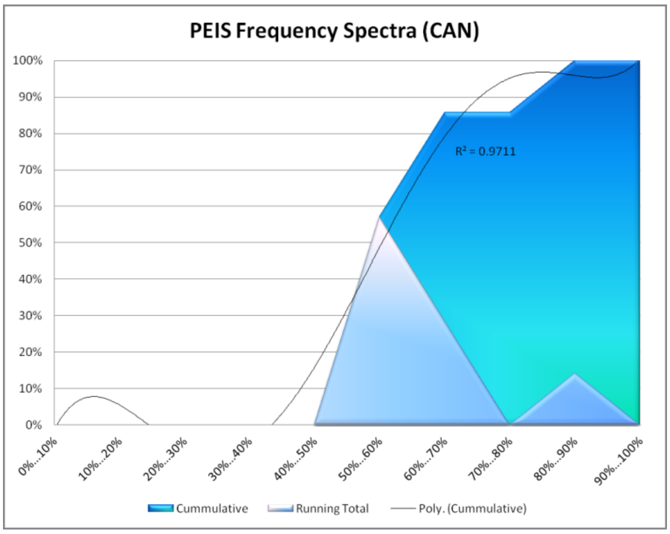
- Availability of results for development of business: (***ARDBE***): 0%
- Availability of results for private investments (***ARPI***): 0%
- Growth expressed in TEUs or other units (***GET***): 0%
- Inclusion of hinterland (***IH***): 0%
- Productivity (***P***): 0%
- Use of surveys (***US***): 0%

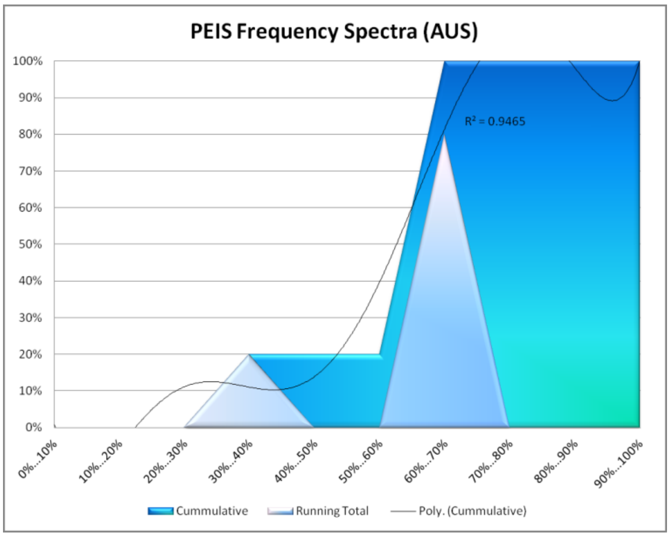
The extreme statistics describing the Australian setting are responsible for the observed inequalities in the high frequency set populations of the two groups, with the Canadian one consisting of 6 variables and the Australian set doubling its size at 12 variables, resulting from the collapse of its medium and high zones into one.



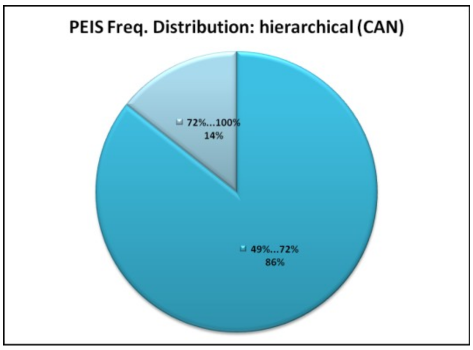


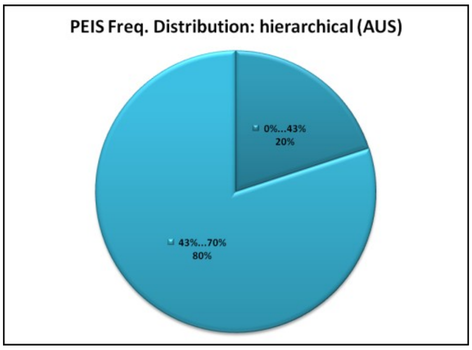
Concerning study adequacy levels, the greatest majority of studies in both settings lie in the middle range, albeit with relative high scores, mirroring the distribution observed in the USA. With a *statistical mean* of 60.32% and a narrow *standard deviation* of 11.75% for Canada and respective values of 56.67% and 13.26% for Australia, peaks accumulate predominantly in the middle zone of significance, reflecting the shift to the left of adequacy scores observed in the USA. Consequently, both the cumulative distributions rise steeply, formulating into wide blocks at the right of the spectrum: *kurtosis* 208.00% and *skewness* 144.21% for the Canadian cumulative distribution, *kurtosis* 442.29% and *skewness* -204.28% for the Australian one, with "difficult" spline fits with $R^2 \cong 97\%$ and $R^2 \cong 95\%$ respectively.





