



UNIVERSIDAD DE CANTABRIA



E.T.S. de Ingenieros de Caminos, Canales y Puertos

Tesis Doctoral

Modelización de demanda por transporte portuario de mercancías: Un enfoque de autómatas celulares

Autora

MABEL LEVA HENRÍQUEZ

Directores

ÁNGEL IBEAS PORTILLA

RODRIGO GARRIDO HIDALGO

Santander, febrero 2017

A Rodrigo, el complemento de mi alma.

RESUMEN

La modelización de la elección de puertos por parte de exportadores de carga, suele realizarse asumiendo que esta elección depende esencialmente de los costos de transporte entre el puerto y el origen o destino final dentro del hinterland. Sin embargo, este proceso de elección está dominado por una multiplicidad de variables y actores, que dificultan su modelización con técnicas estándar como los modelos de elección discreta. Una de las razones que explican la incapacidad de estos métodos para predecir el comportamiento futuro es la característica de sistema complejo que se observa en este proceso. En efecto, la elección de puerto es el resultado de la interacción de muchos actores a través de reglas relativamente simples pero que en conjunto evolucionan en forma difícil de predecir.

En esta investigación se utiliza una herramienta que ha demostrado ser de una enorme capacidad predictiva y descriptiva de sistemas complejos de distinta naturaleza: la teoría de autómatas celulares (AC). Estos modelos logran captar con facilidad interacciones espacio-temporales que en otros contextos de modelización son extremadamente difíciles de incorporar y calibrar.

En esta tesis se ha desarrollado un modelo de elección portuaria basado en un autómata celular probabilístico, que busca replicar la decisión de múltiples exportadores que interactúan entre sí y que ponderan una multiplicidad de aspectos y comportamientos espaciales que configuran una elección dinámica.

La especificación del autómata celular diseñado es aplicada a un caso práctico de exportación de automóviles en Brasil, utilizándose un algoritmo genético para su calibración.

Los resultados obtenidos muestran que el modelo logra replicar con un alto porcentaje de aciertos el comportamiento dinámico de la elección, obteniéndose que el parámetro espacial, asociado a la componente de vecindad del autómata, es significativo al igual que el correspondiente a las características y ubicación de los puertos analizados.

Así, el principal producto de esta tesis, es una herramienta de modelización que permite aproximarse, de manera simple, a un problema no resuelto: el efecto de la complejidad

en las interacciones espacio-temporales de los exportadores en la elección del puerto y en la definición del hinterland.

Esta herramienta, de rápida calibración y aplicación, es una base de modelización que puede ser de enorme valor para los tomadores de decisiones, ya que permite evaluar el impacto de diversas políticas (como mejoras en accesibilidad, planificación territorial, localización portuaria, etc.) y de acciones que se vinculen directamente a la operación portuaria.

Como producto de esta tesis se desarrollaron los siguientes artículos:

- I. Modelación de elección portuaria: Un enfoque de autómatas celulares probabilísticos. Actas XIX Congreso Panamericano de Ingeniería de Tránsito, Transporte y Logística, Ciudad de México, 2016.
- II. A probabilistic cellular automaton to forecast port choice decisions. *Transport*. En revisión.
- III. Port hinterland formation: A Probabilistic Cellular Automaton Approach. *Naval Research Logistics*. En revisión.

ABSTRACT

Modeling the port choice process from the shippers' perspective is usually done assuming that this choice depends (almost) exclusively on transportation costs between the port and the origin (or destination) within a certain hinterland. Nevertheless, several factors and stakeholders in a complicated set of interactions that makes difficult to find a standard model specification, such as discrete choice models. One of the reasons why standard transportation modeling approach fail to explain the future behavior of the system is that stakeholders exhibit characteristics of complex behavior. In fact, the rules governing the actions of the various actors involved in the port choice problem are relatively simple and yet the aggregate system's behavior is difficult to predict.

This research presents a new modeling and forecasting tool with a significant predictive power, based on the theory of cellular automaton (CA). The latter are mathematical structures capable of representing complex interactions of many agents in a relatively simple manner. In particular, CA models can represent spatio-temporal decisions, which are rather complicated to describe with other methodological tools. In those contexts the CA are easy to calibrate and implement.

In this thesis a probabilistic port choice model based on CA was developed. The model was aimed to replicate the decision processes made by multiple exporters that interact with each other and face similar options with regard to the choice of a seaport for exporting their shipments within a certain time span.

The specification of the CA model is applied to a practical case of automobile exports from the main ports in Brazil and the calibration of this model is done through a genetic algorithm.

The results of the model's application show that the probabilistic CA is able to replicate the dynamic behavior of the automobile exports through Brazilian seaports with a high percentage of success. The spatial component of the CA turned out to be o major relevance in the dynamic decision process along with the attributes and geographical location of ports.

Thus, the main contribution of this thesis is a modeling tool that allows tackling an unresolved problem: the effect of complex spatio-temporal interactions between exporters on the port choice and the hinterland definition.

This tool, of easy calibration and implementation, is a modelling base that can be of great value for decision makers, since it allows assessing the impact of various policies (e.g. accessibility improvements, land use planning, port location, etc.) and actions directed related with port activities.

The following articles summarize the result of this research:

- I. Modelación de elección portuaria: Un enfoque de autómatas celulares probabilísticos. Proceedings of the XIX Congreso Panamericano de Ingeniería de Tránsito, Transporte y Logística, Mexico City, 2016.
- II. A probabilistic cellular automaton to forecast port choice decisions. *Transport*. Under review.
- III. Port hinterland formation: A Probabilistic Cellular Automaton Approach. *Naval Research Logistics*. Under review.



AGRADECIMIENTOS

Quisiera agradecer en primer lugar a Ángel Ibeas por brindarme la oportunidad de desarrollar estas tesis y por apoyarme durante todo su desarrollo.

A Rodrigo Garrido por su creatividad, conocimiento, paciencia y las muchas horas dedicadas en esta tesis.

A Alejandro León por su interés en participar de esta investigación y su apoyo en el desarrollo y programación de los modelos.

A mis compañeros de trabajo del Ministerio de Transporte y Telecomunicaciones de Chile, quienes me aportaron su conocimiento y el sentido de realidad que implica trabajar para el país.

A mis hijos Tomás y Nicolás, por existir y porque los amo.

Y por supuesto, a mi esposo, mi norte y mi compañero de vida, por no soltar mi mano en este nuevo paso.

ÍNDICE DE CONTENIDO

RESUMEN	i
ABSTRACT	iii
AGRADECIMIENTOS.....	vii
ÍNDICE DE CONTENIDO	viii
ÍNDICE DE FIGURAS	x
ÍNDICE DE TABLAS	xi
1. INTRODUCCIÓN Y OBJETIVOS	2
1.1. Motivación	2
1.2. Objetivos	6
1.3. Metodología	7
1.4. Estructura de la Tesis	8
1.5. Aportaciones	9
2. REVISIÓN BIBLIOGRÁFICA.....	12
2.1. Agentes Tomadores de Decisión	12
2.2. Factores de Elección Portuaria.....	15
2.2.1. Factores de Elección desde el Punto de Vista de los Exportadores	15
2.2.2. Factores de Elección desde el Punto de Vista Naviero	21
2.2.3. Factores de Elección desde el Punto de Vista de más de un Actor.....	25
2.3. Modelos de Elección Portuaria	28
2.2.4. Modelo Probit Multinomial Multidimensional Espacio-Tiempo.....	33
2.4. Conclusiones.....	45
3. DISEÑO DE AUTÓMATA CELULAR PARA ELECCIÓN PORTUARIA.....	56
3.1. Modelos de Autómatas Celulares	56
3.2. Estructura y Características del Autómata Celular Probabilístico de Elección Portuaria (ACPEP)	59
3.2.1. Vecindad.....	61
3.2.2. Función de Actualización.....	62
3.2.3. Algoritmo de Calibración.....	64
3.3. Conclusiones.....	67
4. ESTRUCTURA EQUIVALENTE MODELO PROBIT MULTINOMIAL.....	71
4.1. Modelo Probit de Elección Portuaria con Correlación entre Exportadores.....	72
4.2. Conclusiones.....	75
5. APLICACIÓN PRÁCTICA ACEP	77
5.1. Análisis de Datos	77

5.2.	Calibración ACPEP	83
5.2.1	Interfaz Gráfica del Algoritmo	83
5.2.2	Resultados Calibración	84
5.3.	Ejecución del Modelo	89
5.3.1	Interfaz Gráfica	89
5.3.2	Rendimiento del Modelo	90
5.4.	Conclusiones.....	92
6.	CAPACIDAD PREDICTIVA DEL MODELO.....	95
6.1.	Redistribución de ZGC por Cambio en el Nivel de Servicio Portuario.....	95
6.2.	Redistribución de ZGC por Desactivación de un Puerto	98
6.3.	Conclusiones.....	109
7.	CONCLUSIONES FINALES Y LÍNEAS DE INVESTIGACIÓN FUTURA.....	111
7.1.	Conclusiones.....	111
7.2.	Líneas Futuras de Investigación	117
	REFERENCIAS.....	119
	ANEXOS	126
	Anexo A: Modelización de elección portuaria: Un enfoque de autómatas celulares probabilísticos.....	126
	Anexo B: A probabilistic cellular automaton to forecast port choice decisions	144
	Anexo C: Port hinterland formation: A Probabilistic Cellular Automaton Approach.....	163
	Anexo D: Port of destination and carrier selection for fruit exports: a multi- dimensional space-time multinomial probit model.....	192
	Anexo E: Código ACPEP	208

ÍNDICE DE FIGURAS

Figura 1: Tomadores de decisión claves en el proceso de elección portuaria.....	13
Figura 2: Esquema conceptual de un AC en dos dimensiones.....	57
Figura 3: Ejemplo de evolución AC según el juego de la vida	58
Figura 4: Esquema conceptual de las zonas y puertos en el ACP	60
Figura 5: Esquema de vecindad definida en el ACPEP	62
Figura 6: Ejemplo cruce entre individuo 1 y 2 en la generación t y el individuo descendiente en la generación t+1.....	66
Figura 7: Esquema del operador de mutación que afecta al individuo 1 en la generación t y el individuo mutante en la generación t+1	67
Figura 8: Puertos y ZGC.	79
Figura 9: Porcentaje de ZGC activas en cada período observado.....	80
Figura 10: Distribución histórica de ZGC según su elección portuaria.....	80
Figura 11: ZGC según puerto más cercano.....	81
Figura 12: Distribución histórica promedio de la elección de cada ZGC.....	82
Figura 13: Interfaz del programa desarrollado para calibrar el algoritmo genético.....	83
Figura 14: Interfaz del programa una vez finalizado el algoritmo genético.	84
Figura 15: Total de vecinos por cada ZGC	87
Figura 16: Probabilidad de activación de ZGC.....	89
Figura 17: Interfaz inicio para simulación del ACPEP	89
Figura 18: Visualización de la reasignación de ZGC con el puerto de Santos desactivado	90
Figura 19: Predicción ZGC activas a 20 años	91
Figura 20: Distribución de ZGC según su elección portuaria con $\gamma_1 = 1250, 1300, 1350$ y 1400	97
Figura 21: Distribución de ZGC según su elección portuaria sin la existencia del Puerto 1.....	99
Figura 22: Comparación de la distribución histórica promedio de elección portuaria de cada ZGC con la distribución simulada de elección sin la existencia del Puerto 1.	100
Figura 23: Distribución de ZGC según su elección portuaria sin la existencia del Puerto 2.....	101
Figura 24: Comparación de la distribución histórica promedio de elección portuaria de cada ZGC con la distribución simulada de elección sin la existencia del Puerto 2.	103
Figura 25: Distribución de ZGC según su elección portuaria sin la existencia del Puerto 3.....	104
Figura 26: Comparación de la distribución histórica promedio de elección portuaria de cada ZGC con la distribución simulada de elección sin la existencia del Puerto 3.	105
Figura 27: Distribución de ZGC según su elección portuaria sin la existencia del Puerto 4.....	106
Figura 28: Comparación de la distribución histórica promedio de elección portuaria de cada ZGC con la distribución simulada de elección sin la existencia del Puerto 4.	108

ÍNDICE DE TABLAS

Tabla 1: Factores de elección portuaria desde el punto de vista de los exportadores de carga o freight forwarders.	21
Tabla 2: Factores de elección portuaria desde el punto de vista naviero	25
Tabla 3: Factores de elección portuaria considerando más de un actor	28
Tabla 4: Actores, factores de elección portuaria y metodologías de modelización	54
Tabla 5: Parámetros obtenidos para el ACPEP	85
Tabla 6: Bondad de ajuste de modelos de elección discreta aplicados a elección portuaria.	86
Tabla 7: Comparación de la distribución simulada de ZGC por puerto respecto a la distribución real.	91
Tabla 8: Variación en la distribución de ZGC según su elección portuaria sin la existencia del Puerto 1.	99
Tabla 9: Variación en la distribución de ZGC según su elección portuaria sin la existencia del Puerto 2.	101
Tabla 10: Variación en la distribución de ZGC según su elección portuaria sin la existencia del Puerto 3.	104
Tabla 11: Variación en la distribución de ZGC según su elección portuaria sin la existencia del Puerto 4.	107



Capítulo 1 **INTRODUCCIÓN**

1. INTRODUCCIÓN Y OBJETIVOS

1.1. Motivación

Más del 85% de las mercancías movilizadas globalmente utilizan el transporte marítimo como modo principal. Esta característica parece ser una tendencia de largo plazo al observar los tamaños de naves que se utilizan para el transporte marítimo. En efecto, en la actualidad los tamaños de buques portacontenedores y buques tanque han llegado a tener tamaños muy superiores a los estándares anteriores.

Hoy los buques Postpanamax y Postpanamax plus, entre otros, están siendo utilizados con frecuencia en rutas que habitualmente preferían embarcaciones más pequeñas. Estas tendencias han ido de la mano con mejoramientos en la capacidad de transferencia de los puertos y en la profesionalización de todos los servicios requeridos por estas embarcaciones. Así, por ejemplo, la tecnología de transferencia portuaria ha experimentado cambios sustantivos tanto en las dimensiones de la maquinaria como en los sistemas logísticos, para permitir un trasbordo eficiente y seguro de las mercancías tanto desde la nave hacia el puerto, como de éste hacia el sistema que consume estas mercancías.



Dentro de los efectos indirectos que trae el aumento de transferencia de carga están las demoras en puerto, la congestión intra y extraportuaria de vehículos pesados, la contaminación ambiental y el ruido, entre otros. Por esto, la planificación estratégica de un sistema portuario va más allá de las actividades que tienen lugar en el puerto y abarca toda una gama de interacciones entre nodos de una red portuaria y los asentamientos urbanos con que estos puertos interactúan.

Dado el escenario anterior, la planificación portuaria hoy en día requiere de un estudio prospectivo de la utilización de los distintos nodos portuarios de una red para mantener un balance eficiente de capacidad de transferencia (incluyendo todos sus subprocesos, esto es muelle, acopio y acceso), minimización de costos y mantención de las externalidades negativas bajo control. En esta misma línea, la localización de un nuevo nodo portuario requiere responder al menos las siguientes preguntas fundamentales: ¿Dónde se debiera localizar un nuevo nodo portuario?, ¿Qué demanda por servicios enfrentará en el futuro?, ¿Cuál es el impacto que generará en la red y en su entorno?. Esta tesis busca ayudar a responder estas preguntas, las cuales a su vez abren una serie de dimensiones a considerar. Entre ellas, el tipo de mercancías a las que el puerto estará orientado, la capacidad de transferencia asociada a los tipos de mercancía definidos y la competencia (disminución de demanda) o cooperación (aumento de demanda) con los nodos portuarios correlacionados con la puesta en marcha de un nuevo nodo en la red.

La proyección de demanda por transporte de mercancías, fundamental para la planificación portuaria, es típicamente incorporada en los programas maestros de desarrollo de los puertos. Los estudios más tradicionales, proyectan la demanda a partir de un modelo basado en datos históricos de transferencia portuaria y tasas de crecimiento del PIB.

La captura de nuevas cargas, en tanto, se trata generalmente en base a escenarios o, en el caso de estudios más elaborados, con modelos de elección portuaria que consideran el potencial de captura de nuevas cargas en función de los costos (generalizados) de



transporte de mercancías entre el puerto y el origen o destino final dentro del hinterland¹.

Respecto al proceso de elección portuaria, existe numerosa literatura que analiza los factores que influyen en la elección, considerando diferentes agentes tomadores de decisión. La mayoría de los enfoques analizados consideran como principal tomador de decisión a las líneas navieras y/o a los exportadores de carga. En el caso de los primeros, localización, eficiencia portuaria, tarifas y tamaño de mercado son algunos de los factores señalados como determinantes de la elección (ver por ejemplo Ha, 2003; Guy y Urli, 2006; Tongzon y Sawant, 2007; Chang et al., 2008; Meersman et al., 2008; Wiegmans et al., 2008; Moya, Song et al., 2016, Moya y Feo-Valero, 2016). Desde el punto de vista de los exportadores, en tanto, los costos generalizados de transporte (asociados a la localización del puerto respecto a los exportadores) y la disponibilidad de servicios navieros, son ejemplos de los atributos más mencionados (ver por ejemplo Slack, 1985; Nir et al., 2003; Tiwari et al., 2003; Malchow y Kanafani, 2004; Tongzon, 2009; García-Alonso y Sánchez-Soriano, 2009; Ng et al., 2013; Veldman et al., 2015; Moya y Feo-Valero, 2016).

En ambos casos, estas son sólo algunas de las variables que entran en juego en la elección (De Langen, 2007). En efecto, la capacidad negociadora, especialización en un tipo de mercancía, contratos de largo plazo, integración comercial, son otros factores relevantes en la captura del mercado y que son de muy difícil predicción. Las condiciones comerciales señaladas, dependerán además del tamaño de los exportadores. Mientras mayor volumen de carga posean, mayor será su capacidad de influir en la decisión de puerto de recalada de las naves (Slack, 1985; Tiwari et al., 2003; Chang et al., 2008; García-Alonso y Sánchez-Soriano, 2009).

¹ El concepto de hinterland no es fácil de definir. Weigend (1958), fue uno de los primeros investigadores en utilizar este concepto para el transporte de mercancías, definiéndolo como un "espacio de tierra organizado y desarrollado que está conectado por un puerto por medio de líneas de transporte y que recibe o envía mercancías a través de ese puerto". Así, el hinterland estará formado por la agregación de las unidades geográficas desde las cuales los exportadores deciden transportar sus mercancías a un puerto determinado. Weigend señaló también, que un puerto podría tener más de un hinterland si se considera el tipo de mercancías transportadas.



La evidencia empírica, ha mostrado que, al generarse nuevos servicios navieros (producto, por ejemplo, de un contrato de un gran exportador con una naviera), se producen a su vez mejores condiciones de negociación para exportadores más pequeños. En ese nuevo escenario, muchos de los exportadores seguirán el comportamiento de sus vecinos, pudiendo variar su elección portuaria. Lo anterior se refuerza con el hecho de que la ubicación espacial de ellos se vincula frecuentemente con actividades económicas similares y, por lo tanto, sus necesidades portuarias y navieras suelen ser compartidas.

Por otra parte, la proximidad de los exportadores tiene también un efecto en las redes, ya que, al utilizar las mismas vías, se genera mayor congestión, y por consiguiente, un aumento de costos de transporte.

Este efecto de red, así como cambios en las condiciones ofrecidas por los puertos competidores, define un comportamiento dinámico del hinterland dando cuenta de la interacción espacial que existe en el proceso (ver Wang et al., 2016). Su variación, sin embargo, depende a su vez de la inercia del proceso de elección portuaria, la cual varía dependiendo de quién sea el tomador de decisión. De Langen (2007), muestra, por ejemplo, que los exportadores tienen una elasticidad precio de la demanda portuaria menor que los freight forwarders². Lo anterior puede explicarse por el tipo de negocio asociado a cada uno de ellos. Así, en el caso de los exportadores, puede existir una mayor disposición al pago debido a que las condiciones ofrecidas por el puerto son muy adecuadas para sus productos. Sin embargo, en el caso de los freight forwarders, dado que su ganancia está en la disminución de costos, responderán más rápidamente a una variación en las tarifas.

Tanto la interacción espacial como temporal señalada, son características que han sido recogidas en muy pocos estudios. De hecho, la exhaustiva revisión bibliográfica realizada, indica que este podría ser el primer intento en modelización de elección

² Los freight forwarders son conocidos también como agentes de carga. Son personas o compañías que organizan envíos desde el fabricante o productor al mercado de destino. Realizan los contratos de transporte (carretero, marítimo y/o aéreo), actuando como intermediarios entre exportadores o importadores y las empresas de transporte. Generalmente ofrecen soluciones de transporte, almacenaje, embalaje de los productos, encargándose también de los procesos documentales.



portuaria y de formación del hinterland con un tratamiento de sistema complejo en las dimensiones antes señaladas.

Hasta ahora la interacción espacio-temporal del proceso analizado, había sido abordada principalmente con modelos de elección discreta, desconociendo en muchos casos, los supuestos en que se basan dichos modelos. Sólo hay un estudio que ha intentado reflejar esta interacción cumpliendo con los supuestos de modelización de elección discreta. Sin embargo, este estudio, realizado por la autora de esta tesis (Garrido y Leva, 2004), considera una estructura de modelización altamente compleja en calibración y aplicación.

En resumen, las características del proceso de elección portuaria dan cuenta de un sistema complejo, debido entre otras cosas, a la cantidad de actores, sus interrelaciones y los numerosos factores que influyen en él. Meersman et al. (2010), por ejemplo, señala que además de la multiplicidad de actores, existen variables exógenas como aumento del comercio mundial o el rápido desarrollo de tecnologías en equipamiento portuario, que dificultan aún más esta modelización. En Moya y Feo-Valero (2016), se indica, además, que las decisiones de los diferentes agentes están altamente interrelacionadas y que la naturaleza iterativa del proceso de elección hace que los determinantes varíen con las características, necesidades y objetivos del tomador de decisión.

Todo lo anterior plantea un gran desafío en términos de modelización, sin embargo es tal la importancia que tiene este proceso en la estimación de demanda portuaria, así como en el impacto en el resto de la red, que se requiere encontrar enfoques de modelización que recojan la dinámica del sistema y que permitan aproximar de manera simple el comportamiento del hinterland frente a cambios en las condiciones de éste.

1.2. Objetivos

Esta tesis tiene como objetivo desarrollar una herramienta de modelización de la captura de mercado por parte de un nodo portuario que enfrenta competencia por parte de otros nodos dentro de una red de transporte marítimo-terrestre.

En cuanto a los objetivos específicos de esta tesis, éstos son los siguientes:



- Revisión de la bibliografía técnica y científica sobre los métodos existentes en la literatura respecto de la elección portuaria, las variables decisivas y los agentes involucrados en este proceso.
- Identificación de los factores relevantes en la toma de decisiones de los agentes que intervienen en la elección portuaria
- Identificación de las relaciones entre los factores relevantes y el comportamiento de los agentes envueltos en el transporte de mercancías
- Desarrollo de un modelo para predecir la elección portuaria para el transporte de mercancías.
- Comprobación de la validez del método de modelización y análisis de su potencial aplicación en la práctica, con especial énfasis en las necesidades de datos.

1.3. Metodología

Para cumplir los objetivos anteriores se desarrolló un nuevo modelo de elección portuaria, en el que se consideran variables vinculadas a los puertos del conjunto de elección, así como la influencia espacial que tiene la elección portuaria de una zona exportadora sobre otra espacialmente cercana (interacción espacial).

Se utilizó una herramienta que ha demostrado ser de una enorme capacidad predictiva y descriptiva de fenómenos de distinta naturaleza: la teoría de autómatas celulares (AC).

En efecto, tal como se muestra en Goles y Martínez (1998) este tipo de herramienta permite modelar fenómenos complejos y no lineales (como la demanda por servicios portuarios) con gran eficiencia computacional y uso de información. Principalmente, estas herramientas logran captar con facilidad interacciones espacio-temporales que en otros contextos de modelización son extremadamente difíciles de incorporar y calibrar (Garrido y Leva, 2004).

Se postula aquí, que la elección portuaria es una decisión de múltiples actores que interactúan entre sí y que ponderan una multiplicidad de aspectos y comportamientos espaciales que configuran una elección dinámica que un AC puede imitar sin necesidad de conocer a todos los actores o a todos los determinantes de un proceso discreto que cambia en el tiempo.



1.4. Estructura de la Tesis

Esta tesis ha sido estructurada en siete capítulos, que permiten comprender los distintos hitos de la investigación y guiar al lector en el proceso metodológico seguido para llegar a la definición de la herramienta de modelización que responde al objetivo de esta tesis.

En el Capítulo 1 se presenta una introducción al contexto general en que se enmarca esta tesis doctoral. En este Capítulo se presenta una motivación al tema escogido y se plantea el problema que se abordará.

En el Capítulo 2 se describe el estado del arte en las materias relacionadas con el objetivo de esta tesis. Aquí se encuentran los elementos que resultan clave para el diseño de la estructura de modelización propuesta. Se definen los agentes, variables e interacciones que diversos investigadores han presentado en la literatura especializada, y se analizan las diferentes metodologías de modelización que se han utilizado para reflejar el comportamiento de elección portuaria.

En el Capítulo 3 se presenta la metodología utilizada en la tesis, es decir, se comienza a aportar al estado del arte en función de una propuesta de modelización novedosa y eficaz para el análisis de captura de mercado por parte de un nodo portuario. Se realiza aquí la descripción de la estructura del modelo de elección portuaria y del algoritmo de calibración utilizado.

En el Capítulo 4, para efectos de comparación, y considerando el estudio precursor de esta tesis, se describe la estructura que tendría un modelo de elección discreta si se quisieran considerar las características del sistema recogidas por el modelo predictivo escogido en esta tesis.

En el Capítulo 5, se presentan los resultados obtenidos de la calibración del modelo propuesto. Se aborda también el tema de la existencia y calidad de datos en la práctica para poder alimentar modelos de elección portuaria del tipo autómatas celular.

En el Capítulo 6, se describen los resultados de los ejemplos utilizados para probar la capacidad predictiva del modelo, mostrando para esto dos enfoques posibles de aplicación.



Finalmente, en el Capítulo 7, se presentan las conclusiones y recomendaciones para estudios futuros y para la implementación práctica de una herramienta de modelización predictiva basada en autómatas celulares.

1.5. Aportaciones

En esta tesis se han establecido tres preguntas fundamentales que guiaron el curso de la investigación:

- ¿Cuál es el estado del arte en modelización de elección portuaria?
- ¿Qué estructura de modelización de sistemas puede ser diseñada para representar el fenómeno de captura de mercado en un sistema portuario?
- ¿Qué valor añadido se puede aportar a los tomadores de decisiones frente a la elección portuaria y cómo se puede hacer éste más accesible?

Para lograr las respuestas señaladas, esta tesis propuso avances en los siguientes aspectos:

- a) Revisión bibliográfica actualizada de los estudios publicados en el campo de la elección portuaria.
- b) Síntesis y clasificación de los estudios más relevantes publicados a la fecha.
- c) Generación de una estructura de modelización basada en autómatas celulares apropiada a los objetivos planteados.
- d) Generación de un método de validación de la calidad predictiva de la estructura de modelización propuesta.
- e) Aplicación de la estructura diseñada, a un conjunto de datos reales con un análisis detallado de las implicancias de utilizar el método propuesto.
- f) Propuesta de implementación de mejoras/adiciones a la estructura propuesta, a fin de facilitar el proceso de toma de decisiones en el ámbito portuario.

Como un sub-producto de la tesis, se desarrolló una herramienta de modelización que permitió aproximarse, de manera simple, a un problema no resuelto: el efecto de la complejidad en las interacciones espacio-temporales de los exportadores en la elección del puerto y en la definición del hinterland.

Esta herramienta, de rápida calibración y aplicación, es una base de modelización que podría ser de enorme valor para los tomadores de decisiones, ya que permite evaluar el



impacto de diversas políticas (como mejoras en accesibilidad, planificación territorial, localización portuaria, etc.) y de acciones que se vinculen directamente a la operación portuaria.

A continuación, se listan los artículos que se realizaron durante el desarrollo de esta tesis. Estos artículos muestran la evolución que tuvo la estructura de modelización a lo largo de la investigación:

- IV. Modelación de elección portuaria: Un enfoque de autómatas celulares probabilísticos. Actas XIX Congreso Panamericano de Ingeniería de Tránsito, Transporte y Logística, Ciudad de México, 2016.
- V. A probabilistic cellular automaton to forecast port choice decisions. *Transport*. Segunda revisión.
- VI. Port hinterland formation: A Probabilistic Cellular Automaton Approach. *Naval Research Logistics*. Primera revisión.

Los artículos mencionados, son presentados en los anexos, en el mismo orden que fueron listados.

Adicionalmente, se incluye en anexos el siguiente artículo desarrollado por la autora de esta tesis durante su Magíster en Ciencias (ver Leva, 2002), y que fue precursor de esta investigación.

- Port of destination and carrier selection for fruit exports: a multi-dimensional space-time multinomial probit model. *Transportation Research Part B* 38 (7). pp. 657-667. (2004)



Capítulo 2

ESTADO DEL ARTE



2. REVISIÓN BIBLIOGRÁFICA

Se ha realizado una amplia investigación sobre el estado del arte en modelización de elección portuaria, la cual fue agrupada en los siguientes ámbitos:

- Agentes tomadores de decisión en el proceso de elección portuaria.
- Factores que afectan la elección portuaria desde el punto de vista de los distintos agentes.
- Metodologías de modelización de elección portuaria, especificándose aquellas que consideran los efectos de interacción espacial y temporal dentro de su especificación.

2.1. Agentes Tomadores de Decisión

El proceso de elección portuaria se caracteriza por tener una gran cantidad de actores interactuando entre sí, y que toman su decisión en base a factores que no son siempre comunes entre ellos o que no ponderan de igual manera.

De acuerdo a la literatura revisada, no hay consenso respecto a cuál es el principal agente tomador de decisión, sin embargo, las investigaciones se han centrado básicamente en los señalados en la Figura 1.



Figura 1: Tomadores de decisión claves en el proceso de elección portuaria.

De acuerdo a Moya y Feo-Valero (2016), la identificación del tomador de decisión es uno de los principales problemas en modelización de elección portuaria. Señalan que las decisiones de los diferentes agentes están altamente interrelacionadas y que la naturaleza iterativa del proceso de elección hace que los determinantes varíen con las características, necesidades y objetivos del tomador de decisión. Destacan que, desde el punto de vista marítimo, las líneas navieras diseñan su red de servicios para maximizar economías de escala, mientras que, desde el punto de vista de los actores terrestres, el objetivo es minimizar los costos.

Leva (2002), por su parte, indica que el tipo de acuerdo de compra es un factor clave en la definición del tomador de decisión. En su investigación estudió el proceso de elección de puerto de desembarque de exportaciones frutícolas chilenas, determinándose que, cuando los productos son vendidos con un acuerdo de compra de venta firme³ en el país de destino, será el importador quien tomará la decisión respecto a la naviera y puerto. En el caso de venta a consignación⁴, es el exportador quien tendrá la mayor influencia en la elección. Lo anterior es también señalado en Tongzon (2009).

Aronietis et al. (2010), en tanto, analizan 30 artículos de elección portuaria con el fin de identificar a los actores clave del proceso, así como sus criterios de elección. Dentro de los actores identificados se encuentran: exportadores, freight forwarders, compañías navieras y operadores de terminales. Señalan también, que en algunos casos, la decisión que tomen estos actores puede ser influenciada por las autoridades portuarias y las

³ La venta a firme corresponde a un acuerdo de compra donde el precio de la mercancía exportada se fija previo a su envío. En este caso el importador pasa a ser propietario de la mercancía (Leva, 2002).

⁴ La venta a consignación es un acuerdo de compra entre exportador e importador, en el que el importador recibe una comisión por la venta de la mercancía. En este caso el exportador es dueño de la mercancía hasta el minuto de su venta en el destino final (Leva, 2002).



agencias de gobierno. De acuerdo a la revisión realizada por los autores, en la mitad de los estudios analizados se identifica al exportador como el principal actor en la elección portuaria. Las compañías navieras, en tanto, son consideradas en 7 artículos como agentes decidores del proceso de elección, seguidas por los freight forwarders y operadores de terminales, quienes son destacados en ese rol, en 5 y 3 artículos, respectivamente.

Sin embargo, hay investigaciones (Tongzon, 2009; Malchow y Kanafani, 2004; Ugboma et al. 2006; Ng et al., 2013) que a pesar que reconocen al exportador como actor clave en el proceso de elección, consideran que la decisión que éste tome se encontrará limitada por los contratos con líneas navieras o freight forwarders. En este caso, ven a la elección del exportador implícita en la elección de estos actores y por tanto, podría también enfocarse el análisis en ellos.

No obstante, Moya y Feo-Valero (2016), indican que considerar la elección portuaria desde el punto de vista naviero, es hoy aún más complejo, debido a los procesos de fusiones y alianzas estratégicas de las empresas de transporte marítimo, que dificultan la determinación del real tomador de decisión.

Existe también una discusión respecto al rol que juega el tamaño del exportador en el proceso de elección portuaria (Slack, 1985; Aronietis et al., 2010; Steven y Corsi, 2012; y Moya y Feo-Valero, 2016). Slack (1985), por ejemplo, considera que los grandes exportadores tienen mayor peso en la elección, a diferencia de Steven y Corsi (2012), que ven en los medianos y pequeños una mayor influencia. Estos últimos autores justifican su postura, en que los exportadores de gran tamaño poseen en general, contratos de largo plazo con compañías navieras y, por lo tanto, tienen menor incidencia en el proceso de elección.

Sin embargo, asumir que los contratos de largo plazo entre líneas navieras y grandes exportadores, disminuyen el peso que estos últimos puedan tener en la decisión, podría ser un enfoque inadecuado. La decisión por sí misma existe, pero ocurre con una frecuencia menor que en los casos de exportadores que no poseen dichos contratos. En otras palabras, independiente del tamaño del exportador, éste participa del proceso de



elección portuaria, pero con una periodicidad variable, dependiente, entre otras cosas, de la duración de sus contratos.

Moya y Feo-Valero (2016), indican que al no existir claridad en quién es el tomador de decisión, el cual puede variar además según tipo de industria y/o país, el enfoque correcto sería estudiar conjuntamente el proceso de decisión de los distintos actores. Sin embargo, la mayoría de los estudios limitan su análisis a solo uno de ellos, lo cual se explicaría por la complejidad matemática del análisis del proceso de decisión conjunto y la dificultad en la identificación de los objetivos individuales de la elección y su interrelación.

Así, el desafío de la modelización del proceso de elección portuaria, está no sólo en definir una estructura de modelización que recoja las relaciones del sistema, sino que además sea de fácil calibración y aplicación.

2.2. Factores de Elección Portuaria

Tal como se señaló en la sección anterior, existen distintas aproximaciones para representar el proceso de elección portuaria. En la mayoría de los artículos analizados, las variables de este proceso han sido estudiadas dependiendo del actor que se considere como agente tomador de decisión.

Teniendo en cuenta lo anterior, se presenta a continuación una descripción de los hallazgos reportados en la literatura, respecto a las características y factores identificados en el proceso de elección portuaria, según el punto de vista del tomador de decisión considerado.

2.2.1. Factores de Elección desde el Punto de Vista de los Exportadores

Desde el punto de vista de los exportadores, la localización del puerto, entendida como distancia, tiempo o costo de viaje, es una de las variables más comúnmente mencionadas en la literatura especializada (ver por ejemplo Slack, 1985; Nir et al., 2003; Tiwari et al., 2003; Malchow y Kanafani, 2004; Tongzon, 2009; García-Alonso y Sánchez-Soriano, 2009; Steven y Corsi, 2012; Ng et al., 2013; Veldman et al., 2011, 2013, 2015; Vega et al., 2014; Kramberger et al., 2015; Wang et al., 2016; Moya y Feo-Valero, 2016).



García-Alonso y Sánchez-Soriano (2009), indican, por ejemplo, que los exportadores tienden a elegir el puerto más cercano, y que incluso al momento de decidir la localización de sus instalaciones, tienen a la vista los puertos que ofrecen los servicios que necesitan. Sin embargo, los autores señalan que la localización no es suficiente para explicar el éxito de un puerto. Hay otras variables, como el número de líneas navieras, que también afectan la decisión del exportador (esta variable es mencionada también en Slack, 1985; Nir et al., 2003; Tiwari et al., 2003; Malchow y Kanafani, 2004; Ugboma et al., 2006; De Langen, 2007; Tongzon, 2009; Steven y Corsi, 2012; Vega et al., 2014). García-Alonso y Sánchez-Soriano (2009), señalan, además, que existe una retroalimentación entre el volumen de carga del hinterland y los servicios navieros disponibles en un puerto, ya que al haber mayor volumen de carga hay más navieras interesadas en recalar en dicho puerto, generando con esto una mayor atraktividad para los exportadores.

Además de la componente espacial asociada a la localización del puerto recién señalada, se menciona también en la literatura, una cierta inercia temporal en la elección portuaria (Leva, 2002; Nir et al., 2003; Garrido y Leva, 2004; Langen, 2007). Esta inercia podría explicarse, por ejemplo, por la rigidez de la cadena logística a la que pertenece la mercancía movilizad, o por la duración de los contratos entre exportadores y otros actores del sistema (por ejemplo, navieras, puertos y freight forwarders).

El efecto de la cadena logística en el proceso de elección, es mencionado en Magala y Sammons (2008), Tongzon (2009) y Moya y Feo-Valero (2016), entre otros. Magala y Sammons (2008), señalan, por ejemplo, que la influencia del exportador en el proceso de decisión ha disminuido producto de la coordinación al interior de la cadena de suministro, la cual ha permitido acceder a diversas opciones de servicios de transporte desde el origen hasta el destino final, disminuyendo el costo total de la logística y maximizando el valor para el cliente y el proveedor. En ese caso el puerto es sólo un eslabón en esa cadena, y por lo tanto la elección no se centra en él. El impacto de estos servicios de transporte puerta a puerta es también reportado en Malchow y Kanafani (2004).

Sin embargo, Tongzon (2009), señala que a pesar de que los puertos se consideran parte de las cadenas de suministro y, por lo tanto, no deben ser vistos por sus usuarios de



forma aislada, en muchos casos, sobre todo en los países en desarrollo, los puertos aún no están bien integrados con otros elementos de la cadena.

Respecto a las diferencias que pudiesen existir en los factores de elección portuaria, dependiendo del tamaño del exportador, Steven y Corsi (2012), señalan que los grandes exportadores son más sensibles a factores que permiten una entrega más rápida de la carga, mientras que los pequeños consideran más relevante los costos de transporte marítimo que un menor tiempo de entrega.

Sin embargo, la mayoría de los autores consideran al costo de transporte terrestre, y no al costo de transporte marítimo, una variable relevante para el exportador en su proceso de elección (Slack, 1985; Nir et al., 2003; Ng et al., 2013; Vega et al., 2014; Wang et al., 2016). Vega et al. (2014), determinan además que la probabilidad de elección de puerto es altamente elástica respecto del costo de acceso. Lo anterior es consistente con lo señalado por Halim et al. (2016), quienes recalcan la importancia de los costos asociados al hinterland respecto a los costos totales, esto a pesar que la distancia recorrida es solo el 10% de la distancia total.

Hay estudios, en tanto, en que el costo de transporte terrestre y el marítimo son considerados relevantes en la decisión del exportador. En Veldman et al. (2011), por ejemplo, definen como variables claves en la elección de puertos de contenedores en España, el costo de transporte terrestre, el costo de transporte marítimo y la calidad de servicio del puerto.

Sin embargo, en una investigación posterior, Veldman et al. (2015), indican que es difícil disponer de información de costos marítimos y portuarios, y por lo tanto proponen utilizar variables no monetarias como aproximación de las anteriores. Dentro de estas variables proxy se encuentran: tamaño de la nave, factor de carga y velocidad de transferencia portuaria. Las variables fueron testeadas en modelos de elección discreta, considerando, además, distancia del exportador al puerto, calidad de servicio del puerto (representado por el indicador Mohring⁵) y necesidad de servicios feeder hacia el

⁵ El efecto Mohring se produce cuando al aumentar el uso de facilidades de transporte, se mejora la situación de bienestar de otros usuarios. En este paper la variable Mohring se expresa como uno menos el inverso del movimiento de contenedores de cada puerto.



destino final. Las pruebas demostraron que las variables proxy sugeridas eran exitosas como sustituto a los costos marítimos.

Otra variable relevante considerada en el proceso de elección portuaria tanto del punto de vista de los exportadores como desde el punto de vista marítimo, es el rendimiento portuario (Ugboma et al., 2006; De Langen, 2007; Tongzon, 2009; Steven y Corsi, 2012; Ng et al., 2013; Moya y Feo-Valero, 2016). Moya y Feo-Valero (2016), indican que en la literatura analizada en su investigación, no existía una forma única de medir esta variable, pero se podían distinguir dos componentes: eficiencia y eficacia portuaria. En el caso de la primera, se han utilizados proxys como el número de grúas o el número de sitios de atraque (ver, por ejemplo, Tiwari et al., 2003; Steven y Corsi, 2012). En cuanto a la eficacia portuaria, ésta ha sido considerada más bien en enfoques desde el punto de vista marítimo, por lo que su análisis se realizará en la sección siguiente.

Finalmente, y tal como se señala en Moya y Feo-Valero (2016), los factores de elección portuaria desde el punto de vista del exportador, pueden variar, además, dependiendo del país en que se encuentran localizado, del tipo de industria a la cual pertenecen, del tamaño de la empresa, del valor de los productos, de los requisitos logísticos del envío, de los sistemas de fabricación e inventario, entre otros.

A modo de resumen, se presenta en la Tabla 1, los artículos revisados y los principales factores de elección portuaria considerados en cada uno de ellos.



Autor	Principales factores de elección ⁶
Slack (1985)	– Costos de transporte terrestre – Localización – Congestión portuaria – Frecuencia navieras – Intermodalidad – Costos portuarios – Infraestructura
Nir et al. (2003)	– Tiempo de viaje (puerta/puerto) – Costos de viaje (puerta/puerto) – Frecuencia semanal de navieras – Número de rutas navieras disponibles
Tiwari et al. (2003)	– Distancia – Frecuencia navieras – Tamaño de flota naviera – Congestión portuaria (medida como Teus transferidos) – Número de líneas de atraque – Tamaño de flota naviera
Malchow y Kanafani, (2004)	– Distancia – Frecuencia navieras
Ugboma et al. (2006)	– Localización – Eficiencia – Frecuencia naviera – Infraestructura – Buena reputación en daños y demoras – Calidad de servicio portuario – Costos portuarios

⁶ El orden no representa prioridad de factores



Autor	Principales factores de elección ⁶
Tongzon (2009)	– Localización – Frecuencia naviera – Infraestructura – Eficiencia portuaria – Tarifas portuarias – Respuesta rápida a las necesidades de los usuarios portuarios – Reputación del puerto respecto al daño de la carga
García-Alonso y Sánchez-Soriano (2009)	– Distancia – Constante de atractividad portuaria
Veldman et al. (2011)	– Costos de transporte terrestre – Costos de transporte marítimo – Calidad de servicio del puerto
Steven y Corsi (2012)	– Tiempo de tránsito terrestre – Frecuencia naviera – Tamaño promedio de naves – Número de líneas de atraque – Productividad de la grúa – Congestión portuaria – Administración del puerto (pública o privada) – Costos portuarios – Tiempo de tránsito marítimo – Tamaño del exportador
Veldman et al. (2013)	– Costos de transporte terrestre – Costos de transporte marítimo – Efecto Mohring⁷ – Desbalance de contenedores de importación y exportación

⁷ Ver nota al pie N° 1.



Autor	Principales factores de elección ⁶
Ng et al. (2013)	– Distancia – Costos de transporte terrestre – Eficiencia portuaria – Costos portuarios – Calidad de servicio portuario
Vega et al. (2014)	– Costo de acceso al puerto – Frecuencia naviera – Tarifa marítima – Tiempo de tránsito marítimo
Veldman et al. (2015)	– Distancia terrestre – Índice portuario (factor de carga y velocidad de transferencia portuaria) – Índice marítimo (tamaño de la nave, distancia marítima) – Efecto Mohring⁸ – Dummy por servicio feeder (1 si se necesita servicio feeder para el destino final, 0 sino)
Kramberger et al. (2015)	– Distancia – Tiempo de transporte – Tasa de preferencia
Wang et al. (2016)	– Costos de transporte de rutas intermodales

Tabla 1: Factores de elección portuaria desde el punto de vista de los exportadores de carga o freight forwarders.

2.2.2. Factores de Elección desde el Punto de Vista Naviero

Tal como se señaló en la sección 2.2.1, el rendimiento portuario ha sido una variable considerada en variados estudios de elección portuaria, tanto desde el punto de vista del exportador como desde el punto de vista naviero. Se mencionó también que el rendimiento podía ser entendido como la descomposición en eficiencia y eficacia

⁸ Ver nota al pie N° 1.



portuaria. La eficiencia, de acuerdo a la literatura revisada, se representa de manera similar independiente del tomador de decisión. Así, indicadores como número de grúas o número de muelles, son utilizados en ambos enfoques de modelización (Tiwari et al., 2003; Steven y Corsi, 2012; Moya y Feo-Valero, 2016). Sin embargo, la componente de eficacia, es un factor de elección que se considera más relevante en las navieras, que en los exportadores. Generalmente se representa través de variables como: calidad de servicio, información disponible al usuario y/o reputación del puerto (ver por ejemplo Ha, 2003; Guy y Urli, 2006; Tongzon y Sawant, 2007; Wiegmans et al., 2008; Aronietis et al., 2010).

En un análisis comparativo de niveles de servicio de los 15 principales puertos de contenedores del mundo, Ha (2003), encontró que perfeccionar la calidad y la cantidad de información, era una de las variables fundamentales para mejorar la percepción de las navieras respecto a los puertos, quienes podrían aumentar así, su probabilidad de elección.

Respecto a otros factores considerados por las líneas navieras al momento de escoger puerto, se encuentran por supuesto los costos portuarios (Lirn et al., 2004; Guy y Urli, 2006; Tongzon y Sawant, 2007; Chang et al., 2008; Wiegmans et al., 2008; Moya y Feo-Valero, 2016; Song et al., 2016). Tongzon y Sawant (2007), por ejemplo, señalan que dos puertos que comparten un hinterland, operando con altos niveles de eficiencia, necesitarían bajar sus tarifas, entregar nuevos servicios de valor agregado e infraestructura, para tener una ventaja competitiva sobre el otro.

En la literatura se mencionan también otras variables como volumen de carga local, disponibilidad de muelle, volumen transferido, conexión con alimentadores y localización del puerto (Guy y Urli, 2006; Chang et al, 2008; Wiegmans et al., 2008). Respecto a este último factor, Moya y Feo-Valero (2016), señalan que las líneas navieras buscan lograr un equilibrio entre la ubicación del puerto con respecto a las rutas navieras principales y la distancia a los principales centros productores/consumidores. Esto es destacado también por Malchow y Kanafani (2004), quienes demuestran que la distancia terrestre, resulta incluso ser más significativa que la distancia marítima.



El efecto de los costos de transporte en el hinterland, en la elección portuaria de las navieras, es considerado también por Song et al. (2016), quienes analizaron la competitividad portuaria, considerando además costos de transporte naviero, costos portuarios y costos de servicios de alimentadores.

En Meersman et al., (2008), por su parte, se hace énfasis en la importancia que tiene para un puerto contar con alternativas modales y capacidad en la red del hinterland, en particular para el tráfico de contenedores. Señalan que las compañías navieras tienden a escoger aquellos puertos con menor congestión, incluyendo en esto no sólo la operación en el puerto, sino también las opciones de transporte hacia el hinterland.

Hay también unos pocos estudios en que se mencionan factores, como los temas ambientales, que podrían ser relevantes en la elección portuaria según las navieras (ver por ejemplo Wiegmans et al., 2008).

Todas las variables mencionadas, sin embargo, no tendrán el mismo peso en el momento de la elección. Guy y Urli (2006), por ejemplo, señalan que no existe una solución óptima de los ponderadores de las variables, sino más bien una solución compromiso que permite cumplir con los objetivos de la línea naviera. Lo anterior es aún más complejo en el escenario actual, en que las navieras han aumentado las alianzas entre ellas y por lo tanto son múltiples los objetivos que se deben satisfacer (Jiang et al., 2015).

Además, al momento de definir los factores de elección portuaria desde el punto de vista naviero, es necesario tener en cuenta el tipo de servicio ofrecido por dicha naviera. Esto es señalado en Moya y Feo-Valero (2016), quienes indican que en el caso de las rutas navieras interoceánicas, los costos portuarios, la ubicación geográfica, las conexiones con el hinterland, la infraestructura y eficiencia portuaria son los criterios más importantes. En las rutas regionales, sin embargo, el tamaño de la economía del hinterland y el volumen de carga transferida, son los principales factores que influyen en la elección.

En la Tabla 2 se presenta un resumen de los estudios analizados, destacando las variables de elección portuaria que se consideran en ellos.



Autor	Principales factores de elección ⁹
Guy y Urli (2006)	– Calidad Infraestructura – Costos portuarios – Servicios portuarios – Localización
Tongzon y Sawant (2007)	– Eficiencia portuaria – Localización – Infraestructura – Tarifas portuarias – Conectividad – Volumen de carga – Servicios portuarios
Chang et al. (2008)	– Volumen de carga local – Tarifa del terminal – Disponibilidad de muelle – Localización – Volumen transferido – Conexiones de alimentadores
Wiegmans et al. (2008)	– Localización – Eficiencia portuaria – Reputación – Calidad del servicio – Intermodalidad – Tarifas portuarias – infraestructura
Meersman et al. (2008)	– Capacidad de conexión con el hinterland – Intermodalidad – Calidad de servicios de transporte

⁹ El orden no representa prioridad de factores

Autor	Principales factores de elección ⁹
Song et al. (2016)	– Costos portuarios – Costos de transporte naviero – Costos de transporte en el hinterland – Costos de servicios alimentadores

Tabla 2: Factores de elección portuaria desde el punto de vista naviero

2.2.3. Factores de Elección desde el Punto de Vista de más de un Actor

Tal como se mencionó al comienzo de esta sección, en el proceso de elección portuaria hay una serie de actores involucrados, siendo exportadores y navieras los más influyentes. Sin embargo, hay estudios que se han centrado en conocer cómo varían los factores de elección entre estos actores y otros que están presentes en el proceso. Por ejemplo, Song y Yeo (2004), consideraron la opinión de armadores, exportadores, navieras, operadores de terminales, académicos e investigadores, para determinar los elementos claves que contribuyen a la competitividad global de los principales puertos Chinos. Identificaron 73 elementos, siendo escogida como variable principal, la localización.

De igual forma, De Langen (2007), al consultar a exportadores y freight forwarders, coincide en dicha variable como la más relevante, siendo las características del producto (por ejemplo, su valor o si es o no refrigerado) y el número de rutas y servicios navieros disponibles en el puerto, las que le siguen en importancia. Sin embargo, el mismo autor, indica que existen diferencias en el *comportamiento* del tomador de decisión que no son considerados en el análisis, y que son relevantes en la elección portuaria. Para evaluar estas diferencias, De Langen, realizó una encuesta a exportadores y freight forwarders, en la que solicitaba evaluar una serie de proposiciones respecto a la elección portuaria. Los resultados indicaron que, a pesar que ambos actores tenían opiniones similares en la mayoría de las proposiciones, habían diferencias en las afirmaciones “*un bajo precio puede compensar un nivel de servicio bajo*” y en “*si el puerto que uso actualmente provee un servicio satisfactorio, no tengo razón para cambiarme, incluso si existen desventajas en el precio*”, mostrando los exportadores una menor disposición a aceptar niveles de servicio menores y a cambiarse de puerto por motivos de tarifas. En



su investigación, el autor, analizó luego la distribución del hinterland entre puertos competidores, para lo cual lo clasificó en primario y secundario. Definió hinterland primario como aquel que considera todas aquellas zonas en que su costo generalizado de transporte es notoriamente menor hacia un puerto. El hinterland secundario, en tanto, lo definió como aquel compuesto por zonas en que no existe una ventaja clara del costo generalizado de transporte hacia un puerto sobre otro competitivo. Este enfoque lo aplicó en el análisis de la partición de mercado de importaciones Austríacas entre puertos competidores, observando que existía una cierta inercia al cambio de puerto. Este desfase entre la variación de los costos y la reasignación portuaria, la explicó por la dificultad que implica modificar la estructura de las cadenas logísticas.