The subtle difference between the two distributions originate from the low peaks appearing next to the core high peak, located in the high frequency zone at the Canadian distribution and in the low zone at the Australian one. This reflects the rarely used ***ART*** variable – frequency of employment 20% - in the Australian setting, though it is incorporated into the high frequency variable zone due to the large number of omitted variables; the Canadian distribution seems more uniform in that aspect.





Consequently, the populations for each zone of significance appear polarized between the zone of medium significance encompassing the majority of PEIS, 6 studies attributed to 86% of the sample size for Canada and 4 studies attributed to 80% for Australia, and single-member sets for the high and low significance zones respectively. This effect may also be attributed to the narrow population size of the regional sampled PEIS, but judging from the observed polarization between dominant high-frequency variables versus negligible and omitted low-frequency ones, this effect may as well characterize the general dynamics of larger sample sizes.

3.3. Optimization towards a Common Framework

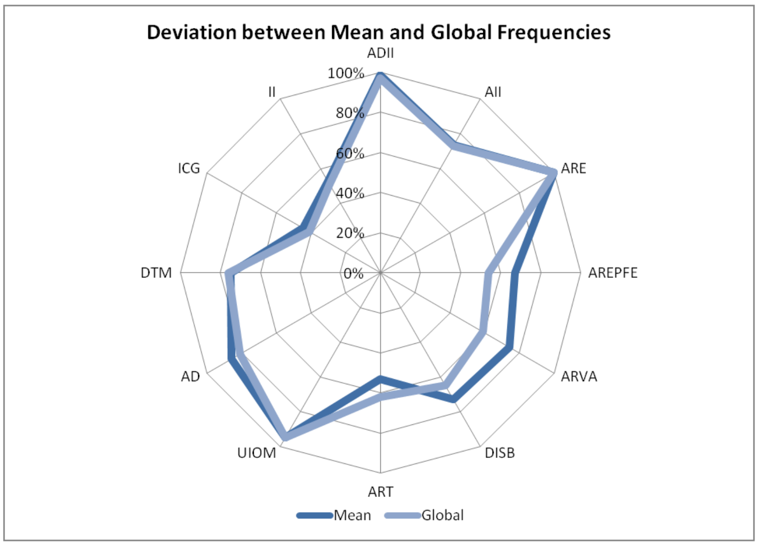
After reviewing global trends in descriptive variable frequency of employment, pertaining to conceptual preferences of socioeconomic impacts assessment, and reviewing studies according to their descriptive adequacy, pertaining to the range of concepts employed towards covering the demands of analytical objectives, we have compared the observed global statistical patterns with ones extracted from sets of case studies belonging to regional groups. Similarities and marked differences have been observed through cross-examination of such regional groups of PEIS and their individual effects on the global trends have been clarified; where the global distributions of variable employment and adequacy levels have appeared uniformly dispersed, rendering the formation of concrete arguments unsafe and even prohibitive, as the case might be, regional statistical patterns have provided much clearer insights instead. Therefore, we will attempt to reroute our approach from the partial to the whole, in order to reconstruct precise arguments concerning the determination of a common framework regarding PEIS.

Recapitulating the tendencies observed in the regional setting of Canada and Australia, the statistical distribution of variable employment dictates pronounced polarization between sets of variables of high frequency of occurrence, usually nearing the 100% upper limit, and sets that have been decisively neglected or omitted, nearing or coinciding with the lower 0% limit of the distribution. Variable sets in the zones of medium significance have exhibited either comparatively high levels of frequency, as the 80% set of variables in the case of Canadian PEIS, or they are altogether nonexistent as an empty set in the case of Australia. This pattern provides a definitive benefit by clearly identifying variable sets that may act as guidelines, facilitating their adoption by future studies that seek to emulate paradigms set out by earlier ones; thus, framework cohesiveness can be assured. This pattern may also be outlined in the regional distribution of the USA, albeit less pronounced as there is a wider dispersal of frequencies over the full spectrum. The polarized pattern is less identifiable in the European set, where there are variable sets of very high frequency, but the effect blends out due to the prominence of middle-range variables instead. However, the European patterns may be justified due to the collective character of the region, being comprised of a number of member-states with less coordinated policies or guidelines.