Lirn et al. (2004), en tanto, determinaron que costos portuarios, proximidad a las principales rutas de navegación, proximidad a las zonas en que se ubican importadores/exportadores, condiciones de la infraestructura y conexión con alimentadores, eran los factores más relevantes según la opinión de las navieras, y de operadores portuarios. Estos actores coincidían en las variables, pero las ponderaban de manera distinta.

Lo anterior se observa también en la investigación de Lee et al. (2010). Estos autores identifican importantes diferencias en la evaluación de 38 factores involucrados en la elección portuaria, al consultar a exportadores de carga, navieras y operadores de terminales de contenedores. Estos dos últimos, poseían un mayor número de atributos coincidentes, dando mayor importancia a las “condiciones de operación del terminal”, que los exportadores de carga. La única variable ponderada igualmente por los 3 actores fue “accesibilidad portuaria”.

Aronietis et al. (2010), por su parte, analizan las variables consideradas por exportadores, freight forwarders, compañías navieras y operadores de terminales. Determinan que factores como costos portuarios, conexión con el hinterland, capacidad portuaria, confiabilidad, localización, carga disponible, flexibilidad, servicio al cliente, frecuencia y riesgo de pérdida/daño de la carga eran valorados por dichos actores.

A continuación, se presenta en la Tabla 3 un resumen con los actores considerados en cada artículo revisado, así como las principales variables que consideran relevantes en la elección portuaria.

Autor	Actores considerados	Principales factores de elección
Song y Yeo (2004)	– Armadores – Exportadores – Navieras – Operadores de terminales – Académicos e investigadores	– Localización – Volumen de carga – Infraestructura – Nivel de servicio – Costos portuarios
Lirn et al. (2004)	– Navieras – Operadores portuarios	– Costos portuarios – Proximidad a las principales rutas de navegación – Cercanía a importadores/exportadores – Infraestructura – Conexión con alimentadores
De Langen (2007)	– Exportadores – Freight Forwarders	– Localización – Características del producto – Número de rutas y servicios navieros – Calidad de las compañías que operan el terminal – Eficiencia en el manejo de la carga – Calidad del equipamiento – Calidad de los servicios de transporte – Información de servicios en el puerto – Buena reputación en daños y demoras – Foco en el cliente – Modos de conexión con el hinterland – Contacto personalizado

Autor	Actores considerados	Principales factores de elección
Lee et al. (2010)	– Navieras – Exportadores – Operadores de terminales de contenedores	– Localización – Accesibilidad portuaria – Atributos del puerto – Condiciones de operación portuarias – Características del hinterland – Características de las navieras – Otros
Aronietis et al. (2010)	– Navieras – Operadores de terminales – Exportadores – Freight forwarders	– Costos portuarios – Conexión con el hinterland – Capacidad portuaria – Confiabilidad – Localización – Carga disponible – Flexibilidad – Servicio al cliente – Frecuencia – Riesgo de pérdida/daño de la carga

Tabla 3: Factores de elección portuaria considerando más de un actor

2.3. Modelos de Elección Portuaria

Para la modelización del proceso de elección portuaria se han aplicado diversas metodologías. Las más simples se han basado en encuestas, cuyos resultados permiten priorizar los factores de elección en base a ranking de importancia (ver Slack, 1985; De Langen, 2007; Wiegman et al., 2008; Meersman et al., 2008; Tongzon, 2009; Aronietis et al., 2010; Ng et al., 2013). Sin embargo, para obtener resultados significativos se requiere además de tener un gran número de encuestas, un control adecuado del ejercicio, de manera de limitar los sesgos en las respuestas.

Hay otras metodologías que se basan en análisis multicriterio. Guy y Urli (2006), por ejemplo, lo utilizan para determinar el peso de variables como calidad de infraestructura, costos portuarios, servicios portuarios y localización.

Otro método ampliamente utilizado es el Proceso Analítico Jerárquico (AHP) (ver Song y Yeo, 2004; Lirn et al., 2004; Ugboma et al., 2006; Kramberger et al., 2015). Lirn et al



(2004), por ejemplo, lo aplican a 47 atributos de servicio que obtuvieron en su revisión bibliográfica. Luego, estos atributos los categorizaron en los siguientes cuatro criterios: infraestructura física y tecnológica, localización geográfica, gerencia y administración y costo de terminales. Los resultados del análisis mostraron que tanto navieras como operadores de servicios portuarios, consideraban similares atributos, pero con distinta ponderación.

Jiang et al. (2015), por su parte, utilizan modelos de optimización para representar el proceso de elección portuaria. Los autores, desarrollan dos modelos; uno que minimiza los tiempos de transporte, y otro que maximiza la capacidad de transporte en una red de puertos. Halim et al. (2016), también escoge un modelo de optimización, siendo en este caso multi-objetivo. Considera una función de costos logísticos totales, generando estructuras de distribución factibles del hinterland. El modelo, a partir de diferentes niveles de servicio, determina la localización de centros de distribución dentro de un hinterland portuario. Este modelo, sin embargo, no considera el impacto de las redes de las navieras en el hinterland, lo cual puede ser interesante de evaluar considerando las integraciones verticales que existen entre navieras y operadores logísticos.

Hay también investigaciones en que se ha utilizado más de un enfoque de modelización. Este es el caso de Kramberger et al. (2015), quienes ocupan los dos modelos recién mencionados, esto es, el Proceso Analítico Jerárquico (AHP) y optimización. En su análisis consideran, además de variables como distancia y tiempo de transporte, factores subjetivos del proceso de elección. Estos últimos los representan a través de una “tasa de preferencia” de puerto de origen y de destino. Las tasas de preferencia son calculadas utilizando un Proceso Analítico Jerárquico de factores subjetivos de la elección. En la primera etapa de este proceso, los factores son rankeados por importancia, de acuerdo a lo declarado por los exportadores en las encuestas realizadas. Luego los puertos son comparados entre sí de acuerdo a este ranking, obteniéndose así una tasa de preferencia respecto a sus competidores. La elección de puerto por su parte, se estima optimizando la ruta de transporte desde los puntos de producción a los puntos de consumo, basándose en los costos de transporte y las tasas de preferencia portuaria. Para esto, se utiliza un modelo de programación matemática mixto (considera variables enteras y continuas).



Los modelos de teoría de juegos, también han sido aplicados a elección portuaria. En Song et al. (2016), por ejemplo, utilizan esta herramienta de modelización, considerando un enfoque de cadena de abastecimiento. Para esto, definen una función de costos logísticos que incorpora costos portuarios, costos de transporte naviero, costos de transporte en el hinterland y costos de servicios alimentadores. Sin embargo, el modelo considera sólo dos puertos competidores y una línea naviera, siendo complejo aumentar cualquiera de estas dimensiones.

Yeo, et al (2014), por su parte, proponen un modelo conceptual de elección de puertos basado en lógica difusa. En este enfoque, diversas fuentes (y formatos) de datos, tanto objetivos como subjetivos, se combinan para evaluar el desempeño de un puerto. Para esto se determina una “puntuación difusa” o puntuación de preferencia. Esta puntuación permite comparar cada puerto, ya sea con el conjunto completo de variables consideradas en la elección, o con cada una de ellas de manera independiente. Permite además establecer medidas de desempeño individual de un mismo puerto respecto a sus propias características en periodos anteriores. El método se aplica en el análisis de elección de los principales puertos de contenedores del noreste asiático, teniendo como agente de decisión a una compañía naviera. Esta metodología, sin embargo, es altamente dependiente de la opinión de los stakeholders considerados, la que puede cambiar en el tiempo y en el espacio (en distintas regiones se le pueden otorgar distintos pesos a los atributos).

Tal como se ha descrito hasta ahora, son variadas las metodologías aplicadas en la modelización de elección portuaria. Sin embargo, la más utilizada en la literatura analizada, corresponde a modelos de elección discreta (Manski y McFadden, 1981). Estos modelos asumen lo siguiente: hay un agente racional que, teniendo información perfecta sobre un conjunto de alternativas disponibles, debe seleccionar una de ellas. Cada alternativa puede ser descrita apropiadamente a través de sus atributos. El agente que está llevando a cabo la selección tiene una función de utilidad aleatoria (FUA) que depende de los atributos de cada alternativa, algunas características individuales y una componente aleatoria que explica las desviaciones del comportamiento racional (dadas las variables explicatorias disponibles). Luego, cada agente selecciona un puerto de acuerdo a una distribución de probabilidad que maximiza la FUA.



Los supuestos sobre la componente estocástica de la FUA determinan la forma funcional del modelo. Los supuestos más comunes para esta componente son los que definen el Logit Multinomial (MNL, ver Hausman y McFadden, 1984). El MNL considera una estructura de modelización que requiere dos condiciones: que no haya correlación entre las alternativas, y que la componente aleatoria de la función de utilidad sea homocedástica, es decir la varianza de los errores aleatorios de todas las alternativas es igual. Ejemplos de aplicación de MNL en la elección portuaria se encuentran en Nir et al. (2003), Tiwari et al. (2003), Malchow y Kanafani (2004) y Veldman et al. (2011, 2015). Estos estudios tienen un buen rendimiento en términos de poder de predicción, sin embargo, es probable que el rendimiento reportado sea altamente sesgado, debido a que en los casos en que han sido aplicados, no se cumplen los más importantes supuestos de los MNL.

En efecto, la proximidad espacial de diferentes terminales portuarios y los contratos de largo plazo entre navieras y exportadores/importadores, son algunas de las características que establecen una condición de correlación espacial y temporal que no es fácil de tratar en el contexto de modelos de elección discreta. La condición espacio-temporal del proceso de elección portuaria, es señalada por ejemplo en Mangan et al. (2002), Leva (2002), Garrido y Leva (2004), García-Alonso y Sánchez-Soriano (2009), Veldman et al. (2013), Wang et al. (2016).

Por otra parte, las condiciones marítimas, la infraestructura de los puertos y la interfaz con el sistema de transporte terrestre (por nombrar algunos), imponen una gran variabilidad en los servicios portuarios, lo que crea condiciones para que alternativas del conjunto de elección, presenten un comportamiento heterocedástico.

Hay sólo unos pocos estudios publicados en revistas especializadas que se ocupan de modelos más robustos que permiten quebrantar las restricciones MNL. Uno de ellos resuelve parcialmente el problema, a través de un modelo Logit Jerárquico (ver Papola, 2004), que si bien no puede trabajar correlaciones generales si puede considerar heterocedasticidad. Aplicaciones de Logit Jerárquico en modelización de elección portuaria podemos encontrar por ejemplo en Veldman et al. (2013) y Vega et al. (2014).



García-Alonso y Sánchez-Soriano (2009), por su parte, incorporan la componente espacial en su análisis de elección portuaria, considerando a ésta desde la perspectiva del hinterland. La metodología utilizada consideró preferencias reveladas, entendiendo la elección portuaria como un problema de multielección con interacción espacial. Para determinar los parámetros de la función que representa la partición de mercado de un puerto, se estimó un modelo Logit Condicional¹⁰ (McFadden, 1974) con datos históricos del origen de la carga transferida en cuatro puertos españoles. Este modelo, sin embargo, dado los supuestos que lo sustentan (esto es, que no haya correlación entre las alternativas), es probable que tenga resultados sesgados.

En Wang et al. (2016), en tanto, consideran la componente espacial en la elección, realizando un análisis probabilístico en que combinan un modelo de elección discreta con la ubicación geográfica de los exportadores. Así, desarrollan un algoritmo que determina el límite del hinterland a partir de las probabilidades de elección de rutas intermodales desde el origen al destino final, por parte de los exportadores. Sin embargo, dado que el modelo requiere una gran cantidad de información, se recurre a la utilización de datos simulados.

Tal vez el único modelo que incorpora explícitamente elementos espaciales y temporales es el de Garrido y Leva (2004). Allí, los autores presentan un Modelo Probit Multinomial (MNP, ver Garrido y Mahmassani, 2000), que considera correlaciones, espaciales, temporales y una matriz de varianza-covarianza general para el conjunto de elección. Este artículo es el precursor de esta tesis, ya que las características del proceso de elección portuaria que considera, son las que se desean incorporar en un modelo de más fácil calibración y aplicación.

En la siguiente sección se detalla el modelo del artículo señalado, con el fin de mostrar la complejidad que implica considerar en modelos de elección discreta las correlaciones espaciales y temporales del proceso de elección portuaria¹¹.

¹⁰ Este modelo se diferencia del Logit Multinomial en que las variables explicativas se refieren a atributos de las distintas alternativas, y no a características específicas de los individuos.

¹¹ En Anexos se presenta el artículo en su completitud.

2.2.4. Modelo Probit Multinomial Multidimensional Espacio-Tiempo

El modelo desarrollado por Garrido y Leva (2004), fue diseñado para modelar la elección de naviera y puerto de destino de exportadores frutícolas chilenos. Dada las características espaciales y temporales del proceso de elección, el modelo probit multinomial (MNP) utilizado consideró correlación serial, dependencia estadual y autocorrelación espacial.

El modelo se describe de la siguiente manera. Considérese un modelo de elección puerto-naviera con múltiples alternativas. Sea P el número de puertos, N el número de navieras y T la cantidad de períodos considerados. Así, el conjunto de elección espacio temporal $C = \{pnt\}$, tendrá $P \times N \times T$ elementos. Cada elemento de C posee atributos que dependen de un puerto específico, de una naviera específica, del par puerto - naviera y del intervalo de tiempo correspondiente.

Siguiendo la teoría de la utilidad aleatoria, una alternativa pnt será seleccionada si su utilidad es la máxima entre todas las alternativas disponibles.

La estructura de la utilidad se define como:

$$U(p, n, t) = V(p, n, t) + \delta(p, n, t) \quad (1)$$

donde $V(p, n, t)$ y $\delta(p, n, t)$, corresponden a la utilidad representativa y al término de error para la alternativa pnt , respectivamente.

La probabilidad de que la alternativa pn sea elegida en el período t , se determina como:

$$P(p, n, t) = P\{U(p, n, t) \geq U(q, m, t), \forall qm \in C\} \quad (2)$$

$$P(p, n, t) = P\{V(p, n, t) + \delta(p, n, t) \geq V(q, m, t) + \delta(q, m, t), \forall qm \in C\}$$

Para definir $V(p, n, t)$ y $\delta(p, n, t)$, se asume la notación utilizada por Williams y Ortúzar (1982), para modelos de elección multidimensionales, obteniendo:

$$V(p, n, t) = V_{pt} + V_{nt} + V_{pnt} \quad (3)$$

$$\delta(p, n, t) = \varepsilon_{pt} + \varphi_{nt} + \eta_{pnt} \quad (4)$$

donde V_{pt} , V_{nt} , V_{pnt} son las utilidades representativas específicas asociadas al puerto (p), naviera (n) y par puerto-naviera (pn) en el período t y ε_{pt} , φ_{nt} , η_{pnt} son los errores

estocásticos asociados a cada una de ellas en el mismo período. Estos errores se consideran independientes entre sí.

Los efectos espaciales y temporales presentes en el proceso, fueron incorporados dentro de la estructura general del error, considerando las siguientes características:

- La componente ε_{pt} asociada a un puerto específico durante un intervalo dado, es el resultado de un proceso dinámico estocástico representado con un modelo AR(1) (Autoregresivo de primer orden)¹² y está espacialmente correlacionada con el error de otros puertos.
- Las componentes φ_{nt} y η_{pnt} asociadas a una naviera y a un par puerto-naviera específico para un intervalo dado, son el resultado de procesos estocásticos independientes representados por modelos AR(1).

Así, para ε_{pt} , se consideró la estructura de error del modelo probit espacio – temporal desarrollada por Garrido (1997), siendo ésta la siguiente:

$$\varepsilon_t = \rho W \varepsilon_t + \mu_t \quad (5)$$

donde ρ es un parámetro espacial, μ_t representa un vector del proceso estocástico y W es la matriz de contigüidad de tamaño $P \times P$ cuyos elementos son:

$$w_{ij} = \begin{cases} 1, & \text{si } i \text{ y } j \text{ son puertos vecinos} \\ 0, & \text{en otro caso} \end{cases} \quad (6)$$

Al despejar el término de error se obtuvo:

$$\varepsilon_t = (I - \rho W)^{-1} \mu_t \quad (7)$$

siendo I la matriz identidad de dimensión $P \times P$ ¹³.

Por su parte el proceso estocástico μ_t lo representan a través de un modelo AR(1), de acuerdo a la siguiente expresión:

$$\mu_t = \alpha \mu_{t-1} + \lambda_t \quad (8)$$

¹² Ver McMillen (1992), Garrido (1997).

¹³ Nótese que para obtener la expresión (7) se requiere que la matriz $(I - \rho W)$ no sea singular.

donde α es un parámetro escalar y λ_t un vector aleatorio independientemente distribuido (i.d) tal que $\lambda_t \sim N(0, \Omega)$. Este vector representa un proceso de ruido blanco¹⁴, cuyas componentes poseen los siguientes momentos:

$$E(\lambda_{pt}) = 0$$

$$E(\lambda_{pt}\lambda_{qs}) = \begin{cases} \omega_p^2 & \text{si } p = q, t = s \\ 0 & \text{en otro caso} \end{cases} \quad (9)$$

$$\text{Es decir, } \Omega = \begin{bmatrix} \omega_1^2 & 0 & \cdots & 0 \\ 0 & \omega_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \omega_p^2 \end{bmatrix}.$$

Considerando el operador de retardo del tiempo L , tal que $L\mu_t = \mu_{t-1}$, se puede despejar μ_t de la ecuación (8), quedando:

$$\mu_t = C(L)\lambda_t, \quad C(L) = \left(\frac{1}{1 - \alpha L} \right) \quad (10)$$

Luego, expandiendo $C(L)$ en una serie geométrica infinita (considerando $|\alpha| < 1$), obtienen:

$$C(L) = \left(\frac{1}{1 - \alpha L} \right) = 1 + \alpha L + \alpha^2 L^2 + \cdots + \alpha^i L^i + \cdots$$

$$C(L) = \sum_{j=0}^{\infty} \alpha^j L^j \quad (11)$$

Así, μ_t , queda expresado como:

$$\mu_t = \sum_{j=0}^{\infty} \alpha^j \lambda_{t-j} \quad (12)$$

Reemplazando (12) en la ecuación (7), obtienen la expresión para el error específico asociado al puerto de destino, el cual considera correlación espacial y serial:

¹⁴ Una variable aleatoria λ_t es un proceso de ruido blanco, sí y sólo sí, los valores de las series de λ_t son independientes e idénticamente distribuidos, es decir $E(\lambda_t) = 0$ y $\text{Var}(\lambda_t) = \Omega$, lo cual implica que el primer y segundo momento son estables a través del tiempo.

$$\varepsilon_t = (I - \rho W)^{-1} \sum_{j=0}^{\infty} \alpha^j \lambda_{t-j} \quad (13)$$

En el caso de las características temporales de los errores específicos de la naviera (φ_{nt}) y del par puerto-naviera (η_{pnt}), fueron expresadas considerando que estos errores son procesos estocásticos representados con un modelo AR(1).

En el caso del error específico asociado a la naviera se tiene:

$$\varphi_t = \theta \varphi_{t-1} + \xi_t \quad (14)$$

siendo θ un parámetro escalar y ξ_t un vector aleatorio independientemente distribuido (i.d) tal que $\xi_t \sim N(0, \Sigma)$. Este vector representa un proceso de ruido blanco, cuyas componentes poseen los siguientes momentos:

$$\begin{aligned} E(\xi_{nt}) &= 0 \\ E(\xi_{nt} \xi_{ms}) &= \begin{cases} \sigma_n^2 & \text{si } n = m, t = s \\ 0 & \text{en otro caso} \end{cases} \end{aligned} \quad (15)$$

$$\text{es decir, } \Sigma = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_N^2 \end{bmatrix}.$$

Al igual que en el caso de la componente estocástica del error específico del puerto, consideraron el operador de retardo del tiempo L , tal que $L\varphi_t = \varphi_{t-1}$. Realizando los mismos pasos anteriores se puede despejar φ_t de la ecuación (14), quedando:

$$\varphi_t = \sum_{j=0}^{\infty} \theta^j \xi_{t-j} \quad (16)$$

Similarmemente, el error específico del par puerto-naviera también lo consideran como un proceso estocástico representado por un AR(1), tal que:

$$\eta_t = \phi \eta_{t-1} + \chi_t \quad (17)$$

con ϕ parámetro escalar y χ_t un vector aleatorio de ruido blanco independientemente distribuido (i.d) tal que $\chi_t \sim N(0, \Phi)$. Los momentos de sus componentes son:

$$E(\chi_{pnt}) = 0$$

$$E(\chi_{pnt}\chi_{qms}) = \begin{cases} \kappa_{pn}^2 & \text{si } p = q, n = m, t = s \\ 0 & \text{en otro caso} \end{cases} \quad (18)$$

$$\text{es decir, } \Phi = \begin{bmatrix} \kappa_{11}^2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \kappa_{12}^2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & 0 & \vdots & \vdots \\ \vdots & \vdots & 0 & \kappa_{np}^2 & 0 & 0 \\ 0 & 0 & \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \kappa_{NP}^2 \end{bmatrix}.$$

Aplicando el procedimiento equivalente, obtienen el error específico del par puerto – naviera, que considera correlación serial:

$$\eta_t = \sum_{j=0}^{\infty} \phi^j \chi_{t-j} \quad (19)$$

Conocidos ε_t , ϕ_t , η_t , el término error asociado con la elección del puerto p y naviera n en el período t , queda expresado como:

$$\delta(t) = (I - \rho W)^{-1} \sum_{j=0}^{\infty} \alpha^j \lambda_{t-j} + \sum_{j=0}^{\infty} \theta^j \xi_{t-j} + \sum_{j=0}^{\infty} \phi^j \chi_{t-j} \quad (20)$$

Las propiedades estadísticas de $\delta(t)$ dependerán de las funciones de densidad de los vectores de ruido blanco λ_t , ξ_t y χ_t , los cuales fueron supuestos normales.

Para determinar el primer y segundo momento de $\delta(t)$, los autores consideraron lo siguiente:

- Tomando la esperanza de la expresión (20), se tiene:

$$E(\delta(t)) = E\left((I - \rho W)^{-1} \sum_{j=0}^{\infty} \alpha^j \lambda_{t-j} + \sum_{j=0}^{\infty} \theta^j \xi_{t-j} + \sum_{j=0}^{\infty} \phi^j \chi_{t-j}\right)$$

$$= (I - \rho W)^{-1} \sum_{j=0}^{\infty} \alpha^j E(\lambda_{t-j}) + \sum_{j=0}^{\infty} \theta^j E(\xi_{t-j}) + \sum_{j=0}^{\infty} \phi^j E(\chi_{t-j}) \quad (21)$$

$$E(\delta(t)) = 0$$



- Dado que los errores ε_t , φ_t , η_t , se supusieron independientes entre sí, la matriz de covarianza de $\delta(t)$ estará determinada por la suma de sus matrices de covarianza.
- En el caso de ε_t , la matriz de covarianza en dos intervalos diferentes está dada por:

$$\begin{aligned} E(\varepsilon_t \varepsilon'_{t-r}) &= E \left[\left((I - \rho W)^{-1} \sum_{j=0}^{\infty} \alpha^j \lambda_{t-j} \right) \left((I - \rho W)^{-1} \sum_{j=0}^{\infty} \alpha^j \lambda_{t-j-r} \right)' \right] \\ &= (I - \rho W)^{-1} E \left[\sum_{j=0}^{\infty} \alpha^j \lambda_{t-j} \sum_{j=0}^{\infty} \alpha^j \lambda'_{t-j-r} \right] (I - \rho W)^{-1'} \end{aligned} \quad (22)$$

siendo r el tiempo de retardo entre dos intervalos.

Considerando que sólo la esperanza del producto de términos contemporáneos es distinta de cero¹⁵, la expresión (22) puede rescribirse como:

$$\begin{aligned} E(\varepsilon_t \varepsilon'_{t-r}) &= (I - \rho W)^{-1} \sum_{j=0}^{\infty} \alpha^{j+r} \alpha^j E[\lambda_{t-j-r}^2] (I - \rho W)^{-1'} \\ &= \sum_{j=0}^{\infty} \alpha^{j+r} \alpha^j (I - \rho W)^{-1} \Omega (I - \rho W)^{-1'} \end{aligned} \quad (23)$$

A su vez, el primer término de esta expresión $\sum_{j=0}^{\infty} \alpha^{j+r} \alpha^j$ es equivalente a:

$$\sum_{j=0}^{\infty} \alpha^{j+r} \alpha^j = \alpha^r \sum_{j=0}^{\infty} \alpha^{2j} = \frac{\alpha^r}{1 - \alpha^2} \quad (24)$$

Al reemplazarlo en (23) se obtiene:

$$E(\varepsilon_t \varepsilon'_{t-r}) = \frac{\alpha^r}{1 - \alpha^2} (I - \rho W)^{-1} \Omega (I - \rho W)^{-1'} \quad (25)$$

expresión que corresponde a la matriz de covarianza del error específico asociado al puerto, con $P + 2$ parámetros desconocidos.

- Efectuando un procedimiento similar al anterior, la matriz de covarianza de φ_t entre dos intervalos de tiempo distintos, estará dada por:

¹⁵ Esto se debe a que λ_t , es un proceso de ruido blanco, y por lo tanto la $E(\lambda_t \lambda'_{t-r}) = 0$.

$$\begin{aligned}
 E(\varphi_t \varphi'_{t-r}) &= E \left[\left(\sum_{j=0}^{\infty} \theta^j \xi_{t-j} \right) \left(\sum_{j=0}^{\infty} \theta^j \xi_{t-j-r} \right)' \right] \\
 &= \sum_{j=0}^{\infty} \theta^{j+r} \theta^j E[\xi_{t-j-r}^2] \\
 &= \sum_{j=0}^{\infty} \theta^{j+r} \theta^j \Sigma
 \end{aligned} \tag{26}$$

la cual puede reescribirse como:

$$E(\varphi_t \varphi'_{t-r}) = \frac{\theta^r}{1-\theta^2} \Sigma \tag{27}$$

que corresponde a la matriz de covarianza del error específico asociado a la naviera, con $N + 1$ parámetros desconocidos.