Drawing from the above, we will attempt to define sets of core variables that may be used as proposed guidelines in a global setting, thus promoting spatial and temporal continuity for future PEIS in a common framework. Thus, we need to compare the Canadian high-frequency variable sets, used in our case as base scenario, with the corresponding variable frequencies from other regions, in order to determine working sets of satisfactory descriptive effectiveness.

	variable	*CAN*	*USA* *frequency*	*AUS*	*EU*	*Mean*	*Global*
CAN	ADII	**100%**	94%	100%	100%	99%	97%
	AII	**100%**	81%	80%	33%	74%	73%
	ARE	**100%**	100%	100%	100%	100%	100%
	AREPFE	**100%**	25%	100%	44%	67%	54%
	ARVA	**100%**	19%	100%	78%	74%	59%
	DISB	**100%**	44%	80%	67%	73%	65%
USA	ART	71%	**88%**	20%	33%	53%	62%
	UIOM	86%	**94%**	100%	100%	95%	95%
AUS	AD	86%	69%	**100%**	89%	86%	81%
	DTM	71%	81%	**80%**	67%	75%	76%
	ICG	43%	44%	**80%**	11%	45%	41%
	II	14%	50%	**80%**	67%	53%	51%
	Mean	81%	66%	85%	66%	74%	71%
	St. Deviation	28%	28%	23%	30%	18%	19%

In the above table, the calculated mean may differ from the global average, as we have assumed an equal weighting factor over all regional sets, thus eliminating the built-in deviation of the global value that corresponds to unequal sample sizes of regional sets. Out of the overall average scores, we select a set of variables within the mean frequency range defined by $[\overline{x'} - s_{x'}, 100\%] = [74\% - 18\%, 100\%] = [56\%, 100\%]$ to comprise the core variable set, with the rest serving as auxiliary variables.



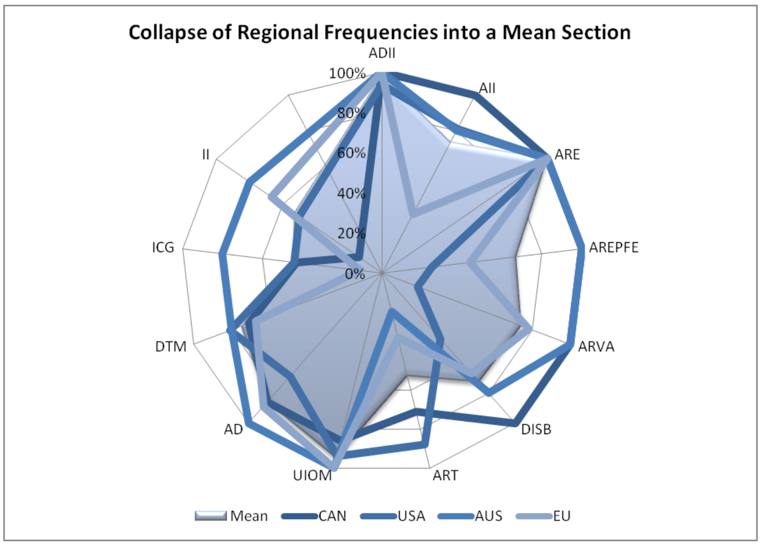
Such auxiliary variables are the following:

- Availability of results for taxes: ***ART***
- Impact per commodity group: ***ICG***
- Impact per industry: ***II***

Thus, the core variable set is comprised of the following variables:

- Availability of definitions: ***AD***
- Availability of direct and indirect impacts: ***ADII***
- Availability of induced impacts: ***AII***
- Availability of results for employment: ***ARE***
- Availability of results for employment in person-years or full-time eq.: ***AREPFE***
- Availability of results for value added: ***ARVA***
- Detailed information on sectoral boundaries: ***DISB***
- Distinction of traffic mix: ***DTM***
- Use of Input-Output methods: ***UIOM***

The above variable set comprises and encapsulates all the necessary information about the totality of port impacts, employment and economic activity, as well as clarifications on methodology procedures. Auxiliary variable sets may act as sources of additional information, even with the reinforcement by variables not in the list, but the core variables of PEIS need to be always present, upholding continuity and robustness.



We also consider the proposal of limiting the use of variables that have exhibited extremely low frequencies in our analysis, specifically the following:

- Availability of results for development of business: ***ARDBE***
- Availability of results for private investments: ***ARPI***
- Productivity: ***P***

This measure can always be uplifted when there is a specific need that calls for their specialized application. As a final note, we stress the fact that in our investigation there were no PEIS originating either from the regions of Africa or Asia; we recommend further investigation in order to address this issue.

Acknowledgements

The author would like to thank Dr. M. Nijdam for supervising this thesis and Mr. T. Tiliakos, *MRes BEng Dip* for his insightful remarks on statistical analysis methodology.

References

1. Chang, S., "*In Defense of Port Economic Impact Studies*", Transportation Journal, Spring, 1978
2. DeSalvo, J., "*Measuring the Direct Impacts of a Port*", Transportation Journal, pp.33-42, Summer, 1994
3. Garcia, L., "*Ports Economic Impact: Literature review and alternative proposal*", Journal of Maritime Research, Vol.1, No.3, pp.85-104, 2004
4. Hall, P. V., "*"We'd Have to Sink the Ships": Impact Studies and the 2002 West Coast Port Lockout*", Economic Development Quarterly, Vol.18 No.4, pp.354-367, November 2004
5. Hudson, I., "*The Use and Misuse of Economic Impact Analysis: The Case of Professional Sports*", Journal of Sports & Social Issues, Vol.25, No.1, pp.20-39, February 2001
6. Leontief, W., "*Input-output economics*", 2nd ed. Oxford University Press, New York, 1986
7. Marco et al. "*On the Economic Impact of Ports: Local vs. National Costs and Benefits*" Paper presented at the 9th World Conference on Transport Research, Seoul, Korea, July 2001
8. Notteboom, T., "*The facilitating role of port authorities in the logistics chain*", ESPO/EFIP International Workshop on "*Ports-facilitators in the logistics chain*", Strasbourg, March 2008,
9. Notteboom,T. & Rodrigue, J., "*Port regionalization: towards a new phase in port development*", Maritime Policy and Management, vol.32, no.3, pp.297-313, 2005
10. P. Cotto-Millan *et al.*(eds.), "*Essays on Port Economics, Contributions to Economics*", Springer-Verlag Berlin Heidelberg, 2010
11. Suykens, F., "*The City and Its Port-an Economic Appraisal*", Geoforum, Vol. 20, No.4, pp. 437-445, 1989
12. Waters, R. C., "*Port Economic Impact Studies: Practice and Assessment*", Transportation Journal, pp 14-18, Spring, 1977
13. Yochum, Agarwal, "*Economic Impact of a Port on a Regional Economy: Note*", Growth and Change, pp.74-87, Summer 1987

List of Port Economic Impact Studies

1. Atkins, "*Economic Impact of the port of Southampton*", Marine South East, August 2011
2. Blue Engineering Ltd., "*Environmental Impact Assessment: Nassau Harbour Port Improvement project, Nassau, Bahamas*", Cox & SHAL Consultants, Government of the Bahamas, November 20, 2008
3. BST Associates, "*Port of Astoria. 2009 Economic Impact Study: Final draft Report*", Port of Astoria, April 15, 2011
4. BST Associates, "*Port of Olympia:2004 Economic Impact Study Final Report*", port of Olympia, July 15, 2005
5. BST Associates, "*Port of Port Angeles: 2006 Economic Impact Study Final Report*", Port of Port Angeles, September 25, 2007
6. BST Associates, "*Trade Impact Study Final Report*", Alameda Corridor Transportation Authority, March, 2007
7. Carstensen F. V. *et al.*, "*The Economic Impact of Connecticut's Deepwater Ports: An IMPLAN and REMI analysis*", Connecticut Center For Economic Analysis, University of Connecticut, May 23, 2001
8. Castillo J. *et al.*, "*Dynamising Economic Impact Studies: the case of the port of Seville*"
9. Charleston Southern University, "*The Economic Impact of the South Carolina Ports Authority*", Center for Economic Forecasting
10. Chris Lowe Planning & Management Group, "*Port of Halifax Economic Impact report*", Halifax Port Authority, February, 2009
11. Coppens F. *et al.*, "*Economic Impact of port activity: a disaggregate analysis. The case of Antwerp*", National Bank of Belgium, Working paper document, vol.110, February 2007
12. Coto-Millan *et al.*, "*Port economic impact: methodologies and application to the Port of Santander*", maritime Transport Applied Economics, Civitas, Chapter 8, pp.131-154, 1999b