- Finalmente, la matriz de covarianza del error específico del par puerto-naviera, obtenida utilizando el mismo procedimiento, es:

$$E(\eta_t \eta'_{t-r}) = \frac{\phi^r}{1-\phi^2} \Phi \tag{28}$$

la cual posee $P \times N + 1$ parámetros desconocidos.

Conocidas las matrices de covarianza de ε_t , φ_t y η_t , el segundo momento de $\delta(t)$ queda determinado por:

$$\begin{aligned}
 E(\delta_t \delta'_{t-r}) &= E(\varepsilon_t \varepsilon'_{t-r}) + E(\varphi_t \varphi'_{t-r}) + E(\eta_t \eta'_{t-r}) \\
 &= \frac{\alpha^r}{1-\alpha^2} (I - \rho W)^{-1} \Omega (I - \rho W)^{-1'} + \frac{\theta^r}{1-\theta^2} \Sigma + \frac{\phi^r}{1-\phi^2} \Phi
 \end{aligned} \tag{29}$$

La dimensión de esta matriz¹⁶ es de $P \times N$ y el número de parámetros desconocidos que posee es igual a $(P + 2) + (N + 1) + (P \times N + 1)$. De estos parámetros sólo

$\left[\frac{(N \times P \times T - 1) N \times P \times T}{2} \right] - 1$ serán identificables (Bunch, 1991). Además, es importante

destacar que α , θ , ϕ y ρ , están restringidos a ser menores que uno para que se cumplan las condiciones de estacionaridad, en el caso de los parámetros temporales, y estabilidad del proceso espacial.

¹⁶ Dado que las matrices de covarianza de ε_t y φ_t tienen dimensiones menores, será necesario completarlas con filas y columnas de ceros, de forma tal que alcancen la dimensión $P \times N$.

Con respecto a la función de utilidad representativa presentada en la ecuación (3), los términos correspondientes a la utilidad representativa específica del puerto (V_{pt}) y de la naviera (V_{nt}), consideran la elección realizada en el período anterior, con el fin de capturar la posible inercia en el comportamiento de elección de puerto y naviera¹⁷.

De esta forma,

$$V_{pt} = \sum_{k=1}^K \beta_k X_{kpt} + \beta_{K+1} Y_{p(t-1)} \quad (30)$$

donde β es el vector de dimensión $K+1$; sus primeros K elementos corresponden a los parámetros de los atributos del puerto $p(X_{kpt})$ y el elemento $K+1$ es el parámetro de la variable que representa la elección en el período anterior, definida como:

$$Y_{p(t-1)} = \begin{cases} 1 & \text{si el puerto } p \text{ fue elegido en } (t-1) \\ 0 & \text{en otro caso} \end{cases} \quad (31)$$

El signo del parámetro β_{K+1} indica cómo se vería afectada la utilidad actual de un puerto que fue escogido en el período previo. Así, si β_{K+1} es positivo, la utilidad del puerto p será alta si éste fue escogido en el período previo. En caso contrario, es decir si β_{K+1} es negativo, la utilidad en el período actual será mayor cuando el mismo puerto no haya sido escogido en el período anterior.

La utilidad representativa relacionada con la elección de naviera, tiene una expresión similar a V_{pt} , dada por:

$$V_{nt} = \sum_{h=1}^H \gamma_h R_{hnt} + \gamma_{H+1} Z_{n(t-1)} \quad (32)$$

donde γ es un vector de dimensión $H+1$; sus primeros H elementos corresponden a los parámetros de los atributos de la naviera $n(R_{hnt})$ y el elemento $H+1$ es el parámetro de la variable que representa la elección en el período anterior. Esta variable se define como:

¹⁷ Train (2002) presenta ésta y otras formas de incorporar aspectos dinámicos del comportamiento de elección, dentro de la función de utilidad representativa

$$Z_{n(t-1)} = \begin{cases} 1 & \text{si la naviera } n \text{ fue elegida en } (t-1) \\ 0 & \text{en otro caso} \end{cases} \quad (33)$$

El signo de γ_{H+1} , indicará el efecto de la naviera escogida en el período anterior sobre la utilidad actual de dicha naviera.

Finalmente, la utilidad representativa asociada al par puerto – naviera es simplemente una función aditiva y lineal en los parámetros, definida como:

$$V_{pnt} = \sum_{g=1}^G v_g W_{gpnt} \quad (34)$$

donde, v es el vector de dimensión G de parámetros de los atributos asociados con el par puerto – naviera pn (W_{gpnt}).

De esta forma, la utilidad representativa $V(p,n,t)$ queda definida como:

$$\begin{aligned} V(p,n,t) &= V_{pt} + V_{nt} + V_{pnt} \\ &= \sum_{k=1}^K \beta_k X_{kpt} + \beta_{K+1} Y_{p(t-1)} + \sum_{h=1}^H \gamma_h V_{hpt} + \gamma_{H+1} Z_{n(t-1)} + \sum_{g=1}^G v_g W_{gpnt} \end{aligned} \quad (35)$$

que junto con la expresión (20) definen la función utilidad en el período t , del puerto p y naviera n , del modelo MNP espacio - tiempo.

El modelo recién presentado fue aplicado para determinar la elección de puerto de destino y naviera por parte de exportadores de uva de mesa chilena hacia EE.UU. La base de datos utilizada contenía información sobre las exportaciones realizadas entre los años 1997 y 2001. Los puertos sobre los que se realiza la elección son Filadelfia y Wilmington, ambos en la costa este de EE.UU., mientras que las navieras consideradas fueron CSAV y Lauritzen. Una segunda aplicación realizada por los autores, fue considerando una tercera alternativa de naviera denominada “otras compañías navieras”. Para ambos modelos se supusieron tres períodos de tiempo representados por las primeras semanas de febrero, marzo y abril, meses que representan el período punta de la temporada de exportación de uva de mesa. A continuación, se presenta la especificación de estos modelos.

En primer lugar, los autores definen la matriz de contigüidad como:

$$W = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad (36)$$

dada la cercanía entre los puertos escogidos (35 millas entre ellos). Por otra parte, las matrices de covarianza Ω , Σ y Φ fueron supuestas homocedásticas.

La matriz de covarianza del proceso en tanto, considera las 12 alternativas (2 puertos x 2 navieras x 3 períodos) resultando una estructura como la siguiente:

$$\begin{bmatrix} \frac{1}{(1-\alpha^2)}\Theta + \frac{1}{(1-\theta^2)}\Sigma + \frac{1}{(1-\phi^2)}\Phi & \frac{\alpha}{(1-\alpha^2)}\Theta + \frac{\theta}{(1-\theta^2)}\Sigma + \frac{\phi}{(1-\phi^2)}\Phi & \frac{\alpha^2}{(1-\alpha^2)}\Theta + \frac{\theta^2}{(1-\theta^2)}\Sigma + \frac{\phi^2}{(1-\phi^2)}\Phi \\ \frac{\alpha}{(1-\alpha^2)}\Theta + \frac{\theta}{(1-\theta^2)}\Sigma + \frac{\phi}{(1-\phi^2)}\Phi & \frac{1}{(1-\alpha^2)}\Theta + \frac{1}{(1-\theta^2)}\Sigma + \frac{1}{(1-\phi^2)}\Phi & \frac{\alpha}{(1-\alpha^2)}\Theta + \frac{\theta}{(1-\theta^2)}\Sigma + \frac{\phi}{(1-\phi^2)}\Phi \\ \frac{\alpha^2}{(1-\alpha^2)}\Theta + \frac{\theta^2}{(1-\theta^2)}\Sigma + \frac{\phi^2}{(1-\phi^2)}\Phi & \frac{\alpha}{(1-\alpha^2)}\Theta + \frac{\theta}{(1-\theta^2)}\Sigma + \frac{\phi}{(1-\phi^2)}\Phi & \frac{1}{(1-\alpha^2)}\Theta + \frac{1}{(1-\theta^2)}\Sigma + \frac{1}{(1-\phi^2)}\Phi \end{bmatrix}$$

donde Θ corresponde a $(I - \rho W)^{-1} \Omega (I - \rho W)^{-1}$, en la ecuación (29).

Los parámetros a estimar en esta matriz fueron α , θ , ϕ y ρ (todos restringidos a ser menores que uno) y, ω , σ y κ correspondientes a las desviaciones estándar de las matrices específicas asociadas a la elección de puerto, naviera y par puerto-naviera, respectivamente.

Una característica importante de esta estructura, es que los efectos espaciales y temporales dependen de sus respectivas matrices de covarianza. Por lo tanto, si se eliminan del modelo los parámetros correspondientes a las desviaciones estándar de alguna de estas matrices, se descartan directamente los parámetros de los efectos asociados a dicha matriz.

Con respecto a la función de utilidad representativa presentada en la ecuación (3), las variables consideradas fueron:

- Elección de puerto en el período anterior ($Y_{p(t-1)}$)
- Elección de naviera en el período anterior ($Z_{n(t-1)}$)
- Costo unitario de transporte (C)

Los autores no consideraron otro tipo de variables que explicaran la elección debido a la falta de información. Sin embargo, dado que la compleja estructura de matriz de

covarianza considera los distintos efectos posibles en el proceso, ésta debiera representar adecuadamente las variables no observadas.

Así, la especificación resultante de la función fue la siguiente:

$$\begin{aligned} V(p, n, t) &= V_{pt} + V_{nt} + V_{pnt} \\ &= \beta_1 Y_{p(t-1)} + \beta_2 Z_{n(t-1)} + \beta_3 C \end{aligned} \quad (37)$$

El parámetro β_3 del costo, debiese ser negativo e inferior en magnitud al resto de los parámetros de la función de utilidad representativa, ya que, de acuerdo a la información recogida en las diversas entrevistas sostenidas por los autores con agentes del proceso, esta variable tiene mucha menor relevancia en la elección, que la elección de puerto y naviera en períodos anteriores. A pesar de esto, los autores decidieron mantener el costo como variable ya que permite considerar la disponibilidad de las alternativas.

Considerando el análisis realizado de la información disponible (datos históricos y entrevistas), los dos parámetros restantes β_1 y β_2 , se esperaba fuesen significativos y de signo positivo, indicando una tendencia a repetir elecciones pasadas.

En relación a los parámetros de la matriz de covarianza, el efecto de correlación temporal y espacial de la elección portuaria, debiese quedar reflejado en los parámetros asociados al efecto temporal, es decir α , θ , ϕ y en el parámetro espacial ρ , los cuales debiesen ser positivos. Lo anterior, debido a la cercanía y características análogas de los puertos considerados.

En cuanto a las varianzas de las matrices asociadas específicamente a los puertos, navieras y pares puerto-naviera, se esperaría que la magnitud de ω , σ_1 y σ_2 fuera menor que la de κ , producto de la falta de atributos significativos que representen al par puerto-naviera. Esto produciría que el término de error asociado al par de elección, sea mayor que el de elección de puerto y de naviera en forma separada.

Este primer ejemplo considerado por los autores, requirió la estimación de once parámetros, ocho de la matriz de covarianza y tres provenientes de la función de utilidad representativa. El total de observaciones con que contaban los autores para realizar la estimación, fue de 616 observaciones.



El segundo ejemplo, en tanto, consideró una tercera línea naviera, siendo la estructura del modelo similar al ejemplo anterior. Sólo para poder distinguir el efecto que tiene la inclusión de la nueva alternativa, se consideró Σ con varianzas distintas para cada una de las navieras.

La matriz de covarianza por su parte, tuvo una estructura similar, pero de mayor tamaño ya que el número de alternativas aumentó de 12 a 18. El total de parámetros a estimar en este caso fue doce con un total de 830 observaciones.

En la calibración del modelo, los autores señalan que obtuvieron múltiples soluciones de acuerdo al punto de partida escogido, debido a la no linealidad de la función de verosimilitud. Entre todas las soluciones seleccionaron la que presentaba mejores características, ya sea en cuanto a convergencia, test t y log verosimilitud.

De los parámetros estimados se destacaron en magnitud y significancia los asociados a las variables que representan la elección en el período anterior de puerto (β_1) y naviera (β_2). El signo positivo de estos parámetros, corroboró la existencia de una inercia en el comportamiento de elección, tal como suponían los autores en la estructura del modelo. Por otra parte, los parámetros temporales asociados a las matrices de covarianza del puerto (α), naviera (θ) y par puerto-naviera (ϕ) dejaron en evidencia la presencia de correlación serial, ya que estos parámetros fueron positivos, significativos y alcanzaron altos valores en ambos modelos.

El parámetro espacial (ρ) en tanto, alcanzó un mayor valor y significancia en el primer modelo. Los autores señalan que el signo positivo obtenido en este parámetro, puede explicarse por la cercanía y características análogas de los puertos considerados, que hacen que los factores espaciales externos (no considerados en la utilidad representativa) produzcan efectos similares en ambos.

El resto de los parámetros, a excepción del asociado al costo de transporte, fueron todos significativos en el primer modelo, disminuyendo esta significancia en las desviaciones estándar de las matrices de puerto (ω) y naviera (σ) en el segundo modelo. La menor magnitud y significancia de los parámetros ω y σ , con respecto a la desviación estándar par puerto-naviera (κ) en ambos modelos, se explica por la falta de atributos significativos que representen al par puerto-naviera. Esto produciría que el término de



error asociado al par de elección, sea mayor que el de elección de puerto y de naviera en forma separada.

Por otra parte, el signo, magnitud y significancia del parámetro β_3 del costo, es el esperado, ya que, de acuerdo a la información recogida de las diversas entrevistas sostenidas con agentes del proceso, esta variable no tendría un efecto importante en el proceso de elección en estudio.

Las log- verosimilitudes promedio, en tanto, prácticamente no varió entre ambos modelos.

Si bien el modelo probit recién presentado entrega buenos resultados y representa correctamente la interacción y decisiones de agentes, su aplicación es altamente compleja debido a las dificultades matemáticas de especificación, calibración y estimación de probabilidades de elección, haciéndolo poco transferible a la práctica de no-expertos en modelización.

Por este motivo, esta tesis busca definir un modelo de mayor simpleza en calibración y aplicación, pero que incorpore todos los efectos antes señalados.

2.4. Conclusiones

Tras la revisión realizada del proceso de elección portuaria, se puede señalar, que respecto al agente tomador de decisión, no existe un consenso en quien es el que posee mayor influencia. Esto dependerá de las condiciones que ocurran al momento de la elección y de múltiples factores tales como: tipo de producto, tipo y duración de contratos, tipo de movimiento (importación o exportación), tamaño del exportador/importador, entre otros. Sin embargo, se destacan a navieras, freight forwarders y exportadores, como los principales actores del proceso, siendo los últimos los más mencionados en la literatura. Se reconoce, además, que todos estos actores se interrelacionan, generando un proceso que se retroalimenta, y que por lo tanto dificulta su modelización.

Lo anterior implica un desafío en términos de definir una estructura de modelización que recoja las relaciones del sistema, y que además sea de fácil calibración y aplicación.



Respecto a los principales factores considerados en el proceso de elección portuaria, se pudo observar que había algunos de ellos que eran independientes del agente tomador de decisión. Entre estos se encontraba la variable localización (entendida como distancia, tiempo o costo de viaje), la cual se explica tanto por el efecto que tiene en los exportadores en términos de costos de transporte, como en las navieras, por el impacto que tiene en el tamaño del hinterland al cual pueden acceder. Esta variable fue mencionada en todos los artículos analizados.

Otros factores que se mencionaron con recurrencia, fueron: servicios navieros disponibles, intermodalidad, costos marítimos, costos portuarios, congestión, infraestructura, eficiencia y servicios portuarios.

Además de las variables mencionadas, hay dos características del proceso de elección que fueron reconocidas en distintas publicaciones. Estas son, los efectos espaciales y temporales presentes en la decisión. Los primeros se manifiestan tanto por la cercanía espacial de los puertos competidores, como de los exportadores que comparten un hinterland. El efecto temporal, en tanto, se observa en la inercia que puede tener el proceso de elección producto, por ejemplo, del tipo de contrato entre exportadores y navieras, o de la rigidez de la estructura de la cadena logística asociada a la mercancía movilizada.

En cuanto a las metodologías de modelización, se pudo distinguir variados enfoques, siendo los modelos de elección discreta los más recurrentes. Le siguen el método de Proceso Analítico Jerárquico, el cual ha sido ampliamente utilizado para escoger y ponderar los principales factores de elección portuaria.

Sin embargo, la gran mayoría de estas investigaciones no se hace cargo de las características espaciales y temporales antes señaladas, así como tampoco del comportamiento heterocedástico de las alternativas del conjunto de elección. Es probable que la complejidad metodológica que implica considerar estos efectos, haya sido una de las causas que generaron su omisión en los artículos revisados.

Hay algunas investigaciones que han intentado incorporar estos efectos o alguno de ellos, utilizando para esto modelos de elección discreta. Sin embargo, las



especificaciones presentadas no se hacen cargo adecuadamente de las correlaciones y por lo tanto, es altamente probable que sus resultados se encuentren sesgados.

Sólo se encontró un modelo que incorporaba explícitamente efectos espaciales (representados a través de correlación serial y dependencia estadual), temporales (incorporados como dependencia estadual) y una matriz de varianza-covarianza general para el conjunto de elección. Este modelo, correspondiente a un Modelo Probit Multinomial, fue desarrollado por Garrido y Leva (2004). Si bien el modelo entregó buenos resultados y representó correctamente las características del proceso de elección, su aplicación resultó altamente compleja debido a las dificultades matemáticas de especificación, calibración y estimación de probabilidades de elección, haciéndolo poco transferible a la práctica de no-expertos en modelización.

Sin embargo, dada la completitud del modelo presentado en la investigación de Garrido y Leva (2004), éste fue considerado como referencia de comparación con el modelo desarrollado en esta tesis.

A modo de resumen, se presenta en la Tabla 4, la bibliografía revisada, indicando para cada artículo, el actor tomador de decisión, los factores de elección considerados y la metodología de modelización utilizada.



Autor	Actor tomador de decisión	Factores de elección ¹⁸	Metodología aplicada
Slack (1985)	- Exportadores - Freight Forwarders	- Costos de transporte terrestre - Localización - Congestión portuaria - Frecuencia navieras - Intermodalidad - Costos portuarios - Infraestructura	Ranking de importancia
Tiwari et al. (2003)	Exportadores	- Distancia - Frecuencia navieras - Congestión portuaria (medida como Teus transferidos) - Número de líneas de atraque - Tamaño de flota naviera	Modelo de elección discreta- Logit Multinomial
Nir et al. (2003)	Exportadores	- Tiempo de viaje (instalación exportador/puerto) - Costos de viaje (instalación exportadora/puerto) - Número de rutas navieras disponibles - Frecuencia semanal de navieras	Modelo de elección discreta- Logit Multinomial
Song y Yeo (2004)	- Armadores - Exportadores - Navieras - Operadores de terminales - Académicos e investigadores	- Volumen de carga - Infraestructura - Localización - Nivel de servicio - Costos portuarios	Proceso Analítico Jerárquico

¹⁸ El orden no representa prioridad de factores

Autor	Actor tomador de decisión	Factores de elección ¹⁸	Metodología aplicada
Garrido y Leva (2004)	Exportadores	- Costos de transporte - Elección portuaria de períodos pasados - Elección de navieras de períodos pasados	Modelo de elección discreta- Probit Multinomial
Malchow y Kanafani (2004)	Exportadores	- Distancia - Frecuencia navieras	Modelo de elección discreta- Logit Multinomial
Lirn et al. (2004)	- Navieras - Operadores portuarios	- Costos portuarios - Proximidad a las principales rutas de navegación - Cercanía a importadores/exportadores - Infraestructura - Conexión con alimentadores	Proceso Analítico Jerárquico
Guy y Urli (2006)	Navieras	- Calidad Infraestructura - Costos portuarios - Servicios portuarios - Localización	Análisis Multicriterio
Ugboma et al. (2006)	- Exportadores - Freight Forwarders	- Localización - Eficiencia - Frecuencia naviera - Infraestructura - Buena reputación en daños y demoras - Calidad de servicio portuario - Costos portuarios	Proceso Analítico Jerárquico



Autor	Actor tomador de decisión	Factores de elección ¹⁸	Metodología aplicada
Tongzon y Sawant (2007)	Navieras	– Eficiencia portuaria – Localización – Infraestructura – Tarifas portuarias – Conectividad – Volumen de carga – Servicios portuarios	Regresión logística binaria
De Langen (2007)	– Exportadores – Freight Forwarders	– Localización – Características del producto – Número de rutas y servicios navieros – Calidad de las compañías que operan el terminal – Eficiencia en el manejo de la carga – Calidad del equipamiento – Calidad de los servicios de transporte – Información de servicios en el puerto – Buena reputación en daños y demoras – Foco en el cliente – Modos de conexión con el hinterland – Contacto personalizado	Ranking de importancia
Chang et al. (2008)	Navieras	– Volumen de carga local – Tarifa del terminal – Disponibilidad de muelle – Localización – Volumen transferido – Conexiones de alimentadores	Análisis exploratorio de factores y análisis factorial confirmatorio



Autor	Actor tomador de decisión	Factores de elección ¹⁸	Metodología aplicada
Wiegmans et al. (2008)	Navieras	– Localización – Eficiencia portuaria – Reputación – Calidad del servicio – Intermodalidad – Tarifas portuarias – infraestructura	Ranking de importancia
Tongzon (2009)	Freight forwarders	– Localización – Frecuencia naviera – Infraestructura – Eficiencia portuaria – Tarifas portuarias – Respuesta rápida a las necesidades de los usuarios portuarios – Reputación del puerto respecto al daño de la carga	Ranking de importancia y regresión lineal
García-Alonso y Sánchez-Soriano (2009)	Exportadores	– Distancia – Constante de atractividad portuaria	Modelo de elección discreta- Logit Condicional
Aronietis et al. (2010)	– Navieras – Operadores de terminales – Exportadores – Freight forwarders	– Costos portuarios – Conexión con el hinterland – Capacidad portuaria – Confiabilidad – Localización – Carga disponible – Flexibilidad – Servicio al cliente – Frecuencia – Riesgo de pérdida/daño de la carga	Ranking de importancia



Autor	Actor tomador de decisión	Factores de elección ¹⁸	Metodología aplicada
Lee et al. (2010)	– Navieras – Exportadores – Operadores de terminales de contenedores	– Localización – Accesibilidad portuaria – Atributos del puerto – Condiciones de operación portuarias – Características del hinterland – Características de las navieras – Otros	Encuestas/Análisis exploratorio de factores
Veldman et al. (2011)	Exportadores/importadores	– Costos de transporte terrestre – Costos de transporte marítimo – Efecto Mohring¹⁹	Modelo de elección discreta-Logit Multinomial
Steven y Corsi (2012)	Exportadores	– Productividad de la grúa – Congestión portuaria – Administración del puerto (pública o privada) – Tamaño promedio de naves – Frecuencia naviera – Costos portuarios – Tiempo de tránsito marítimo – Tiempo de tránsito terrestre – Tamaño del exportador – Número de líneas de atraque	Modelo de elección discreta-Logit Condicional

¹⁹ Ver nota al pie N° 1.



Autor	Actor tomador de decisión	Factores de elección ¹⁸	Metodología aplicada
Veldman et al. (2013)	Exportadores/ importadores	– Costos de transporte terrestre – Costos de transporte marítimo – Efecto Mohring²⁰ – Desbalance de contenedores de importación y exportación	Modelo de elección discreta - Logit Multinomial y Logit Jerárquico
Ng et al. (2013)	Exportadores	– Distancia – Costos de transporte terrestre – Eficiencia portuaria – Costos portuarios – Calidad de servicio portuario	Ranking de importancia
Vega et al. (2014)	Exportadores	– Costo de acceso al puerto – Tarifa marítima – Frecuencia naviera – Tiempo de tránsito marítimo	Modelo de elección discreta - Logit Jerárquico
Kramberger et al. (2015)	Exportadores	– Distancia – Tiempo de transporte – Tasa de preferencia	Proceso Analítico Jerárquico- Programación Matemática mixta

²⁰ Ver nota al pie N° 1.



Autor	Actor tomador de decisión	Factores de elección ¹⁸	Metodología aplicada
Veldman et al (2015)	Exportadores/ importadores	– Distancia terrestre – Índice portuario (factor de carga y velocidad de transferencia portuaria) – Índice marítimo (tamaño de la nave, distancia marítima) – Efecto Mohring²¹ – Dummy por servicio feeder (1 si se necesita servicio feeder para el destino final, 0 sino)	Modelo de elección discreta - Logit Multinomial
Halim et al. (2016)	-	– Costos logísticos totales – Nivel de servicio de la red	Programación matemática multiobjetivo
Song et al. (2016)	Navieras	– Costos portuarios – Costos de transporte naviero – Costos de transporte en el hinterland – Costos de servicios feeder	Teoría de juegos
Wang et al. (2016)	Exportadores	– Costos de transporte de rutas intermodales	Análisis probabilístico y modelo de elección discreta

Tabla 4: Actores, factores de elección portuaria y metodologías de modelización

²¹ Ver nota al pie N° 1.



Capítulo 3

DISEÑO DE AUTÓMATA CELULAR PARA ELECCIÓN PORTUARIA



3. DISEÑO DE AUTÓMATA CELULAR PARA ELECCIÓN PORTUARIA

3.1. Modelos de Autómatas Celulares

Un Autómata Celular (AC) es un modelo matemático (usualmente vinculado al campo de la física computacional), propuesto inicialmente por John von Neumann en la década del 1950, quien no concluye la escritura de su obra “Theory of Self-reproducing Automata”. La publicación de su libro es editada y completada por A. W. Burks en 1966 (ver Von Neuman, J. y Burks, A. W., 1966).

Un AC está compuesto por 4 componentes: unidades llamadas “células”, un número discreto de estados (por ejemplo, activo o inactivo), una vecindad, y reglas de transición. Las células interactúan localmente, y su estado siguiente depende de su estado actual y el de sus vecinos. El AC usualmente opera en una grilla o en una forma conocida. El conjunto de células forma un sistema, que evoluciona de manera compleja.

Este enfoque de modelización es especialmente adecuado para procesos donde una célula se ve influenciada por sus vecinos inmediatos, como los sistemas urbanos. Ha sido aplicado exitosamente en la modelización de sistemas complejos, autorganizados y emergentes.

A finales de los años 60, se inicia una segunda fase de estudio, en donde John Horton Conway, propone el autómata celular conocido como *El Juego de la Vida*, (Gardner,

1970), llamado así por su relación con la modelización de sistemas biológicos, tales como: comportamientos colectivos de seres vivos, ecosistemas, entre otros. El desarrollo de este juego depende del estado inicial de un conjunto de células localizadas en una grilla ortogonal bidimensional infinita y de las normas del juego.

Las condiciones del estado futuro de las células dependerán solo de las condiciones del estado anterior. Cada célula interactúa con sus ocho vecinos (células que están adyacentes horizontal, vertical y diagonalmente), pudiendo vivir, morir o reproducirse, en función del estado de sus vecinos. El sistema formado por ellas puede adoptar numerosos patrones durante el juego, el cual dependerá fuertemente de las condiciones iniciales.

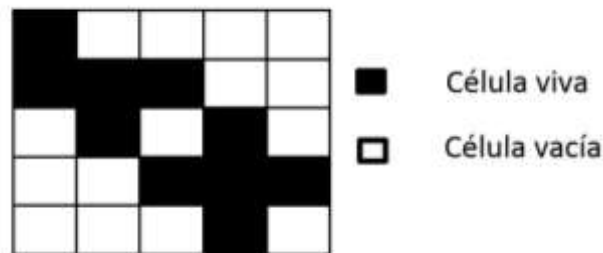


Figura 2: Esquema conceptual de un AC en dos dimensiones.

En cada período de tiempo, el estado de cada célula puede cambiar, de acuerdo a las siguientes reglas:

- **Nacimiento.** Una célula muerta en el tiempo t pasa a vivir en el tiempo $t + 1$ si tiene exactamente tres de sus vecinos vivos en el tiempo t .
- **Muerte por sobre-población.** Si una célula vive en el tiempo t y tiene cuatro o más vecinos vivos en ese período, esta célula morirá en el tiempo $t + 1$.
- **Muerte por aislamiento.** Si una célula vive en el tiempo t y tiene un vecino vivo o ninguno en el mismo período, esta célula morirá en el tiempo $t + 1$.
- **Sobrevivencia.** Una célula que vive en el tiempo t permanecerá viva en el tiempo $t + 1$, sí y solo sí, tiene dos o tres vecinos vivos en el mismo período.

Si la célula muere, su celda correspondiente en la grilla queda vacía.

Tal como se señaló, el jugador es quien determina las condiciones iniciales de las células, creando así una “población inicial” o “generación cero”. Al aplicar las reglas señaladas a



todas las células en un período dado, se crea una nueva generación, ocurriendo todos los nacimientos y muertes simultáneamente.

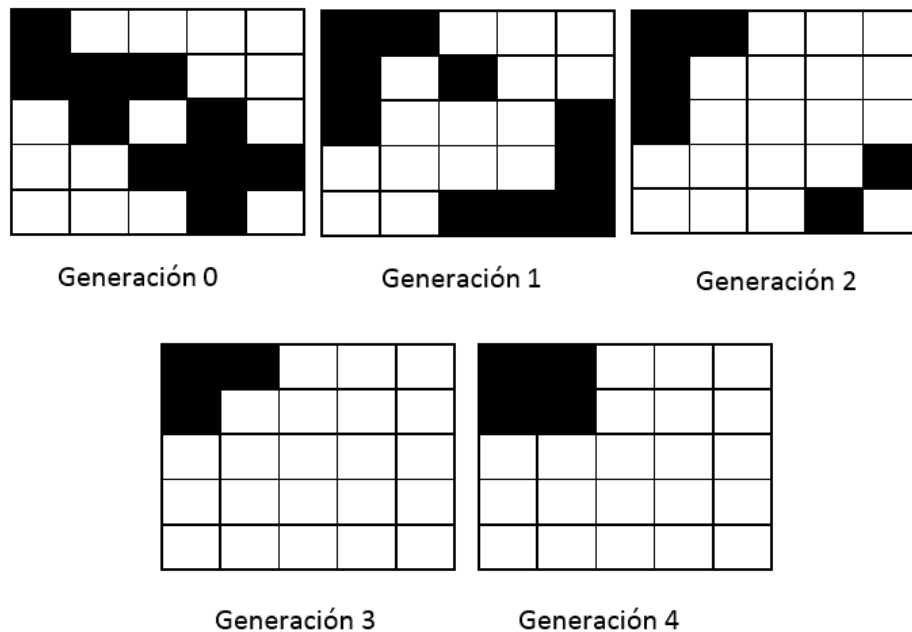


Figura 3: Ejemplo de evolución AC según el juego de la vida

El ejemplo de *El Juego de la Vida*, tiene todos los ingredientes que contiene un AC: un conjunto de células homogéneas, un conjunto de estados de estas células, un arreglo espacial, una definición de vecindad y las reglas de transición de estados. Estos ingredientes proveen así, una herramienta muy flexible para la modelización de fenómenos dinámicos, no lineales y complejos (ver por ejemplo Griffeath y Moore, 2003; Chopard y Droz, 2005; Goles y Martínez, 1998).

Sabemos que el proceso de elección portuaria es dinámico y no lineal, y además que cuenta con las siguientes características que definen un sistema complejo:

- Hay muchos actores que se relacionan entre sí con reglas simples. Sin embargo, el resultado de su interacción es difícil de predecir.
- Las interacciones entre los actores no pueden explicarse a partir del comportamiento aislado de cada actor. Esto es conocido como propiedades emergentes.
- No existe una regla única de elección.

- El sistema no es reductible a un solo nivel de descripción, es decir, siempre se requiere más de una visión para describirlo.
- Ningún actor conoce todas las características del sistema en un instante dado.

De acuerdo a lo anterior y considerando la dinámica espacio temporal observada en el proceso de elección portuaria, la teoría de AC parece especialmente adecuada para tratarlo. Su aplicación permitiría especificar modelos computacionalmente eficientes y menos intensivos en datos.

En la exhaustiva revisión bibliográfica realizada, no se encontró literatura especializada en que se aplicaran AC como herramienta de modelización de elección portuaria. Sin embargo, los AC han sido ampliamente aplicados en estudios de dinámica espacial, como por ejemplo, en el análisis del crecimiento urbano a nivel local y regional (ver Barredo et al., 2003; Aguilera, F., 2006; Li et al., 2007; Almeida et al., 2008; Santé et al. 2010).

Este análisis de comportamiento del sistema territorial, es parte de lo que se desea analizar como resultado de la modelización de elección portuaria. Con este enfoque es posible, por ejemplo, estudiar la dinámica del hinterland de un puerto frente a cambios en las condiciones del sistema.

A continuación, se presenta la estructura y características del AC propuesto para abordar la modelización del proceso de elección portuaria.

3.2. Estructura y Características del Autómata Celular Probabilístico de Elección Portuaria (ACPEP)

El problema que se busca resolver es el de la predicción de elección de puerto por un conjunto de exportadores que se extienden dentro de una región acotada. Estos exportadores, comparten una porción significativa del hinterland. Asumimos que hay m puertos compitiendo por el hinterland y todos estos puertos están a disposición de los exportadores. Los puertos disponibles se encuentran en una red que los conecta con todos sus potenciales clientes (exportadores). Dividimos esta región en n zonas de generación de carga (ZGC) mutuamente excluyentes y colectivamente exhaustivas.

Cada ZGC puede tener $m+1$ estados de activación: una por cada puerto disponible y un estado de inactividad. Así, si i -ésima ZGC “ha elegido” el puerto j en el período t , entonces su estado actual es $e_{it} = j$, con $j = 0, 1, 2, \dots, m$. 0 significa que la i -ésima ZGC esta inactiva (no elige ningún puerto) y j que elige el puerto j .

La Figura 4 muestra un esquema de las ZGC y puertos en que opera el modelo de Autómata Celular Probabilístico de Elección Portuaria (ACPEP). Cada círculo representa un ZGC, correspondiente a su vez a una célula del ACPEP, que puede estar activa o inactiva. Una ZGC estará activa en el período t si utiliza para exportar uno de los puertos disponibles. El color indica el puerto escogido por la ZGC en el tiempo t .

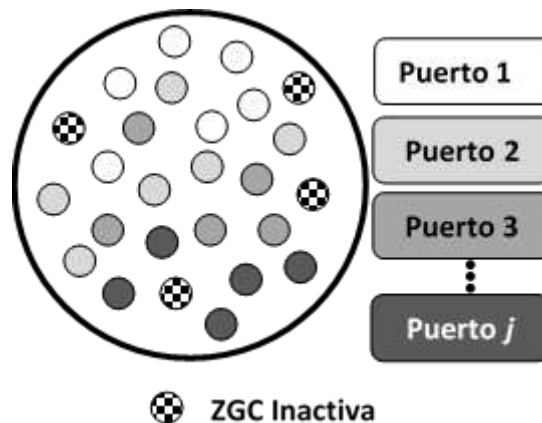


Figura 4: Esquema conceptual de las zonas y puertos en el ACP

La condición de actividad de la i -ésima ZGC en el tiempo t dependerá dos aspectos: las características operacionales y la interacción espacial. Las características operacionales son atributos que conforman la utilidad derivada de la elección de un puerto dado. Ejemplos de estos atributos son la distancia a cada puerto disponible, tarifas, conectividad, capacidad de transferencia, contratos de largo plazo, capacidad negociadora, etc. En el caso de interacción espacial, el nuevo estado de la ZGC, dependerá de su situación actual y de los estados de otras ZGC espacialmente correlacionadas, es decir pertenecientes a su “vecindad” (ver sección 3.2.1).

Por otra parte, dado lo compleja de la decisión, un modelador no puede describir completamente el proceso de elección portuaria sólo con un conjunto de atributos (por ejemplo, conectividad, tarifas, tiempos de espera, etc.), ya que es probable que desconozca diversos aspectos involucrados en esta decisión.

Para evitar este problema, incorporamos una componente estocástica, que explicaría (similar a la hipótesis de la teoría de la utilidad aleatoria estándar), al menos dos fenómenos: el modelador no tiene información perfecta (desconoce con certeza los atributos de cada puerto en un tiempo dado), y una ZGC puede activarse/desactivarse por variables exógenas que no están incluidas en la especificación del modelo. Esta componente estocástica no implica que el puerto sea elegido al azar por los exportadores, sino que indica que, en ausencia de información perfecta por parte del modelador, la elección de puerto no puede determinarse de forma determinística por un conjunto incompleto de atributos disponibles para el modelador. Esta hipótesis explicaría por lo menos dos fenómenos que de otra manera podrían parecer erróneos desde la perspectiva del modelador: en primer lugar, un exportador puede elegir un puerto que no sea su mejor elección según los atributos medidos. En segundo lugar, dos exportadores "idénticos" (desde el punto de vista de los atributos) pueden elegir diferentes puertos. Ambos fenómenos pueden ser descritos como decisiones racionales, cuando el modelador conoce sólo una parte de los atributos que activan el proceso de selección de puertos.

3.2.1 Vecindad

Hay varias formas de definir e interpretar el concepto de vecindad de un AC, el que se traduce en la definición a priori de la influencia espacial entre las células dentro de la región en estudio. Dos de los ejemplos más comúnmente encontrados de vecindad son las definidas por Von Neumann y la de Moore. El primero considera una grilla rectangular, en que las células vecinas son aquellas que comparten un borde común (cuatro vecinos en una grilla cuadrada). El segundo considera como un vecino, cada una de las ocho celdas que rodean una célula dada en una grilla cuadrada (ver ejemplo en sección 3.1).

En el caso de elección portuaria, la ZGC no está necesariamente ordenada en una matriz métrica, por lo que el concepto de grilla regular no se ajusta al patrón espacial real. En lugar de ello, incorporamos el concepto de zona radial, en la que dos células son vecinas si ambas están circunscritas en un círculo de radio R , tal como se muestra en la Figura 5.

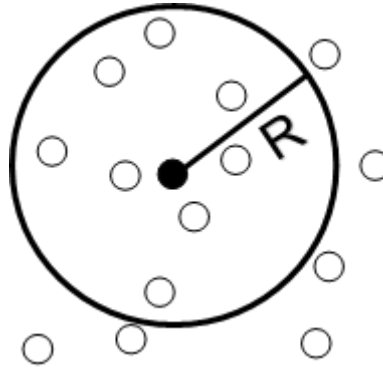


Figura 5: Esquema de vecindad definida en el ACPEP

3.2.2 Función de Actualización

Tal como se señaló en la sección 3.1, un AC debe contener un conjunto de células homogéneas, en este caso las ZGC, un conjunto de estados de estas células (e_{it}), un arreglo espacial (hinterland), una definición de vecindad y las reglas de transición de estados.

Dada las características del proceso de elección portuaria y considerando la necesidad de incorporar aleatoriedad por los motivos ya señalados, en el ACPEP cada célula correspondiente a una ZGC transitará del estado actual al siguiente (en pasos de tiempo discreto) de acuerdo a una función de transición estocástica.

Incorporando estocasticidad en la interacción espacial entre células, incluso bajo simples reglas de actualización, estos modelos son capaces de identificar el comportamiento complejo que puede surgir desde el sistema.

Así, el ACPEP actualizará el estado de cada ZGC, siguiendo una regla estocástica, tal como se señala a continuación:

1. Cada ZGC tiene una probabilidad μ de activación. Esta probabilidad es un número real comprendido entre $0 \leq \mu \leq 1$, que ha sido calibrado por un algoritmo genético (ver sección 3.2.3). Para imponer la activación con probabilidad μ , en cada paso del ACPEP, se obtiene un número real z obtenido de una distribución uniforme (0,1), el cual se compara con μ . Si $z \leq \mu$ entonces la ZGC se activa o permanece activa, adoptando el estado correspondiente al paso 4 de esta regla.

2. Cada ZGC i tiene una función $\Phi^i \in \mathbb{R}^m$ cuyo dominio es la vecindad de la ZGC i que se está actualizando. La función, que entrega un vector de m componentes, es la siguiente:

$$\Phi^i = \alpha \Phi_1^i + \beta(d_{i1}, d_{i2}, \dots, d_{im}) + (\gamma_1, \gamma_2, \dots, \gamma_m) \quad \forall i = 1, 2, \dots, n \quad (38)$$

donde d_{ij} representa la distancia desde la ZGC i al puerto j (con $j=1,2,\dots,m$), y el parámetro constante γ_j , representa otros aspectos del atractivo de un puerto que no han sido considerados explícitamente en la especificación de la ecuación (38). Note que por razones de identificabilidad, solo pueden ser estimados $(m-1)$ parámetros constantes γ_j , debiendo dejarse arbitrariamente fijo en uno de los puertos.

Φ_1^i es un vector m -dimensional, formado por la razón entre el número de ZGC dentro del círculo de radio R de la vecindad (R -vecindad) que escogen el puerto j , respecto del total de ZGC activas dentro de la vecindad, esto es:

$$\Phi_1^i = \left(\frac{n_1}{N_i}, \frac{n_2}{N_i}, \dots, \frac{n_m}{N_i} \right) \quad \forall i = 1, 2, \dots, n \quad (39)$$

en que $n_1, n_2, \dots, n_m$ es el número de ZGC activas vecinas que escogieron el puerto 1 a m , respectivamente en el período t , y N_i es el total de ZGC activas dentro de la R -vecindad de la ZGC i .

De acuerdo a la especificación de la ecuación (38), el parámetro α representa la influencia del comportamiento de los vecinos en la elección portuaria, es decir, la componente espacial de la decisión asociada al estado de las ZGC vecinas, y el parámetro β la influencia de la distancia en el proceso de elección. En la calibración, β debiese tomar un valor negativo, ya que representa una desutilidad (a mayor distancia menor utilidad). α , en tanto, debiese tomar un valor positivo, debido a que implica el peso del proxy de participación de mercado.

3. Cada componente de la función Φ^i representa la utilidad o atractivo que genera en la ZGC i la elección del puerto j . Estas componentes son mapeadas de la siguiente manera: a su mayor componente se le asigna una probabilidad p , a la segunda mayor, una probabilidad $p(1-p)$, y así sucesivamente decayendo como ley de potencia hasta $p(1-p)^{m-1}$. Este mapeo evita que el puerto con mayor participación de

mercado y menor distancia sea siempre elegido. Sin embargo, se le asigna una mayor probabilidad de elección al puerto con mayor utilidad.

4. El estado e_i de cada ZGC se actualiza con un valor entre 1 y m , representando la elección correspondiente a la mayor componente de la función Φ^i mapeada.

3.2.3 Algoritmo de Calibración

El ACPEP descrito anteriormente necesita calibrar los siguientes parámetros:

- R : El radio de la vecindad para todas la ZGC.
- α : El peso de los vecinos en la decisión de elección portuaria.
- β : El peso de la distancia en la decisión de elección portuaria.
- γ : Componente determinística de preferencia
- μ : El vector de probabilidad de activación de la ZGC.

Hay varias maneras de realizar esta calibración, sin embargo, dado que los algoritmos genéticos²² (Goldberg, 1989) han tenido un muy buen rendimiento en el campo de AC (ver por ejemplo Mitchell et al., 1996; Breukelaar y Bäck, 2005; Li et al., 2007; Vázquez y Oliver, 2008), se decidió utilizarlos como herramienta de calibración.

El algoritmo es el siguiente:

- Se define como condición inicial una población de 100 individuos. Cada uno de estos individuos representa una simulación completa para todo el horizonte de modelado.
- El genotipo de cada individuo corresponde a los valores de los parámetros del modelo, escogidos al azar en la primera generación. A cada parámetro desconocido se le da un valor aleatorio inicial. Por ejemplo, para m puertos y n ZGC, el genotipo de los individuos sería:

²² Los algoritmos genéticos son métodos de optimización que se inspiran en la evolución biológica y en su base genético-molecular. Estos algoritmos hacen evolucionar una población de individuos sometiéndola a mutaciones y recombinaciones genéticas, seleccionando con algún criterio, a los individuos más adaptados. Estos individuos sobreviven y se reproducen, mientras que los menos aptos son descartados.

$$\begin{aligned}
 I_1 &= (R_1, \alpha_1, \beta_1, \gamma_{1-1}, \gamma_{1-2}, \dots, \gamma_{1-(m-1)}, \mu_{1-1}, \mu_{1-2}, \dots, \mu_{1-n}) \\
 I_2 &= (R_2, \alpha_2, \beta_2, \gamma_{2-1}, \gamma_{2-2}, \dots, \gamma_{2-(m-1)}, \mu_{2-1}, \mu_{2-2}, \dots, \mu_{2-n}) \\
 &\vdots
 \end{aligned} \tag{40}$$

$$I_{100} = (R_{100}, \alpha_{100}, \beta_{100}, \gamma_{100-1}, \gamma_{100-2}, \dots, \gamma_{100-(m-1)}, \mu_{100-1}, \mu_{100-2}, \dots, \mu_{100-n})$$

- El fenotipo asociado a cada individuo, corresponde a los observables involucrados en la dinámica del proceso estudiado: cantidad de ZGC activas, por un puerto determinado; cantidad de ZGC inactivas; porcentaje de aciertos (similitud respecto de los datos considerados); etc.
- Para cada fenotipo, se pueden calcular indicadores que señalen el grado de ajuste entre la elección modelada y los datos reales. Los parámetros óptimos se obtendrán a partir del mejor ajuste asociado a los datos actuales. Para esto, se define la siguiente función que representa el puntaje a ser optimizado:

$$F_0 = \sum_{i,t} f_0(i, t) \tag{41}$$

$$f_0(i, t) = \begin{cases} \delta & \text{if } (e_i = er_i \neq 0) \\ 1 & \text{if } (e_i = er_i = 0) \\ 0 & \text{en otro caso} \end{cases} \quad \delta \in \mathbb{R}^+ \quad \forall t \tag{42}$$

donde i representa una ZGC y t el período observado. e_i y er_i representan el estado de una ZGC en la simulación y en los datos reales respectivamente. Por lo tanto, cuando ambas variables son iguales entonces el ACPEP ha replicado con éxito la decisión de elección de puerto. δ es un parámetro auxiliar que permite tener un grado de libertad para derivar a generaciones más adaptadas, privilegiando los aciertos activos, es decir, si el modelo predice correctamente una ZGC activa, su éxito se valora δ veces más que sobre el acierto de una zona inactiva.

El mejor conjunto de valores se obtendrá cuando F_0 sea máximo. En efecto, cada vez que el conjunto de valores se ajusta exitosamente (es decir, la elección simulada de puerto coincide con la realmente elegida por una ZGC en un período de tiempo dado), el valor de $f_0(i, t)$ aumenta.



- Una vez que el modelo ha evaluado la función objetivo F_0 , los 100 individuos se ordenan decrecientemente.
- Luego, se definen operadores de cruce entre el 30% superior de la lista y el resto de la población. Por ejemplo, en la Figura 6, se muestra el cruce entre dos individuos “padres” (el 1 y el 2) y el “hijo” que sobrevive a la generación siguiente. El punto, en donde se corta el código genético, se obtiene al azar y puede no ser único. En la Figura se muestra sólo una posibilidad, pudiendo existir muchas otras (por ejemplo, el cruce puede ser con la combinación de todos los genes de los padres).

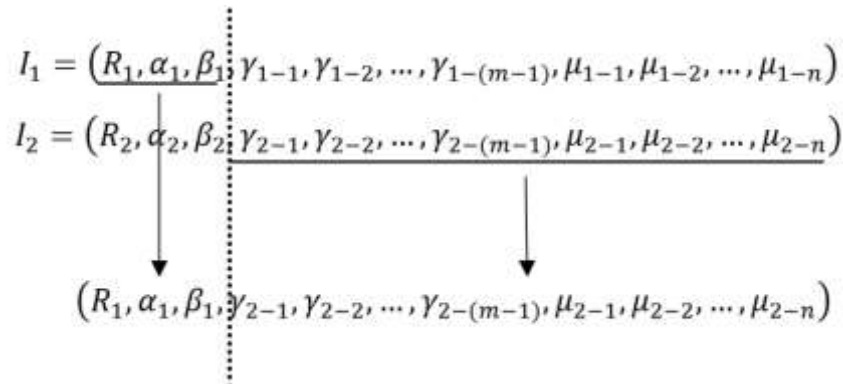


Figura 6: Ejemplo cruce entre individuo 1 y 2 en la generación t y el individuo descendiente en la generación $t+1$.

- Para incorporar diversidad en la generación siguiente se definen operadores de mutación. La Figura 7 presenta esquemáticamente la idea de este proceso de mutación. El gen a ser mutado se escoge aleatoriamente.

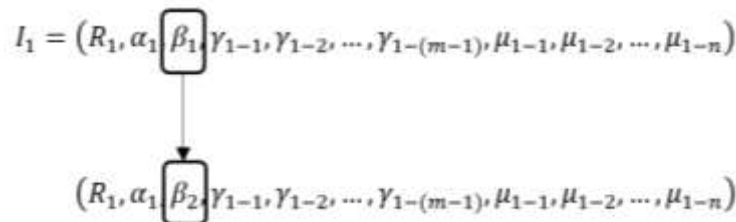


Figura 7: Esquema del operador de mutación que afecta al individuo 1 en la generación t y el individuo mutante en la generación $t+1$.

- La calidad de cada individuo se evalúa junto con un descriptor de diversidad (que permite escapar de soluciones locales) para garantizar una población heterogénea en cada generación. Así, en cada generación futura se conserva el "mejor" individuo de la generación anterior, el cual queda definido como aquél que posee mejor índice total. Este índice se define como:

$$IT = a * IC + (1 - a) * ID$$

donde IT es el índice total, IC el índice de calidad, ID el índice de diversidad y a un número real entre 0 y 1.

- Criterio de convergencia. El algoritmo genético se detiene cuando alcanza un porcentaje determinado de aciertos durante su evolución. Este porcentaje es un imput del modelador.

En la sección 5.2 se presentan los resultados de la calibración del ACPEP, así como la interfaz gráfica desarrollada para facilitar su ejecución.

3.3. Conclusiones

A la hora de elegir el modelo que mejor sirva a los objetivos del estudio, es importante no perder de vista las condiciones que definen la complejidad del proceso de elección portuaria, así como tampoco los enfoques de modelización ya aplicados a este proceso.

Como se pudo observar en el capítulo anterior, eran muy pocos los modelos que se hacían cargo de la complejidad del proceso en términos de sus características espacio temporales, como de las interrelaciones recursivas que se manifiestan entre los actores del sistema. Sólo en Garrido y Leva (2004), se mostró una estructura de modelización completa en términos de las correlaciones y heterocedasticidad de las alternativas de elección. Sin embargo, este modelo resultó altamente complejo y de difícil aplicación.



Así, el desafío fue encontrar un enfoque de modelización que diera cuenta de la complejidad del sistema estudiado, pero que permitiese una calibración y aplicación simple. Estas características son parte de lo que define la teoría de AC.

Los AC son una herramienta muy flexible para la modelización de fenómenos dinámicos, no lineales y complejos, características propias del proceso de elección portuario. Además, los AC son muy eficientes computacionalmente, ya que se basan en reglas simples de decisión que permiten la transición del estado de una célula a otro, en función de la interacción que ésta tenga con sus vecinos. El conjunto de células forma un sistema, que evoluciona de manera compleja.

Un AC se define en función de 4 componentes: unidades llamadas “células”, un número discreto de estados (por ejemplo, activo o inactivo), una vecindad, y reglas de transición.