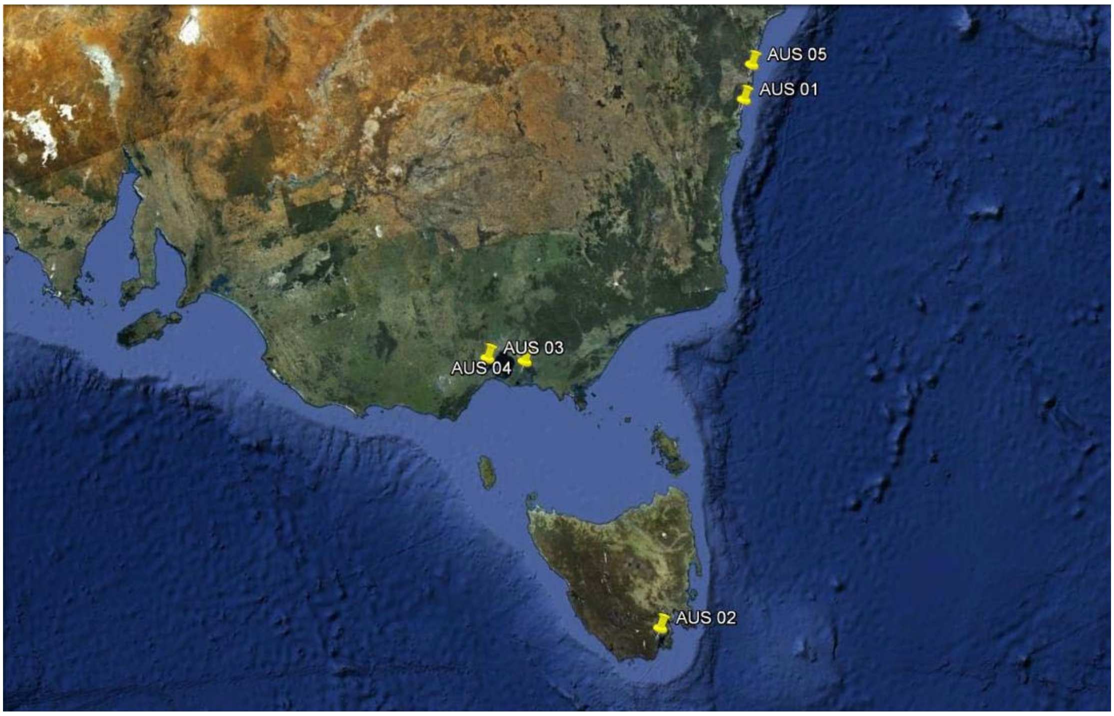
13. Danielis R. *et al.* “*An input-output based methodology to estimate the economic role of a port: the case of the port system of the Friuli Venezia Giulia Region, Italy*”, Working papers SIET, ISSN 1973-3208, 2012
14. Doldán-García X.R. *et al.*, “*Estimating the economic impacts of maritime port development: the case of A Coruna, Spain*”, Ocean& Coastal Management, Vol.54, pp.668-677, 2011
15. EconSearch Pty Ltd, “*Economic Impact of the Port of Geelong*”, Victorian Regional Channels Authority, Department of Transport, April 28, 2010
16. EconSearch Pty Ltd, “*Port of Port Kembla: Economic Impact Study*”, Port Kembla Port Authority, March 7, 2007
17. EconSearch Pty Ltd, “*The Economic Impact of the Port of Esperance, 1999/00*”, Esperance Port Authority, December 2001
18. Fogarty R. *et al.*, “*Benefit Cost Analysis: an economic and environmental impact study for the port of Providence, a study of the long term benefits on income generation, job creation and the environment*”, Port of Providence, August 20, 2010
19. Gardner Pinfold, “*Economic Impact Study of Independent Marine Ports in Atlantic Canada*”, Independent Marine Ports Association of Atlantic Canada, December 2010
20. Guerrero F. *et al.*, “*Economic impact of Western Mediterranean leisure ports*”
21. InterVISTAS CONSULTING INC. , “*Halifax Gateway Council Economic Impact Study Final Report*”, November 2005
22. InterVISTAS CONSULTING INC., “*2008 Port Metro Vancouver Economic Impact Study Final Report*”, Port Metro Vancouver, January 12, 2009
23. InterVISTAS CONSULTING INC., “*Port of Prince Rupert Economic Impact Study version 2*”, Prince Rupert Port Authority, July 19, 2010
24. InterVISTAS CONSULTING INC., “*Port of Prince Rupert: Economic Impact Study Update: Final Report*”, Prince Rupert Port Authority, March 26, 2012
25. Kinsey J. , “*The Economic Impact of the Port of Liverpool on the Economy of Merseyside – Using a Multiplier approach*”, Geoforum, Vol.12, No.4, pp.331-347, 1981