Cada una de estas componentes fue definida para representar el proceso de elección portuaria, llamando Zonas Generadoras de Carga (ZGC) a las unidades del AC. Estas ZGC pueden encontrarse activas o inactivas en cada período de modelización, lo cual dependerá de la evaluación de su función de transición. Esta función se definió a partir de 3 componentes principales: una asociada a la interacción espacial de la ZGC con sus vecinos, que en este caso mide el grado de influencia que tiene en la decisión actual de una ZGC, la elección de puerto realizada por sus vecinos en el período anterior; otra referida al peso de la distancia que implica desplazar la carga desde la ZGC a cada puerto del conjunto de elección; y una tercera, correspondiente a una constante de atractividad propia de cada puerto.

Respecto a la componente de vecindad del AC, esta fue definida considerando que una ZGC será vecina de otra, si ambas se encuentran circunscritas en un círculo de radio R (parámetro a calibrar).

Por otra parte, dado lo compleja de la decisión, y asumiendo que un modelador no puede describir completamente el proceso de elección portuaria sólo con un conjunto de atributos, se incorporó una componente estocástica, que explicaría al menos dos fenómenos: el modelador no tiene información perfecta, y que una ZGC puede activarse/desactivarse por variables exógenas que no están incluidas en la especificación del modelo.



Así, con las componentes antes señaladas, se definió el llamado Autómata Celular Probabilístico de Elección Portuaria (ACPEP).

Respecto a la calibración, y considerando los métodos de optimización que han presentado buen rendimiento en AC, se decidió utilizar algoritmos genéticos como herramienta para la estimación de los parámetros del ACPEP.

Tanto las características del ACPEP como el algoritmo de calibración, fueron descritos en detalle en este capítulo.



Capítulo 4

ESTRUCTURA EQUIVALENTE MODELO PROBIT MULTINOMIAL



4. ESTRUCTURA EQUIVALENTE MODELO PROBIT MULTINOMIAL

Tal como se señaló en Capítulo 2 el modelo Probit Multinomial Multidimensional Espacio-Tiempo (PMMET) desarrollado por Garrido y Leva (2004), es el único modelo de elección discreta que se encontró en la literatura, que ha abordado la elección portuaria considerando adecuadamente las correlaciones espaciales y temporales existentes en el proceso. El resto de las investigaciones que han incorporado estas características utilizando modelos de elección discreta, probablemente han obtenido resultados sesgados, al no cumplir con los supuestos básicos que tienen los modelos más simples, como el MNL.

El modelo PMMET antes descrito, basa su interacción espacial entre las alternativas de elección, mientras que el ACPEP basa dicha interacción en los individuos, quienes implícitamente tienen incorporada la interdependencia entre las alternativas en su proceso de elección. Es interesante, por tanto, conocer como varía la estructura del PMMET, si quisiese aplicarse con características similares al modelo desarrollado en esta tesis. Esto permitiría evaluar el rendimiento de cada estructura de modelización y así escoger aquella que represente de mejor manera el proceso estudiado.

4.1. Modelo Probit de Elección Portuaria con Correlación entre Exportadores

En la Sección 2.2.4 se presentó una estructura de modelización discreta a partir del supuesto de alternativas de elección con interacciones espaciales y temporales. Esto dio origen al modelo dado por las expresiones (29), representando la matriz de varianza-covarianza del proceso estocástico considerado, y la expresión (37) para la función de utilidad sistemática.

Para el caso que ocupa esta tesis, se considera que la elección corresponde a la decisión del puerto de exportación en función de las características propias de los puertos y de la interacción espacial y temporal entre los exportadores que están distribuidos en el área de estudio. Así, las expresiones que habría que utilizar en este caso serían solamente aquellas que guardan relación con la elección portuaria y no con la empresa naviera.

En el caso de un grupo de n exportadores que interactúan entre sí de acuerdo a la elección que sus vecinos toman en el período inmediatamente anterior, correspondería adecuar la matriz de varianza-covarianza de (29), por una que no contengan las matrices Θ y Σ (ya que en este caso no existe la componente naviera dentro de la elección). En este caso, la expresión correspondiente a la nueva matriz de varianza-covarianza sería la siguiente:

$$E(d_t d_{t-r}') = E(e_t e_{t-r}') = \frac{\alpha^r}{1 - \alpha^2} (I - rW)^{-1} W (I - rW)^{-1} \quad (43)$$

En que r es el desfase temporal, α es el parámetro de correlación temporal; ρ es el parámetro de correlación espacial, Ω es la matriz de varianza-covarianza de un ruido blanco²³ cuya dimensionalidad corresponde al número de ZGC y W es la matriz de contigüidad que tendría también la misma dimensión.

La expresión (43) puede simplificarse aún más en el caso bajo estudio, en que la decisión de elección portuaria estaría influida solamente por la decisión en el periodo anterior (es decir r tomará sólo los valores 0 o 1). Adicionalmente, la matriz Ω representa un

²³ Ver nota al pie N°13

proceso IID Normal, por lo que puede escribirse como $\Omega = \sigma^2 I$, en que σ es un escalar que representa la desviación estándar del ruido blanco e I representa la matriz identidad de la misma dimensión que Ω .

Si se reemplaza esta expresión en lugar de Ω en la expresión (43), se advierte que los parámetros α y σ se confundirían en términos de su identificabilidad econométrica, por lo que sin pérdida de generalidad se puede suponer que $\sigma=1$.

Con todo lo anterior, la matriz de varianza-covarianza del término de error espacio-temporal estaría dada por la siguiente expresión:

$$E(\delta_t \delta_{t-r}) = \frac{\alpha^r}{1-\alpha^2} (I - \rho W)^{-1} (I - \rho W)^{-1'} \quad (44)$$

En el caso particular en que la vecindad esté definida en términos radiales (es decir, son vecinas aquellas ZGC que estén dentro de un círculo de radio dado), los elementos de W estarán dados por la siguiente expresión:

$$w_{ij} = \begin{cases} 1 & \text{si ZGC } j \text{ está dentro de un círculo centrado en } i \text{ con radio } R \\ 0 & \text{en otro caso} \end{cases} \quad (45)$$

Por su parte la función de utilidad sistemática, dada por la expresión (37) quedaría reducida a las componentes de la combinación puerto-período, tal como se señala en la siguiente expresión:

$$V(p, t) = V_{pt} = b_1 Y_{p(t-1)} + b_2 C \quad (46)$$

En esta función de utilidad representativa (o sistemática) las variables consideradas serían:

- Elección de puerto en el período anterior ($Y_{p(t-1)}$)
- Costo unitario de transporte (C)

El parámetro β_2 del costo, debiese ser negativo, pero en el caso de β_1 se desconoce a priori cuál debiese ser su signo.

Así por ejemplo para un total de n ZGC, y dos períodos consecutivos, la matriz de varianza covarianza (VCV) del sistema completo estaría dada por la siguiente expresión:

$$VCV(\delta) = \begin{bmatrix} \frac{1}{1-\alpha} \Theta & \frac{\alpha}{1-\alpha} \Theta \\ \frac{\alpha}{1-\alpha} \Theta & \frac{1}{1-\alpha} \Theta \end{bmatrix} \quad (47)$$

donde Θ corresponde al producto matricial $(I - \rho W)^{-1}(I - \rho W)^{-1'}$, de la ecuación (44), cuya dimensión sería de $n \times n$. En el caso particular que se analiza en esta tesis, se cuenta con 69 ZGC, en cuyo caso, la matriz de varianza-covarianza del término de error del modelo de elección discreta, dado por la expresión (47), tendría una dimensión de 138 filas y 138 columnas (cada una de las componentes de la expresión (47) tendría dimensión 69×69). Esta alta dimensionalidad es una complicación desde el punto de vista econométrico para la calibración de los parámetros desconocidos. En efecto, tan solo para formar la matriz (47) es necesario invertir la matriz $(I - \rho W)$ cada vez que un algoritmo de calibración realice la búsqueda de valores óptimos para ρ . Cabe mencionar además, que no hay garantía de que la matriz $(I - \rho W)$ sea invertible para cualquier valor de ρ dentro de la búsqueda realizada por el algoritmo de calibración.

Adicionalmente, los algoritmos de calibración de modelos de elección discreta (por ejemplo, el método de máxima verosimilitud simulada) suelen requerir la inversión de la matriz de varianza-covarianza del término de error, dada por la expresión (47), lo cual tampoco está garantizado para cualquier combinación de valores de α y ρ dentro de la búsqueda del algoritmo de calibración.

Todo lo anterior hace muy ineficiente el uso de este tipo de estructura espacio-temporal para la modelización de elección portuaria, ya que la complejidad en la calibración de parámetros aumenta con un factor de cuatro veces el cuadrado del número de zonas (ver forma de la matriz de la expresión (47)). Incluso, para un número modesto de zonas a considerar en una especificación (como la del ejemplo que se considera en esta tesis con 69 ZGC) se obtendría una matriz de varianza-covarianza mucho más grande que cualquiera de las que se ha publicado hasta el momento para la calibración de modelos Probit.



4.2. Conclusiones

Con el fin de tener un modelo de referencia que permitiera comparar la estructura y factibilidad de modelización del proceso de elección portuaria con las características consideradas en el ACPEP desarrollado en esta tesis, se planteó una modificación al modelo PMMET propuesto por Garrido y Leva (2004). La selección del PMMET, se basa en que fue el único modelo de elección discreta encontrado en la literatura que abordaba las características espacio temporales del proceso, sin los sesgos de modelos más simples como el MNL.

La estructura del modelo PMMET fue modificada, ya que la interacción espacial la consideraba a nivel de alternativas, a diferencia del ACPEP que la incorpora a nivel de individuos (en este caso las ZGC), quienes implícitamente tienen incorporada la correlación entre las alternativas en su proceso de elección.

Al realizar la nueva especificación del modelo PMMET, la matriz de VCV resultó con una alta dimensionalidad (dado que esta vez depende del número de ZGC), aumentando la complejidad del proceso de calibración. Lo anterior se debe a que los algoritmos de calibración de modelos de elección discreta (por ejemplo, el método de máxima verosimilitud simulada) suelen requerir la inversión de la matriz de VCV del término de error, lo cual incluso, para un número acotado de ZGC, genera una matriz mucho más grande que cualquiera de las que se ha publicado hasta el momento para la calibración de modelos Probit.

Todo lo anterior muestra que el uso de este tipo de estructura espacio-temporal en modelos de elección discreta, no es factible de incorporar y por lo tanto, no son adecuadas para representar el proceso de elección portuaria con las características recogidas por el ACPEP.



Capítulo 5

APLICACIÓN PRÁCTICA ACPEP



5. APLICACIÓN PRÁCTICA ACEP

5.1. Análisis de Datos

La metodología descrita en las secciones anteriores necesita tres conjuntos de datos. En primer lugar, un conjunto de exportadores espacialmente localizados en una región determinada. En segundo lugar, una lista de los puertos elegidos por cada exportador en cada intervalo de tiempo, y en tercer lugar, la distancia entre cada ZGC y los puertos del conjunto de elección. Si además se dispone de atributos portuarios que sirvan como proxy de su nivel de servicio, es posible agregar dichas variables en la especificación del modelo.

Son muy pocos los países que disponen de un registro público de los orígenes y destinos de carga de comercio exterior al interior de su país. En el caso de Brasil, existe una base de datos pública (<http://www.aliceweb2.mdic.gov.br/>) que permite obtener la información antes señalada desagregada por tipo de carga, con un nivel de desagregación de hasta 6 dígitos del código arancelario.

Dentro de las cargas principales no graneleras que genera Brasil, se encuentra la exportación de vehículos. Esta carga requiere naves especializadas llamadas RoRo (roll on roll of), las cuales en general tienen una menor prioridad de atención en puertos que no poseen terminales de vehículos especializadas. Esto ha llevado a que marcas como



Ford, construyan terminales portuarios privados, evitando así los problemas de congestión de los puertos de uso público.

Sin embargo, en Brasil se encuentran instaladas más de 30 marcas de vehículos, distribuidas en plantas ubicadas principalmente en el sur y sureste del país (Anfavea, 2016), que utilizan los puertos públicos disponibles.

Para aplicar el modelo propuesto se consideró un conjunto de puertos de uso público con hinterland común, asegurando así, que las ZGC tuviesen el mismo conjunto de alternativas disponibles. En este ejemplo de aplicación los puertos escogidos fueron Paranaguá (Puerto 4), Santos (Puerto 3), San Sebastián (Puerto 1) y Rio de Janeiro (Puerto 2).

El puerto de San Sebastián, posee un muy bajo movimiento de vehículos, sin embargo, se decidió mantener en el conjunto de elección, para probar la capacidad del modelo frente a una alternativa que pudiese mejorar sus condiciones de operación.

Las ZGC, en este ejemplo, corresponden a municipios, es decir, a las divisiones geográficas administrativas, que habían exportado vehículos entre los años 1997 y 2015, a través de alguno de los cuatro puertos seleccionados. Como resultado de esta clasificación, se obtuvo un conjunto de 69 ZGC.

La unidad temporal escogida fue trimestres, generándose así, una serie de tiempo de 74 observaciones.

La distribución espacial de las 69 ZGC y los 4 puertos considerados se presenta en la Figura 8.



Figura 8: Puertos y ZGC.

Como se puede observar en la figura anterior, las ZGC no se encuentran uniformemente distribuidas dentro del hinterland considerado. Existen clusters correspondientes a los principales nodos logísticos de la industria automotriz de esta región de Brasil.

Respecto al porcentaje observado de ZGC activas en cada período, éste varía entre 12% y 39%, siendo el promedio de los últimos 5 años (20 períodos), igual a 29%. La Figura 9, muestra este porcentaje para cada período disponible, así como la curva logarítmica ajustada.

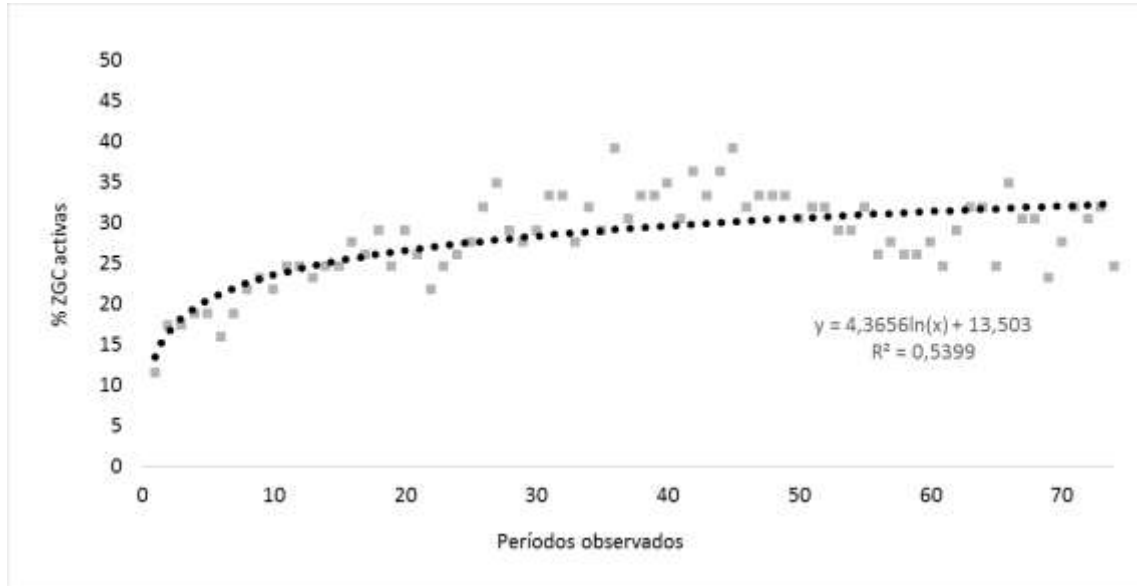


Figura 9: Porcentaje de ZGC activas en cada período observado.

En cuanto a la participación histórica de mercado de cada puerto, es posible observar en la Figura 10, que el Puerto 1 (San Sebastián) tiene una participación muy menor, a pesar de ser el puerto más cercano al puerto con mayor captura de mercado (Puerto 3, Santos). Lo anterior se explica por la falta de condiciones atractivas que tiene el Puerto 1 para el mercado de exportación de vehículos (infraestructura, accesibilidad, equipos, tarifas, etc.).

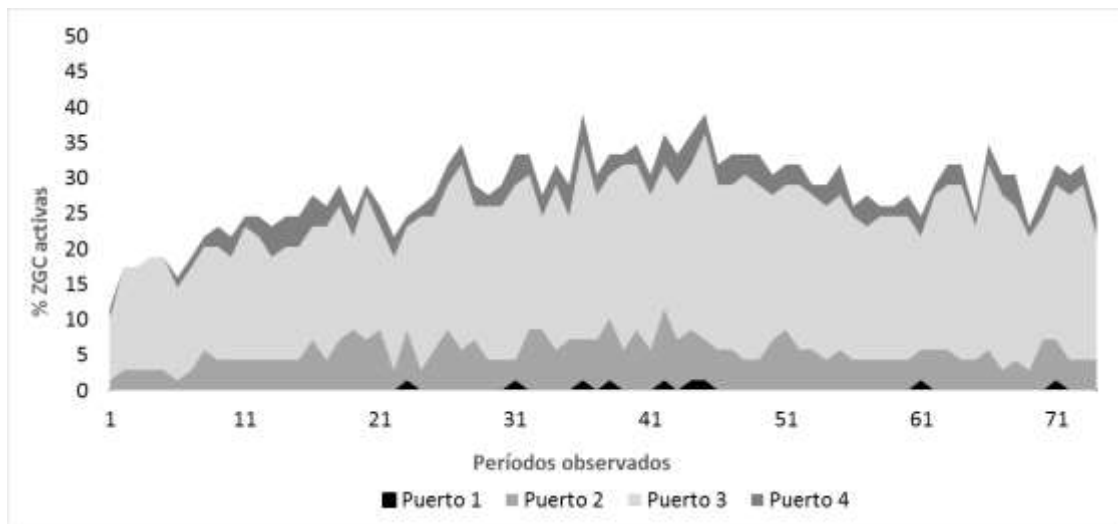


Figura 10: Distribución histórica de ZGC según su elección portuaria.

Respecto a la localización de las ZGC y su distancia a puertos competidores, el 4% de ellas se localiza más cerca del Puerto 1 (San Sebastián), 23% del Puerto 2 (Río de Janeiro),

61% del Puerto 3 (Santos) y un 12% del Puerto 4 (Paranaguá). En la Figura 11 se distinguen las ZGC según el puerto más cercano a cada una de ellas.



Figura 11: ZGC según puerto más cercano.

Las proporciones de ZGC según distancia al puerto más cercano, son de órdenes de magnitud similares a la distribución de ZGC activas por puerto en los últimos 5 años observados (0.5% Puerto 1, 16.5% Puerto 2, 73.7% Puerto 3, 9.3% Puerto 4). Lo anterior puede ser un indicador de la importancia que tiene la distancia en la elección portuaria.

En vista de esto, se analizaron las elecciones portuarias de cada ZGC durante el período de tiempo estudiado, y se comparó dicha elección, con la localización del puerto más cercano a cada una de ellas.

En la Figura 12 se muestra la distribución promedio de la elección de cada ZGC en el período observado (considerando sus períodos de actividad), así como el puerto más cercano a dicha zona (columna derecha).

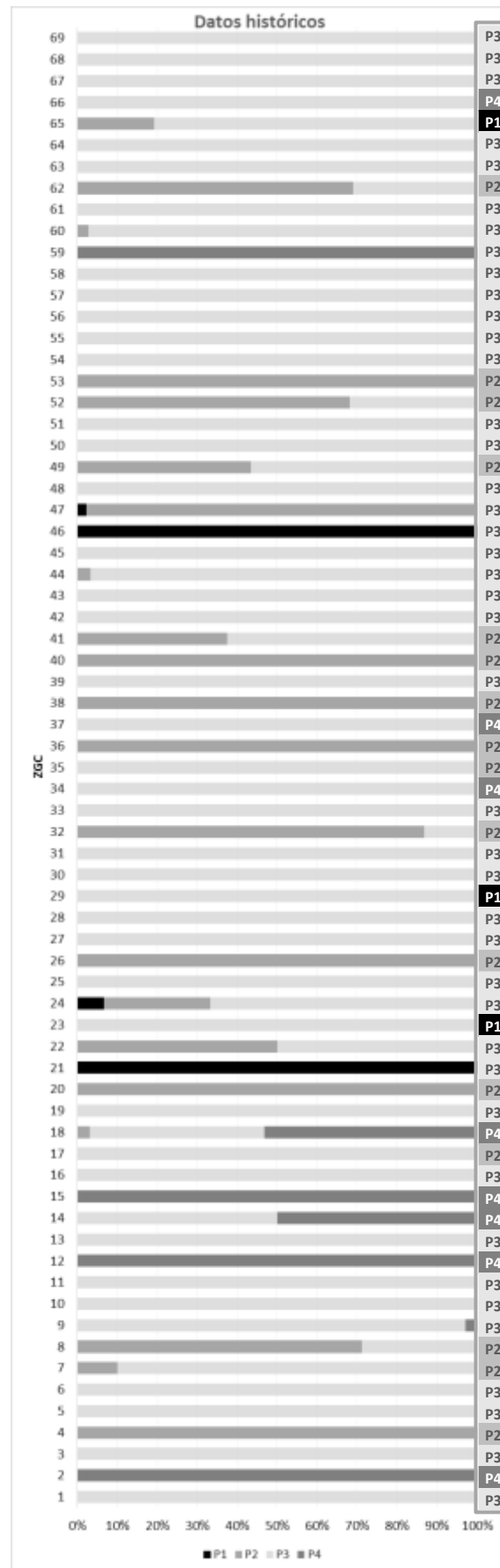


Figura 12: Distribución histórica promedio de la elección de cada ZGC.

Los datos muestran que de las 69 ZGC, hubo 25 que escogieron un puerto distinto al más cercano, en alguno de los períodos analizados. Esto confirma la importancia de la distancia en la elección portuaria, pero da cuenta también de la existencia de otros factores que afectan este proceso de elección, que hacen que el 36% de las ZGC escoja puertos más lejanos.

5.2. Calibración ACPEP

5.2.1 Interfaz Gráfica del Algoritmo

El algoritmo genético presentado en la sección 3.2.3 fue implementado en el lenguaje de programación Lazarus²⁴ (Van Canneyt et al., 2011).

Para facilitar la ejecución de este algoritmo, se desarrolló una interfaz gráfica, en la cual se ingresan los parámetros requeridos, los que corresponden a: porcentaje de aciertos, porcentaje de hijos (cruce), porcentaje de mutación, peso zonas activas (δ en la ecuación (42)) y el valor máximo de iteraciones. También se dispone de la opción “*usar semilla*”, la cual permite entregar un valor de partida al algoritmo.

El programa se detendrá cuando se alcance el porcentaje de aciertos o el máximo de iteraciones indicadas por el usuario. En la Figura 13 se muestra la interfaz señalada.



Figura 13: Interfaz del programa desarrollado para calibrar el algoritmo genético.

²⁴ En el Anexo A.1 se presenta el código del algoritmo descrito.

Tal como se observa en la imagen anterior, una vez que finaliza el algoritmo, se muestra la opción de graficar en pantalla el porcentaje de ZGC obtenido para cada período considerado. Además, para facilitar la comparación respecto a los datos observados, se grafica simultáneamente el porcentaje de ZGC activas que hubo en la serie de tiempo utilizada para la calibración. En el mapa, en tanto, se muestra el estado de las ZGC, en el último período observado y simulado, tal como se puede observar en la Figura 14.



Figura 14: Interfaz del programa una vez finalizado el algoritmo genético.

5.2.2 Resultados Calibración

Los parámetros resultantes del modelo ACPEP, se presentan en la Tabla 5. Cabe señalar que como el parámetro constante γ_j de la ecuación (38) (fijo para cada puerto) fue calibrado para tener en cuenta posibles variables no observadas, sólo es posible estimar tres de ellos (con el fin de evitar problemas de identificabilidad). Así, el valor de γ_3 se fijó en 0 (sin pérdida de generalidad). Análogamente, el valor de la probabilidad p se fijó en 0,75.

Parámetro	Valor	μ_{19}	0.53	μ_{44}	0.22
Radio (km)	267	μ_{20}	0.04	μ_{45}	0.55
α	828	μ_{21}	0.05	μ_{46}	0.17
β	-12	μ_{22}	0.09	μ_{47}	0.54
γ_1	-818	μ_{23}	0.13	μ_{48}	0.04
γ_2	-417	μ_{24}	0.27	μ_{49}	0.07
γ_4	-297	μ_{25}	0.82	μ_{50}	0.06
μ_1	0.17	μ_{26}	0.02	μ_{51}	0.42
μ_2	0.06	μ_{27}	0.004	μ_{52}	0.97
μ_3	0.56	μ_{28}	0.01	μ_{53}	0.09
μ_4	0.27	μ_{29}	0.01	μ_{54}	0.16
μ_5	0.05	μ_{30}	0.07	μ_{55}	0.34
μ_6	0.02	μ_{31}	0.82	μ_{56}	0.13
μ_7	0.00	μ_{32}	0.39	μ_{57}	0.83
μ_8	0.76	μ_{33}	0.55	μ_{58}	0.83
μ_9	0.08	μ_{34}	0.55	μ_{59}	0.15
μ_{10}	0.01	μ_{35}	0.05	μ_{60}	0.89
μ_{11}	0.08	μ_{36}	0.09	μ_{61}	0.88
μ_{12}	0.03	μ_{37}	0.15	μ_{62}	0.66
μ_{13}	0.37	μ_{38}	0.17	μ_{63}	0.04
μ_{14}	0.05	μ_{39}	0.06	μ_{64}	0.93
μ_{15}	0.01	μ_{40}	0.21	μ_{65}	0.99
μ_{16}	0.52	μ_{41}	0.12	μ_{66}	0.01
μ_{17}	0.17	μ_{42}	0.13	μ_{67}	0.03
μ_{18}	0.95	μ_{43}	0.84	μ_{68}	0.01
				μ_{69}	0.37

Tabla 5: Parámetros obtenidos para el ACPEP

El proceso de calibración alcanzó un 73% de aciertos en la predicción de elección de puerto de exportación por parte las ZGC. Esto es un alto grado de éxito considerando

que hay 69 ZGC realizando simultáneamente su decisión durante 74 períodos (trimestres).

Tal como se señala en la ecuación (42) de la sección 3.2.3, el porcentaje de aciertos, se obtiene considerando pesos desiguales dependiendo de si se acierta que la ZGC está inactiva (es decir, tanto el modelo como los datos observados coinciden en que la ZGC no exporta por dichos puertos), o si se acierta en que está activa (es decir, tanto el modelo como los datos observados coinciden en que la ZGC exporta por alguno de los puertos considerados). En esta aplicación del modelo, el peso asignado a un éxito inactivo fue de 1.0, mientras que el éxito de una ZGC activa (δ) fue de 4.0.

Este número de aciertos alcanzado también es alto si se compara con el éxito de otros modelos econométricos de transporte, calibrados con datos reales. En el caso de los modelos de elección discreta la forma econométricamente correcta de medir su bondad de ajuste ha sido a través de la comparación de los logaritmos de la función de verosimilitud respecto de porciones de mercado fijas (modelos solo constantes, representados por ρ^2) o bien, respecto de elección de eventos de igual probabilidad (modelos equiprobables, representados por ρ_0^2). En la literatura se reporta con mayor frecuencia ρ^2 , alcanzando valores cercanos a 0.5 en los modelos de elección portuaria, tal como se muestra en la Tabla 6.

Autor*	Modelo	ρ^2
Nir et al. (2003)	MNL	0.645
Tiwari et al. (2003)	MNL	0.57
Malchow y Kanafani, (2004)	MNL	0.54
Vega et al. (2014)	HL	0.568
Veldman et al. (2011)	MNL	0.377
Veldman et al. (2015)	MNL	0.507

Tabla 6: Bondad de ajuste de modelos de elección discreta aplicados a elección portuaria.

*Las variables consideradas en estos modelos pueden encontrarse en la Tabla 4 presentada en el Capítulo 2.

Respecto al radio de la vecindad obtenido en la calibración, esto es 267 km, se pudo observar que el 83% de las ZGC tenía 5 o más vecinos al considerar dicho radio, siendo el promedio 24 vecinos. En la Figura 15 se muestra el detalle de vecinos por ZGC.

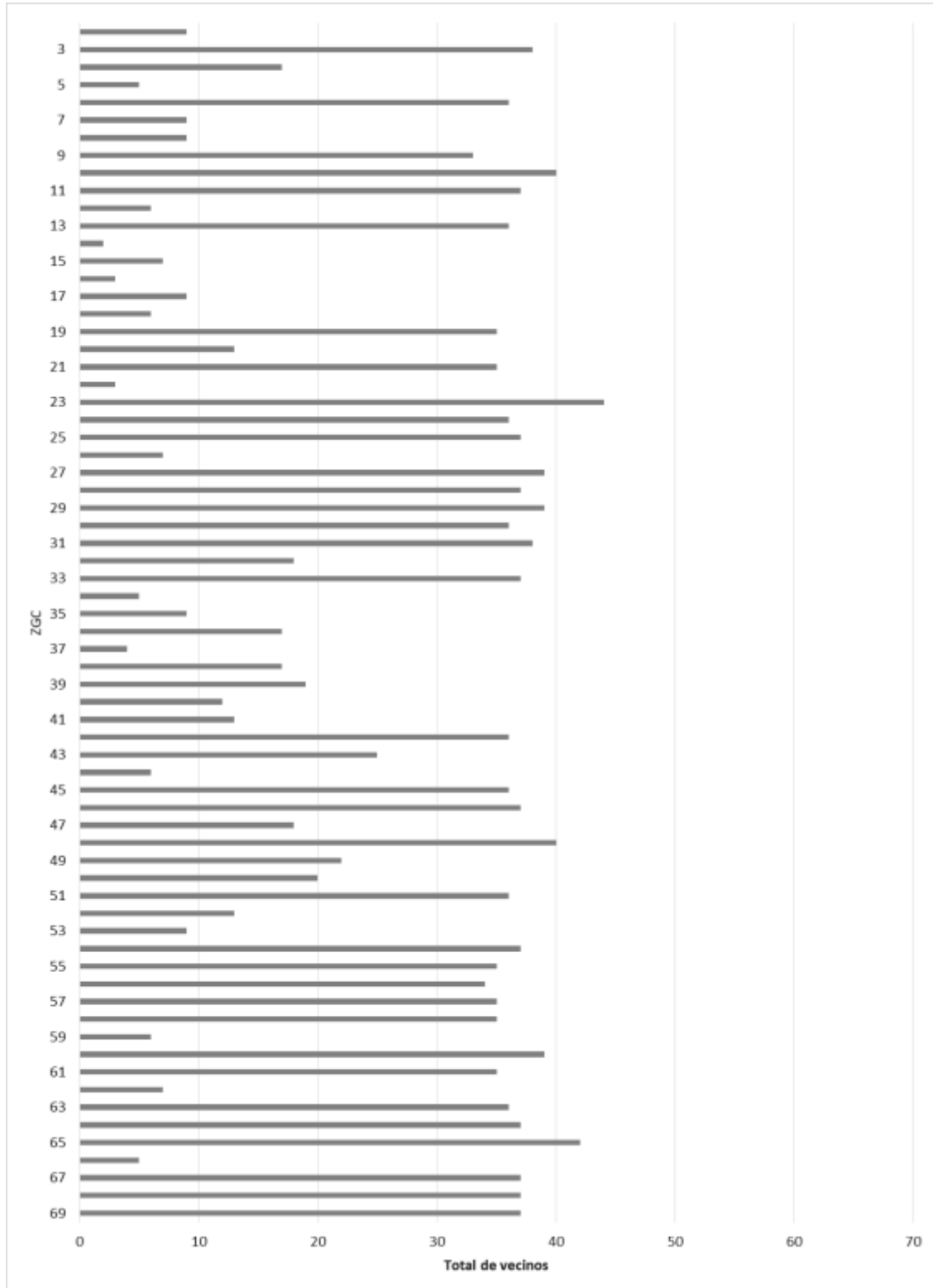


Figura 15: Total de vecinos por cada ZGC

El que exista un promedio de 24 vecinos por ZGC, junto con el valor obtenido para el parámetro espacial α , es un indicador del alto grado de influencia que tienen los vecinos, sobre la elección portuaria de una ZGC dada.

Respecto a las constantes de atractividad de los puertos (γ), todas tuvieron signo negativo, lo cual era esperado debido que la constante del puerto con mayor participación de mercado histórica fue fijada en cero. De igual manera, el puerto que obtuvo la menor constante coincidió con el que tenía menor participación de mercado.

El parámetro β , en tanto, que acompaña a la variable distancia, obtuvo signo negativo, tal como era esperado.

El resto de los parámetros estimados, tienen relación con el umbral de activación de una ZGC (μ). La Figura 16 presenta un histograma de la probabilidad de activación de las ZGC, observándose que solo un 29% de ellas tiene una probabilidad mayor a 0.5. Este valor indica, por tanto, una relativa resistencia de activación en cada período.

Lo anterior es consistente con lo observado en los datos reales, donde cerca del 30% de las ZGC permanecen activas durante el lapso de tiempo estudiado. Esta característica dinámica es fácilmente representada por el ACPEP, a diferencia de la mayoría de los enfoques estándar de modelización, como lo son los modelos de elección discreta, en los que se producen grandes errores cuando hay pocas observaciones con actividad. Esto último se debe al hecho de que la varianza del término de error en los modelos de elección discreta es independiente del valor de la probabilidad de elección. Por lo tanto, la magnitud del error es mayor cuando se comparan alternativas con baja participación de mercado. En este caso, la componente del error podría ser mucho mayor que su componente sistemática.

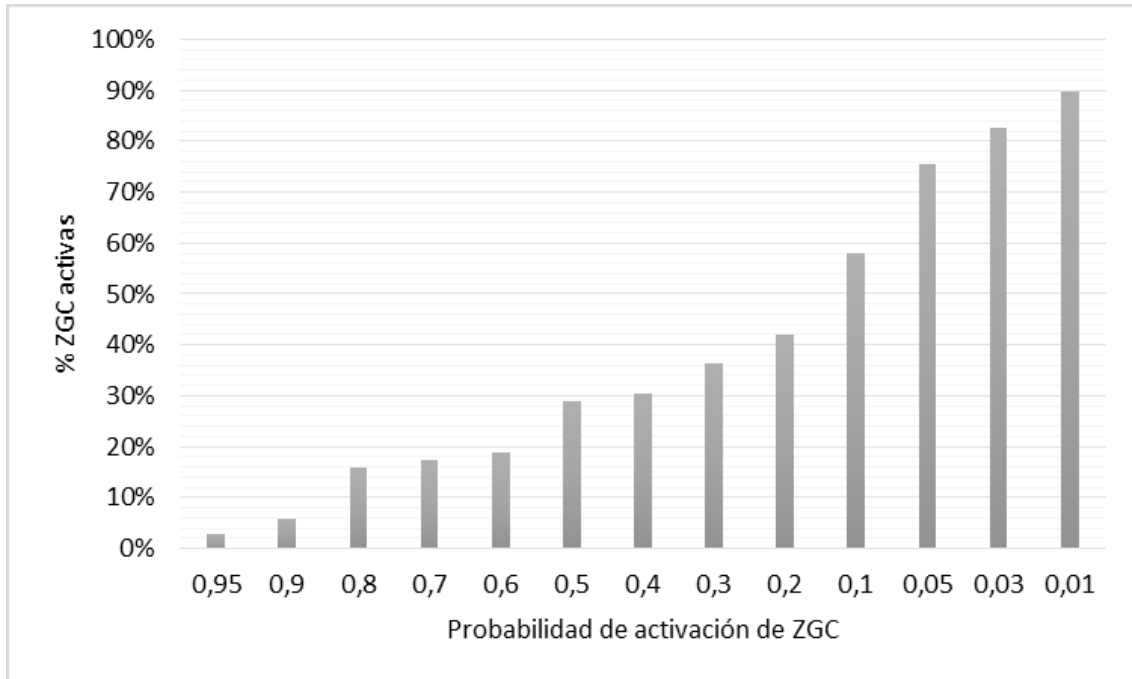


Figura 16: Probabilidad de activación de ZGC

5.3. Ejecución del Modelo

5.3.1 Interfaz Gráfica

Al igual que en el caso del algoritmo de calibración, se desarrolló una interfaz gráfica que permitiera realizar las simulaciones del ACPEP con facilidad. En la Figura 17 se muestra dicha interfaz.

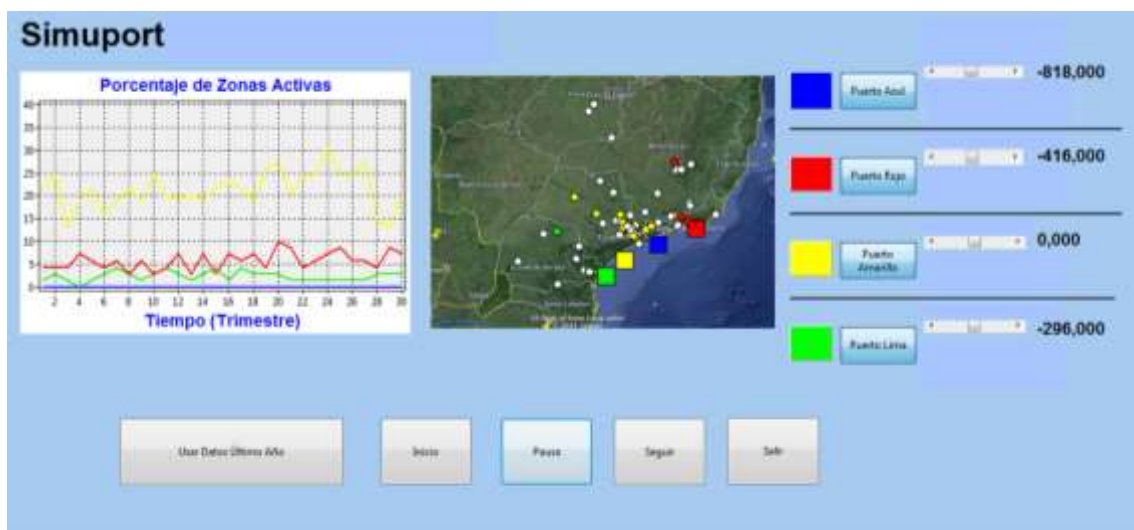


Figura 17: Interfaz inicio para simulación del ACPEP

Como se puede observar, a medida que avanza la simulación, se presenta en pantalla el porcentaje de ZGC activas por puerto, así como su distribución espacial en el mapa. Es posible, además, durante la simulación, variar las constantes de atractividad de los puertos (indicadas en el borde derecho de la pantalla), observando así el impacto que se genera en la elección portuaria de las ZGC.

Otra opción que presenta esta interfaz es la posibilidad de desactivación de uno o más puertos, visualizando en pantalla, la reasignación de las ZGC entre los puertos restantes. En la Figura 18 se muestra un ejemplo en que se desactiva el puerto de Santos (amarillo). En ella se puede observar como el puerto de San Sebastián (azul), captura parte importante de las ZGC que han dejado de ser atendidas por Santos.



Figura 18: Visualización de la reasignación de ZGC con el puerto de Santos desactivado

5.3.2 Rendimiento del Modelo

Para evaluar el rendimiento del modelo como herramienta predictiva se realizaron simulaciones de prueba sin alterar los parámetros asociados a cada puerto. Así, el cambio en la actividad de las zonas y su asignación a cada puerto, se explicaría principalmente por la componente $\alpha\Phi_1^i$ de la ecuación (38) (efecto espacial de la vecindad), y por la aleatoriedad frente al umbral de activación (μ).

El modelo fue ejecutado 20 veces durante 80 períodos (cada período representa un trimestre, por lo que el periodo de predicción equivale a 20 años), con el fin de observar su comportamiento en el largo plazo.

La Figura 19 muestra el porcentaje de ZGC activas en cada período simulado (trimestre). Los puntos representan los resultados de la simulación, y la línea sólida, el promedio por período de las 20 ejecuciones.

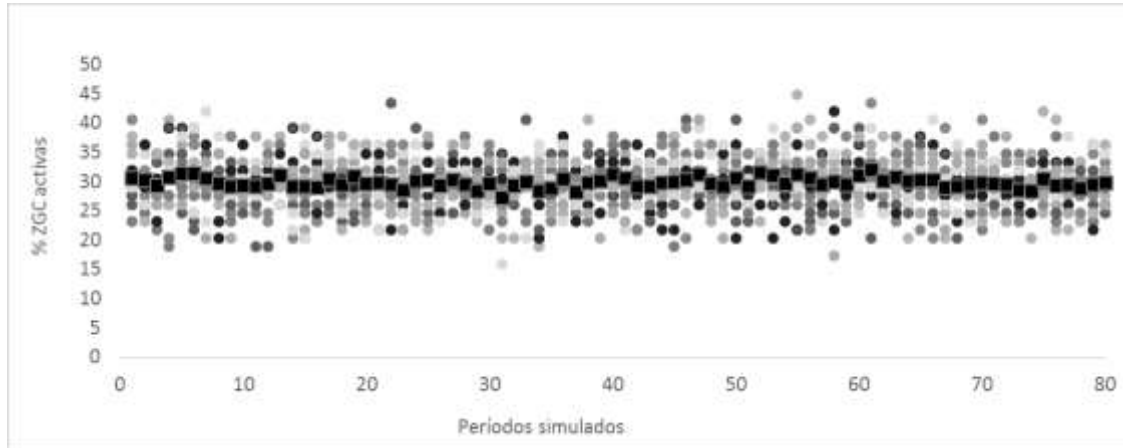


Figura 19: Predicción ZGC activas a 20 años

Tal como se puede observar, y considerando las condiciones señaladas, el modelo muestra un comportamiento estable en el tiempo. El número promedio de ZGC activas es de 29.8% versus el 28.1% del promedio histórico de los datos observados. Más aún, si nos concentramos en los datos observados en los últimos 10 años, la diferencia entre el porcentaje de ZGC activas y el obtenido con el simulador es de sólo 0.9 puntos (29.9% modelado por el ACPEP versus 30.8% en los datos observados).

La herramienta de modelización es precisa también en la estimación de la distribución de ZGC por puerto. En la Tabla 7 se puede observar que la mayor variación con respecto a los datos observados se encuentra en el Puerto 3, con una diferencia de 1.5 puntos porcentuales, lo cual es considerado de alta precisión en la práctica.

	Puerto 1	Puerto 2	Puerto 3	Puerto 4	ZGC activas
Promedio simulación	0.0%	5.9%	21.3%	2.6%	29.8%
Promedio datos observados*	0.2%	5.4%	19.8%	2.7%	28.1%

Tabla 7: Comparación de la distribución simulada de ZGC por puerto respecto a la distribución real.

*Promedio de la distribución de ZGC por puerto de los 74 períodos observados

5.4. Conclusiones

El ACPEP fue aplicado al caso de exportaciones de vehículos de Brasil, para lo cual se consideraron las exportaciones realizadas entre los años 1997 y 2015. Los datos fueron tratados de manera trimestral, por lo que se contó con una serie de 74 períodos.

Con el fin de asegurar que las ZGC contaran con el mismo conjunto de alternativas disponibles, se escogieron puertos de uso público y con un hinterland común. Los puertos que cumplían estas condiciones fueron: Paranaguá, Santos, San Sebastián y Rio de Janeiro. El puerto de San Sebastián, a pesar de poseer un muy bajo movimiento de vehículos, se decidió mantener con el fin de probar la capacidad del modelo frente a una alternativa que pudiese mejorar sus condiciones de operación.

Respecto a las ZGC, estas fueron definidas como aquellos municipios que exportaron vehículos por alguno de los puertos del conjunto de elección, durante el período analizado. Hubo un total de 69 municipios que cumplieron esta condición.

Así, la instancia de aplicación del ACPEP, quedó definida por 69 ZGC, 4 puertos y 74 períodos de observación.

Una de las características de los datos observados, es que en promedio sólo un 30% de las ZGC se mantenía activa. Esto en general implica un desafío mayor en la calibración de modelos de elección discreta, ya que estos no responden adecuadamente cuando se tiene un bajo número de observaciones activas. Sin embargo, los AC pueden representar fácilmente esta dinámica.

Respecto a la distribución histórica de las ZGC a los puertos, y concordante con la revisión bibliográfica analizada, se detectó que un porcentaje importante de ZGC escogía el puerto más cercano a ella. Sin embargo, hubo un 36% que no siguió este comportamiento, dando cuenta así de la existencia de otros factores que afectan este proceso de elección.

En cuanto a la calibración del modelo, se obtuvo un porcentaje de éxito del 73%, siendo superior incluso a muchos modelos de elección discreta que han tratado el proceso de elección portuaria.

Los resultados confirmaron la hipótesis realizada respecto al peso de la interacción espacial entre las ZGC, tanto por el parámetro obtenido que la representa dentro del ACPEP, como por el radio resultante de la vecindad con la cual cada ZGC analiza la decisión de sus vecinos. De acuerdo a este radio, el 83% de las ZGC tendría en su vecindad 5 o más vecinos.

Respecto a los parámetros que definen la probabilidad de activación de las ZGC, se observaron valores que indican una relativa resistencia de activación de las ZGC. Lo anterior es consistente con lo observado en los datos reales, donde sólo el 30% de las ZGC permanecen activas durante el lapso de tiempo estudiado.

Las constantes de atractividad de los puertos (γ), en tanto, tuvieron signo negativo tal como era esperado, debido que la constante del puerto con mayor participación de mercado histórica fue fijada en cero. Por este mismo motivo, el puerto que obtuvo la menor constante coincidió con el que tenía menor participación de mercado.

Finalmente, el parámetro que acompaña a la variable distancia obtuvo signo negativo, tal como era esperado.

Con los parámetros calibrados se procedió luego a la ejecución del modelo, realizándose 20 simulaciones durante 80 períodos. Los resultados mostraron que el ACPEP representa adecuadamente la dinámica del sistema, obteniendo participaciones de mercado y porcentajes de ZGC activas, similares a las observadas en los datos reales.



Capítulo 6

CAPACIDAD PREDICTIVA DEL ACPEP

6. CAPACIDAD PREDICTIVA DEL MODELO

6.1. Redistribución de ZGC por Cambio en el Nivel de Servicio

Portuario

Uno de los muchos tipos de análisis que se pueden realizar con una herramienta como los ACPEP, es la predicción de la participación de mercado como resultado de un cambio en el nivel de servicio en uno o más puertos. En este modelo en particular, las características portuarias son representadas por el parámetro γ_j . Cambiar los valores de γ_j es equivalente a alterar los atributos de cada puerto que no están relacionados por la distancia. Así, un mejor nivel de servicio, estará representado por un γ_j mayor.

A modo de ejemplo, se estudió el comportamiento del mercado cuando el nivel de servicio de los puertos 1 y 3 era igual ($\gamma_1 = \gamma_3$). Estos puertos se encuentran muy cercanos (aproximadamente a una distancia de 160 km entre ellos), por lo que, si los exportadores perciben el mismo nivel de servicio, se esperaría que la participación del Puerto 1 aumentara. Cuando este escenario fue simulado, el crecimiento del Puerto 1 fue relativamente pequeño (sólo 1.6%). La distancia entre las ZGC y ambos puertos explican lo anterior. En efecto, en promedio, las ZGC están 100 km más cercanas al Puerto 3 que al Puerto 1. Este resultado muestra que, ante iguales niveles de servicio, los exportadores preferirían el puerto más cercano. Se decidió entonces estudiar, cuán

mayor debería ser el nivel de servicio de un puerto sobre otro, de manera que superara el efecto que genera la distancia en la elección. En términos de modelización, esto implica determinar en cuánto debe aumentar el valor del parámetro γ_j (para el puerto de menor participación de mercado), con el fin de lograr una participación de mercado similar entre ambos puertos.

La Figura 20 presenta el cambio en la captura de ZGC de cada puerto en función de valores crecientes de γ_j . En ella se observa que cuando el valor de γ_1 se aproxima a 1,250, la demanda capturada por el Puerto 1 aumenta considerablemente. Sin embargo, este cambio se inicia luego de 130 periodos, llegando a superar al Puerto 3 recién en el período 495.

La explicación de este comportamiento dinámico, es que cuando la expresión (38), es evaluada para los puertos 1 y 3, la diferencia en el valor del término $\beta(d_{i1}, d_{i2}, \dots, d_{im})$ entre ambos puertos, alcanza un valor similar a la diferencia entre sus correspondientes γ_j , es decir, 1,250. Lo anterior ocurre porque hay 43 ZGC, que están aproximadamente 100 km más cerca del Puerto 3 que del Puerto 1. Este hecho, hace que el valor β (estimado en -12 en la calibración) multiplicado por la diferencia de distancias, entregue un valor de aproximadamente 1,250.

Note que el tiempo en que se estabiliza la captura de demanda está lejos de ser inmediata. La demora en llegar a este punto se debe al efecto de la vecindad en la función de elección de la expresión (38), es decir, $\alpha\Phi_1^i$, dando cuenta de la inercia espacial y temporal de la elección. Así, si el puerto 1 mejorara su nivel de servicio de tal manera que $\gamma_1 \geq 1400$, se esperaría que en 10 periodos (menos de 3 años) igualara la participación de mercado del puerto 3.

El comportamiento descrito concuerda con lo observado en la realidad, ya que cuando dos puertos comparten su hinterland, el puerto más lejano debe ofrecer condiciones considerablemente mejores que su competencia para generar un cambio significativo en el patrón de captura de demanda. El tiempo que demore en captar más ZGC, dependerá a su vez de las ventajas que observen los exportadores respecto a los costos extras que les genere desplazarse al puerto más lejano. Así, mientras mejor sea el nivel de servicio, más rápido será el cambio en la elección de los exportadores.

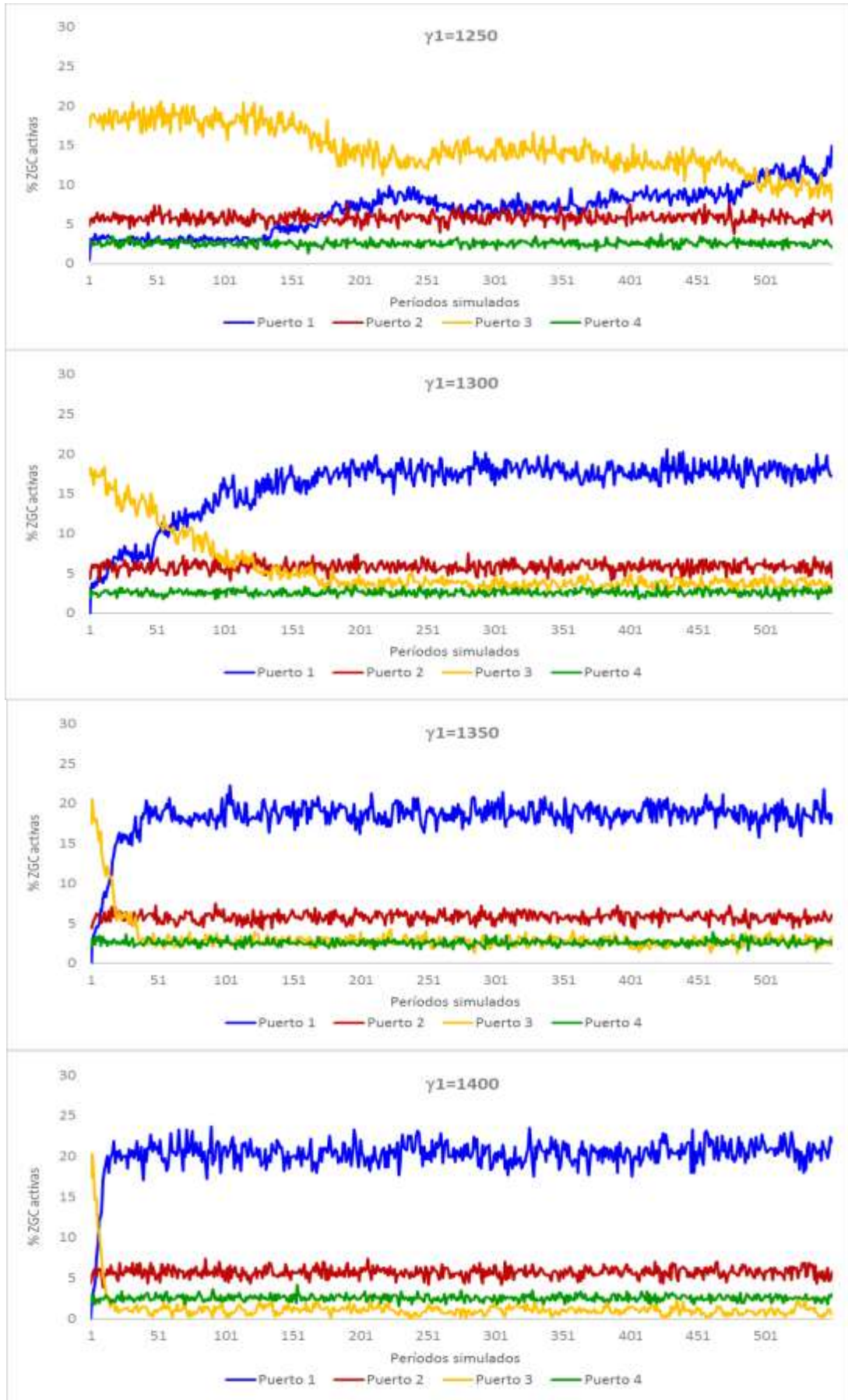


Figura 20: Distribución de ZGC según su elección portuaria con $\gamma_1 = 1250, 1300, 1350$ y 1400

6.2. Redistribución de ZGC por Desactivación de un Puerto

Otro tipo de análisis que se puede realizar es observar cómo varía el hinterland de cada puerto, cuando alguno de ellos deja de estar operativo. Tal como se señaló en la sección 5.1, la distribución por puerto de ZGC activas, considerando el promedio de los 5 últimos años observados (20 períodos) fue de:

- 0.5% Puerto 1 (San Sebastián)
- 16.5% Puerto 2 (Río de Janeiro)
- 73.7% Puerto 3 (Santos)
- 9.3% Puerto 4 (Paranaguá)

Interesa por tanto conocer cuál sería el efecto en estos porcentajes de captura, si desactivamos alguno de los puertos estudiados.