26. Kitchen H. *et al.*, “*Economic and Financial Impact Analysis of Cameco and Zircatec in Port Hope and Northumberland County*”, Cameco and Zircatec Port Hope, July 21, 2006
27. Kleinsteuber S., “*The Jade-Weser Port: An Economic Impact Assessment*”, European Business School, March 7, 2002
28. Loren C. Scott & Associates, “*The Economic Impacts of Port Fourchon on the National and Houma MSA Economies*”, Louisiana Economic Development, South Louisiana Economic Council, April 2008
29. Martin Associates, “*Economic Impact Study of the Great Lakes St. Lawrence Seaway System*”, The U.S. Saint Lawrence Seaway Development Corporation, August 1, 2001
30. Martin Associates, “*Economic Impacts of the Port of Los Angeles*”, Port of Los Angeles, August 20, 2007
31. Martin Associates, “*The Economic Impact of the Port of Lake Charles*”, Lake Charles Harbor and Terminal District, September 29, 2007
32. Martin Associates, “*The economic impact of the port of Tacoma*”, Port of Tacoma, May 24, 2005
33. Martin Associates, “*The economic impacts of the port of Baltimore*”, Maryland Port Administration, January 28, 2008
34. Martin Associates, “*The local and regional economic impacts of the port of Houston,2011*”, Port of Houston Authority , May 22, 2012
35. Mathys C., “*Economic importance of the Belgian Ports: Flemish maritime ports, Liege port complex and the port of Brussels - Report 2009*”, National Bank of Belgium, Working paper document, No215, June 2011
36. Meyrick and Associates & EconSearch Pty Ltd, “*Economic Impact Study: Port of Hastings*”, Port of Hastings Corporation, Victorian Department of Infrastructure, April 2007
37. Nijdam M. *et al.*, “*HavenMonitor 2010*”, Erasmus Universiteit Rotterdam (RHV), May 2012
38. Pearson R. *et al.*, “*The fiscal year 2006 Virginia economic and fiscal impacts of Virginia Port Authority operations*”, The Virginia Port Authority, January 8, 2008
39. Siegesmund P. *et al.*, “*Guide to the Economic Value of Texas Ports*”, Center for Transportation Research, Texas Department of Transportation, Research and Technology Implementation Office, February 2008

40. Sinclair J. T., "*The application of economic impact analysis: a case study of Fraser Port*", The University of British Columbia, 1991
41. Sinclair Knight Merz, "*Port Hedland Cumulative Impact Study*", Department of Industry and Resources, Government of Western Australia, December 19, 2007
42. URS, "*Port Botany Expansion Environmental Impact Statement - Volume 2*", Sydney Ports

Appendix I: List of PEIS original publication titles and maps

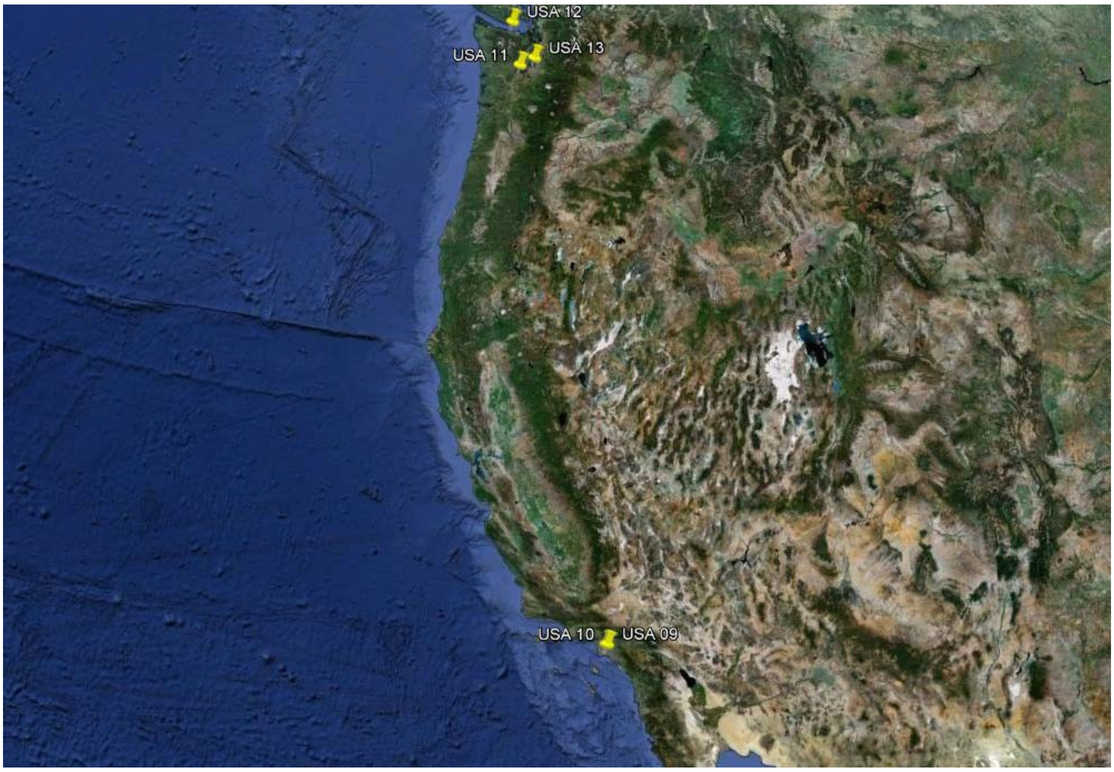
Code	***PEIS Title***
USA01	*Guide to the economic value of Texas ports*
USA02	*The economic impact of port Fourchon on the national and Houma MSA Economies*
USA03	*Port of Astoria - Economic Impact Study*
USA04	*The economic impacts of the port of Baltimore*
USA05	*The economic impact of Connecticut's deep water ports*
USA06	*Economic impact study of the Great Lakes St.Lawrence Seaway system*
USA07	*The local and regional economic impacts of the port of Houston*
USA08	*The economic impact of the port of Lake Charles*
USA09	*The port of Long Beach*
USA10	*Economic impacts of the port of Los Angeles*
USA11	*Port of Olympia*
USA12	*Port of Port Angeles*
USA13	*An economic and environmental impact study for the port of Providence*
USA14	*The economic impact of the port of Tacoma*
USA15	*The economic impacts of the port of Virginia*
USA16	*The economic impact of the South Carolina State Ports Authority*
CAN01	*Halifax Gateway*
CAN02	*Port Metro Vancouver*
CAN03	*The application of economic impact analysis: a case study of Fraser Port*
CAN04	*The application of economic impact analysis: a case study of Fraser Port*
CAN05	*Port of Prince Rupert*
CAN06	*Port of Prince Rupert*
CAN07	*Economic impact study of independent marine ports in Atlantic Canada*
AUS01	*Port of Port Kembla: Economic Impact Study*
AUS02	*The economic Impact of the port of Esperance*
AUS03	*Economic impact of the port of Geelong*
AUS04	*Economic Impact study: port of Hastings*
AUS05	*Ports of Sydney*
ITA01	*An input-output based methodology to estimate the economic role of a port: the case of the port system of the Friuli Venezia Giulia Region, Italy*
ITA02	*Economic impact of western Mediterranean leisure ports*
SPA01	*Dynamising Economic Impact Studies: the case of the port of Seville*
SPA02	*The economic impact of ports: its importance for the region and also the hinterland: the port of Santander*
UK01	*The economic impact of the port of Liverpool on the economy of Merseyside - using a multiplier approach*
UK02	*Economic impact of the port of Southampton*
BEL01	*Economic impact of port activity: a disaggregate analysis: the case of Antwerp*
BEL02	*Economic importance of the Belgian Ports*
NED01	*Havenmonitor HM2010*

Map of European ports corresponding to analyzed PEIS



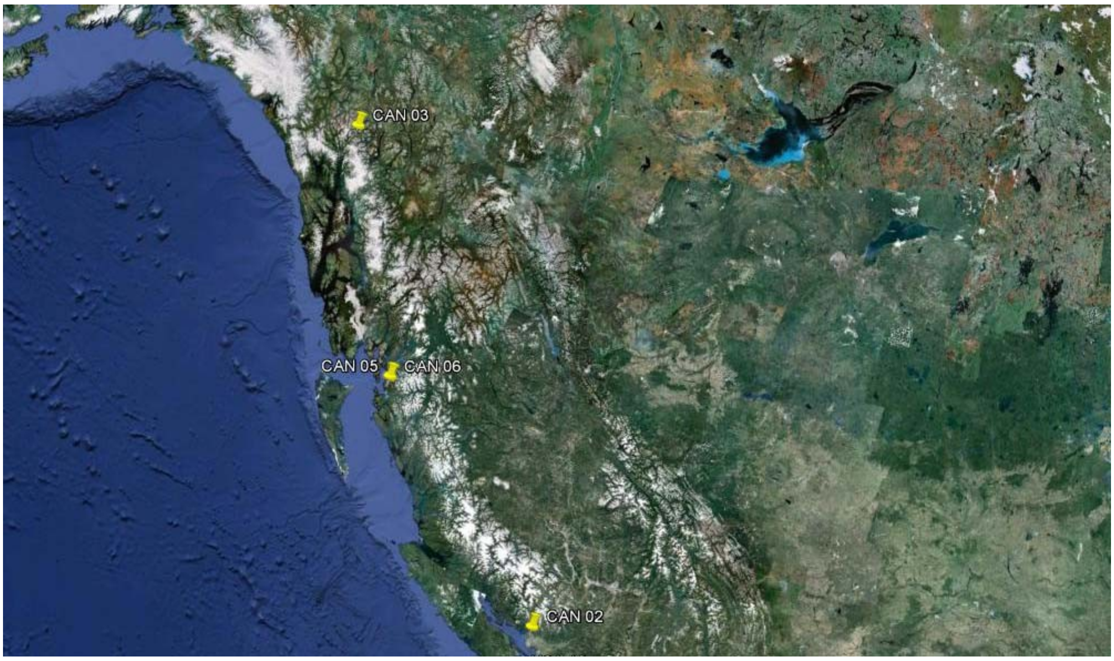
Map of Australian ports corresponding to analyzed PEIS

Map of Eastern USA ports corresponding to analyzed PEIS



Map of Western USA ports corresponding to analyzed PEIS

Map of Eastern Canadian ports corresponding to analyzed PEIS



Map of Western Canadian ports corresponding to analyzed PEIS

Appendix II: List of PEIS codification and reference years

PEIS	*Port of referance*	*Year*
AUS01	Port of Port Kembla	2006
AUS02	Port of Esperance	2000
AUS03	Port of Geelong	2009
AUS04	Port of Hastings	2006
AUS05	Port of Sydney	2002
BEL01	Port of Antwerp	2000
BEL02	Ports of Belgium	2009
CAN01	Halifax Gateway	2004
CAN02	Port Metro Vancouver	2007
CAN03	Port of Fraser	1992
CAN04	Port of Fraser	2007
CAN05	Port of Prince Rupert	2011
CAN06	Port of Prince Rupert	2009
CAN07	Ports of Atlantic Canada	2010
ITA01	Ports of the Friuli Venezia Giulia Region	2007
ITA02	Western Mediterranean leisure ports	2004
NED01	Ports of Netherlands	2010
SPA01	Port of Seville	2003
SPA02	Port of Santander	2005
UK01	Port of Liverpool	1979
UK02	Port of Southampton	2011
USA01	Ports of Texas	2005
USA02	Port Fourchon	2006
USA03	Port of Astoria	2009
USA04	Port of Baltimore	2002
USA05	Ports of Connecticut	2000
USA06	Great Lakes St. Lawrence Seaway system	2000
USA07	Port of Houston	2011
USA08	Port of Lake Charles	2006
USA09	Port of Long Beach	2005
USA10	Port of Los Angeles	2006
USA11	Port of Olympia	2004
USA12	Port of Port Angeles	2006
USA13	Port of Providence	2011
USA14	Port of Tacoma	2004
USA15	Port of Virginia	2006
USA16	Ports of South Carolina State	2002

Appendix III: List of variables and definitions

Code	***Variable definition***
AD	Availability of definitions
ADII	Availability of direct and indirect impacts
AII	Availability of induced impacts
ARDBE	Availability of results for development of business establishments
ARE	Availability of results for employment
AREPFE	Availability of results for employment in person-years or full time equivalent
ARPI	Availability of results for private investments
ART	Availability of results for taxes
ARVA	Availability of results for value added
DISB	Detailed information on sectoral boundaries
DTM	Distinction of traffic mix
GET	Growth expressed in TEUs or other units
ICG	Impact per commodity group
IH	Inclusion of hinterland
II	Impact per industry
P	Productivity
UIOM	Use of input-output models
US	Use of surveys