Para evaluar este efecto, el modelo fue ejecutado 10 veces durante 80 períodos, considerando escenarios en que sólo existían 3 puertos activos. El resto de las condiciones se mantuvo constante, es decir sin variación en los parámetros del modelo. Los valores a comparar con los datos reales, se estimaron como el promedio de los valores obtenidos en cada ejecución del modelo.

En el caso de la desactivación del Puerto 1 (San Sebastián), vemos en la Figura 21 que no hubo una mayor reasignación de ZGC, esto debido a que este puerto contaba con una muy baja captura de ZGC (sólo el 0.5%). Se observa además una baja variación a lo largo de los períodos simulados, indicando, que, de no existir variaciones en los puertos disponibles, la distribución de ZGC se mantiene sin mayores variaciones en el tiempo. Si consideramos el promedio de los valores obtenidos en cada ejecución del modelo, vemos en la Tabla 8, que los puertos tuvieron una baja variación en su captura, respecto a la distribución de ZGC observada en los últimos 5 años, esto, a pesar de la desactivación del Puerto 1.

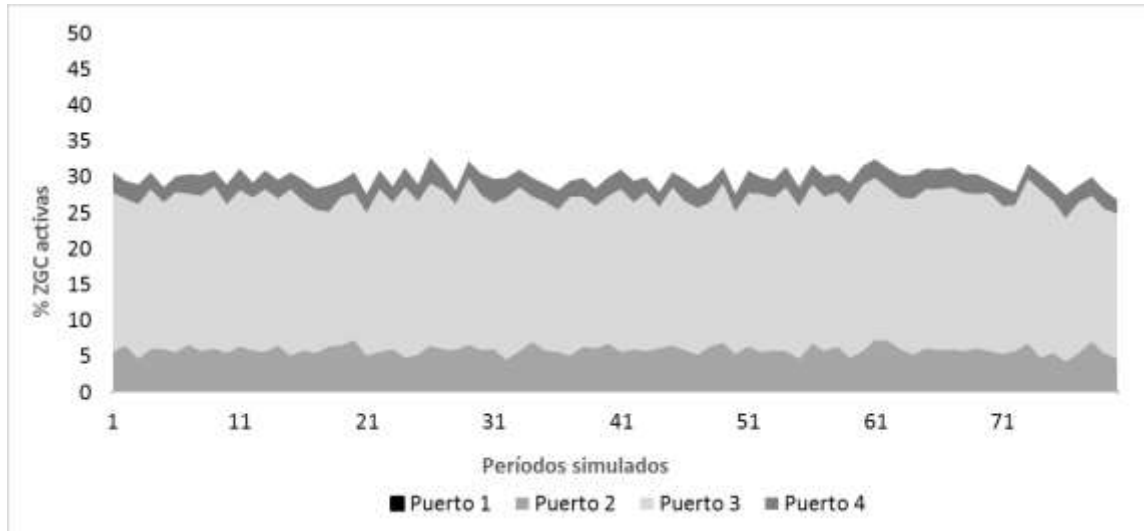


Figura 21: Distribución de ZGC según su elección portuaria sin la existencia del Puerto 1.

	Puerto 1	Puerto 2	Puerto 3	Puerto 4	ZGC activas
Distribución observada de ZGC en últimos 5 años	0.5%	16.5%	73.7%	9.3%	28.6%
Distribución promedio simulada de ZGC sin Puerto 1*	-	19.8%	71.4%	8.8%	29.9%
Desviación estándar de simulación de distribución de ZGC sin Puerto 1*	-	1.9	2.1	1.3	1.2

Tabla 8: Variación en la distribución de ZGC según su elección portuaria sin la existencia del Puerto 1.

*Los valores señalados corresponden al promedio de los 80 periodos simulados.

En la Figura 22, en tanto, se muestra la comparación entre la distribución promedio de la elección portuaria de cada ZGC en el período observado y la redistribución promedio de la elección de cada ZGC en el período simulado (considerando sus períodos de actividad). En la columna derecha se indica el puerto más cercano a dicha zona. Tal como se observa producto de la reasignación de 25 ZGC, queda un 94% del total de ZGC asignada al puerto más cercano.

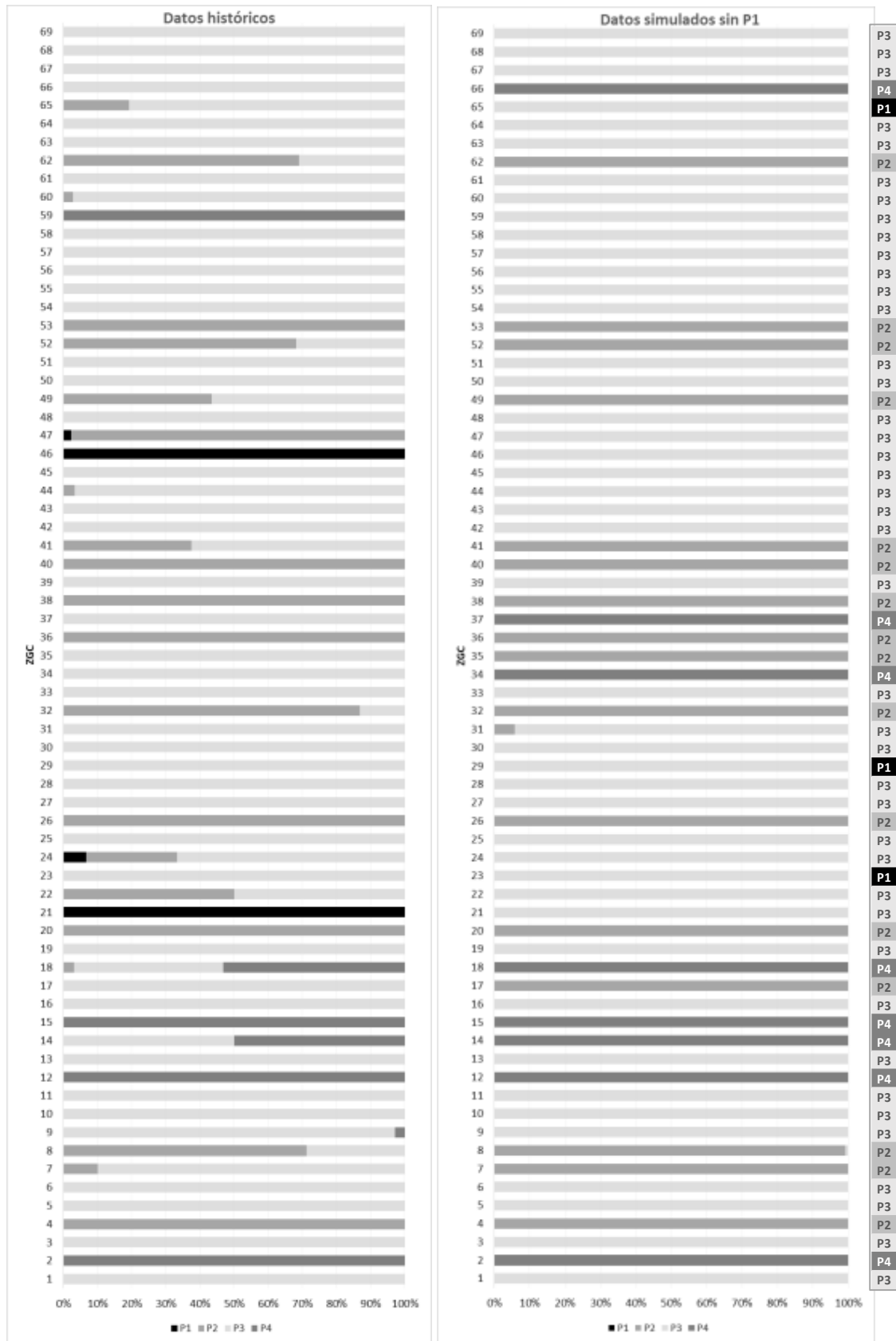


Figura 22: Comparación de la distribución histórica promedio de elección portuaria de cada ZGC con la distribución simulada de elección sin la existencia del Puerto 1.

Al desactivar el Puerto 2, en tanto, se observa en la Figura 23 una mayor redistribución de las ZGC. En este caso el 16.5% de captura que tenía el Puerto 2 en los últimos 5 años, pasa al Puerto 3 aumentado de 73.7% a 91.1%. En la Tabla 9 se muestra la variación respecto a la distribución de ZGC observada en los últimos 5 años.

Al igual que en el caso anterior, se observa además una baja variación a lo largo de los períodos simulados, indicando, que, de no existir variaciones en los puertos disponibles, la distribución de ZGC se mantiene sin mayores variaciones en el tiempo.

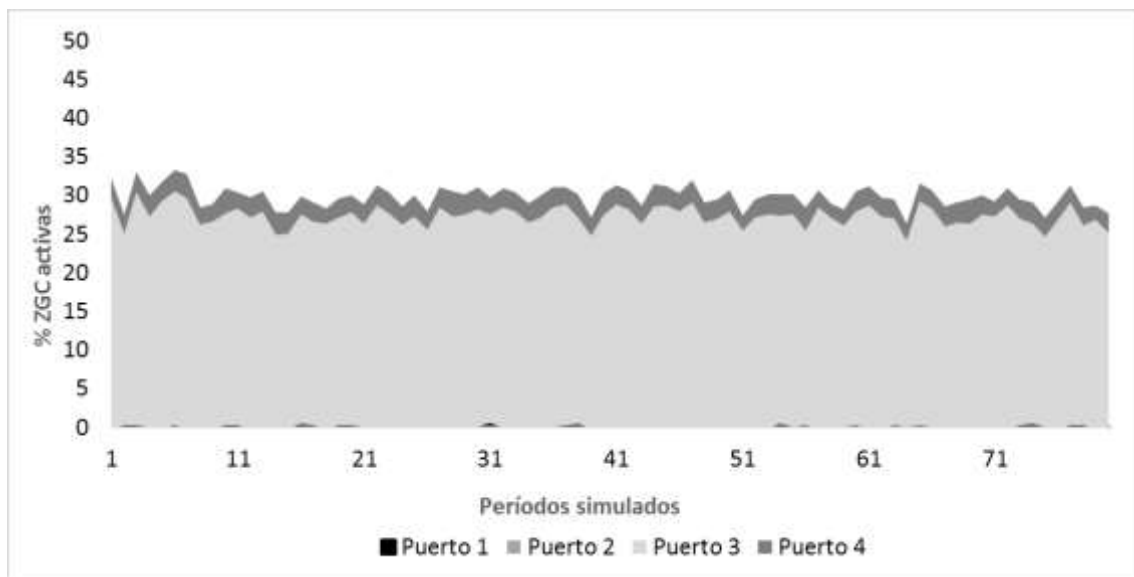


Figura 23: Distribución de ZGC según su elección portuaria sin la existencia del Puerto 2.

	Puerto 1	Puerto 2	Puerto 3	Puerto 4	ZGC activas
Distribución observada de ZGC en últimos 5 años	0.5%	16.5%	73.7%	9.3%	28.6%
Distribución promedio simulada de ZGC sin Puerto 2*	0.4%	-	91.1%	8.4%	29.8%
Desviación estándar de simulación de distribución de ZGC sin Puerto 2*	0.5	-	1.1	1.0	1.4

Tabla 9: Variación en la distribución de ZGC según su elección portuaria sin la existencia del Puerto 2.

*Los valores señalados corresponden al promedio de los 80 períodos simulados.

Respecto a la redistribución en la elección portuaria de las ZGC, se observa en la Figura 24, que la totalidad de las zonas que escogían el Puerto 2, se reasignaron al Puerto 3. En total hubo 30 ZGC que cambiaron su elección portuaria respecto a los datos observados. En este escenario un 29% de las ZGC escoge un puerto que no corresponde a su puerto más cercano, a diferencia del 36% observado en los datos históricos.

En el caso de no contar con el puerto que posee la mayor captura de ZGC, esto es, el Puerto 3, observamos en la Figura 25, que las ZGC atendidas por este puerto se redistribuyen principalmente al Puerto 1. Lo anterior da cuenta de la importancia de la distancia en la elección, ya que a pesar de que el Puerto 1 tiene un bajo nivel de servicio (representado por γ_1), la mayor distancia de la ZGC a los Puertos 2 y 4, hace que éstas prefieran el Puerto 1 por ser más cercano al puerto que deja de estar disponible.

En la Tabla 10, se observa el aumento en la captura del Puerto 1 de 0.5% a 71.3%, cuando el Puerto 3 no está activo. El Puerto 2, en tanto, también aumenta su captura, pero en menor magnitud, pasado de un 16.5% a un 20.1%. El Puerto 4, tiene una leve baja en su captura, pero esta baja es menor que la desviación estándar de los datos simulados. Lo anterior indica que la desactivación del Puerto 3 no tiene efecto en la captura de ZGC del Puerto 4.

Nuevamente vemos, que la distribución de ZGC se mantiene sin mayores variaciones en el tiempo.

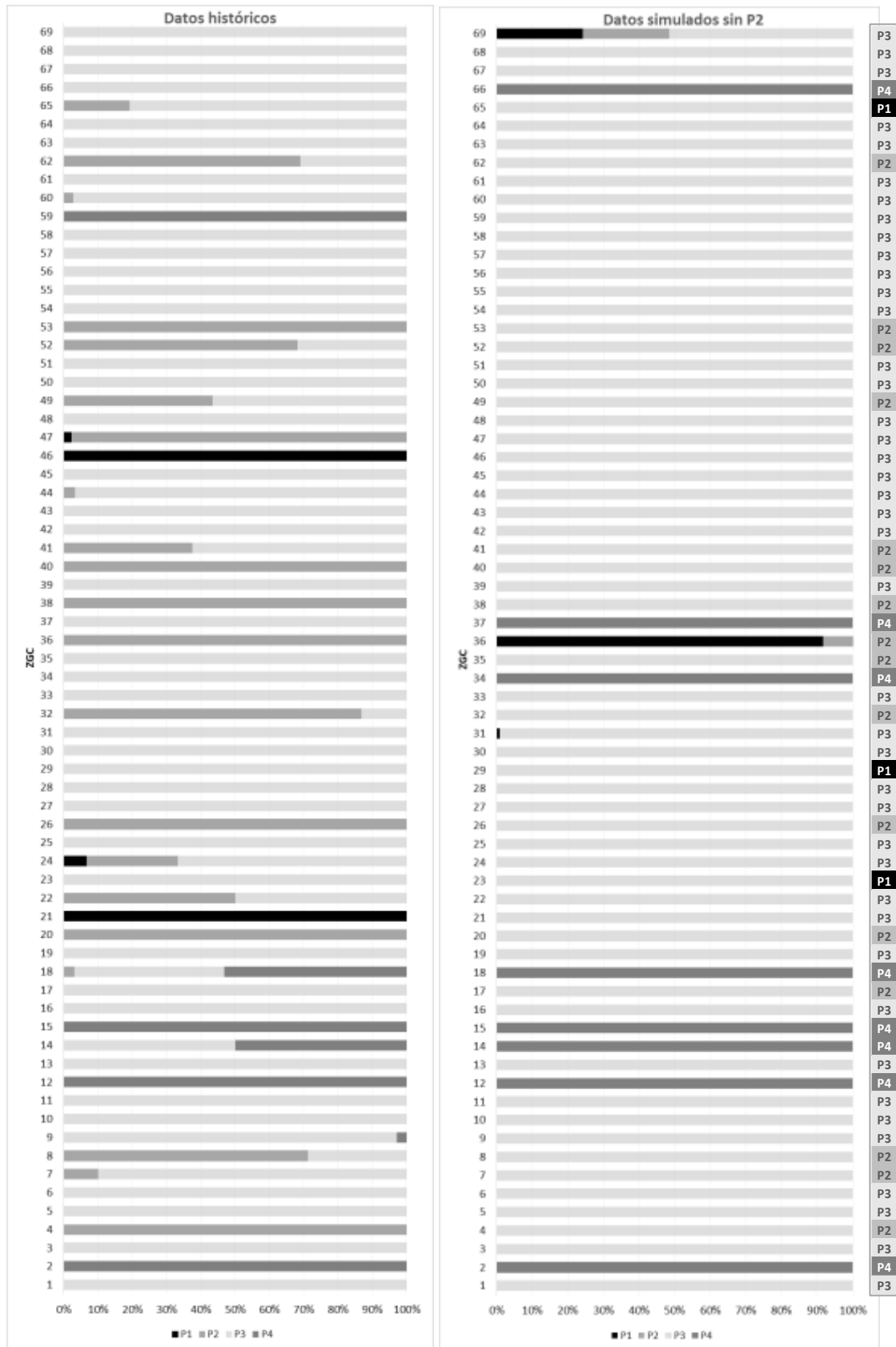


Figura 24: Comparación de la distribución histórica promedio de elección portuaria de cada ZGC con la distribución simulada de elección sin la existencia del Puerto 2.

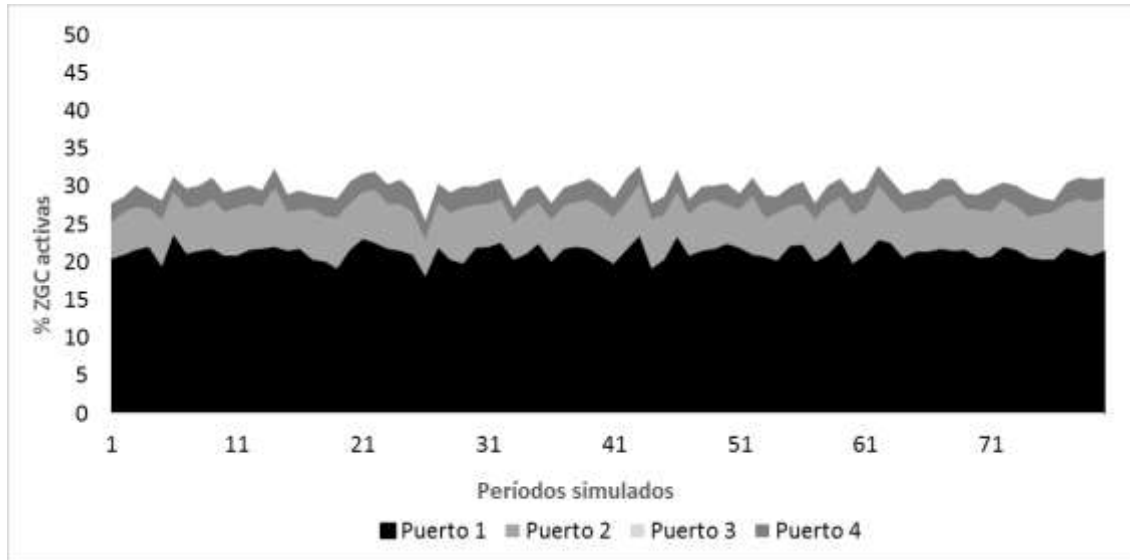


Figura 25: Distribución de ZGC según su elección portuaria sin la existencia del Puerto 3.

	Puerto 1	Puerto 2	Puerto 3	Puerto 4	ZGC activas
Distribución observada de ZGC en últimos 5 años	0.5%	16.5%	73.7%	9.3%	28.6%
Distribución promedio simulada de ZGC sin Puerto 3*	71.3%	20.1%	-	8.6%	29.8%
Desviación estándar de simulación de distribución de ZGC sin Puerto 3*	2.0	1.9	-	1.1	1.3

Tabla 10: Variación en la distribución de ZGC según su elección portuaria sin la existencia del Puerto 3.

*Los valores señalados corresponden al promedio de los 80 períodos simulados.

En la Figura 26, en tanto, es posible apreciar la variación en la elección portuaria de cada ZGC en este nuevo escenario. Como se puede observar, sólo 10 de las 69 ZGC mantuvieron su elección respecto a los datos históricos, mientras que el 38% escogió el puerto más cercano como puerto de salida de sus exportaciones.

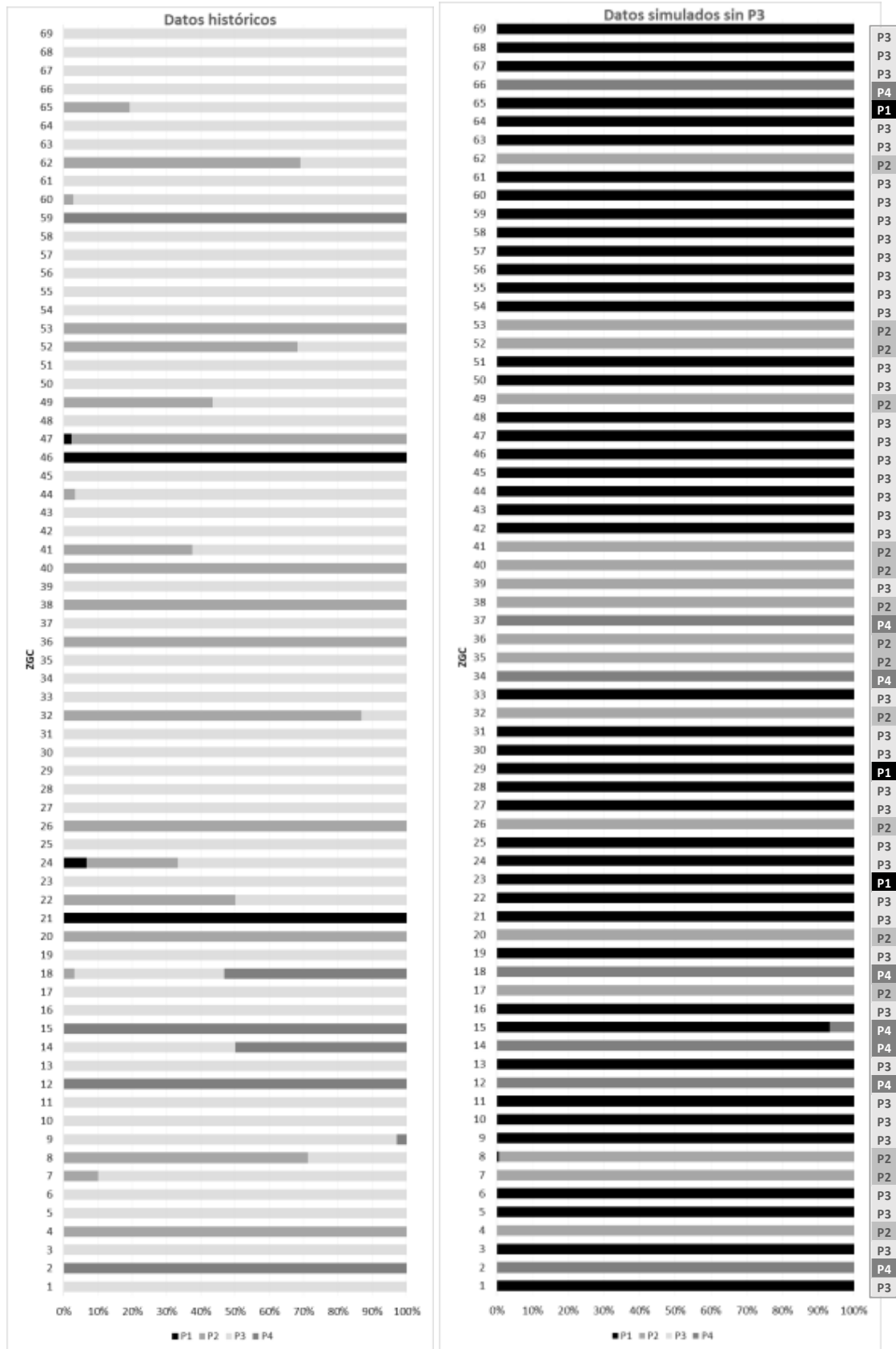


Figura 26: Comparación de la distribución histórica promedio de elección portuaria de cada ZGC con la distribución simulada de elección sin la existencia del Puerto 3.

Finalmente, si el Puerto 4 dejase de estar operativo, vemos en la Figura 27, que las ZGC escogerían al Puerto 2 y al Puerto 3 como puerta de salida de sus exportaciones. Se observa, al igual que en los casos anteriores, que, de no existir variaciones en los puertos disponibles, la distribución de ZGC se mantiene sin mayores variaciones en el tiempo.

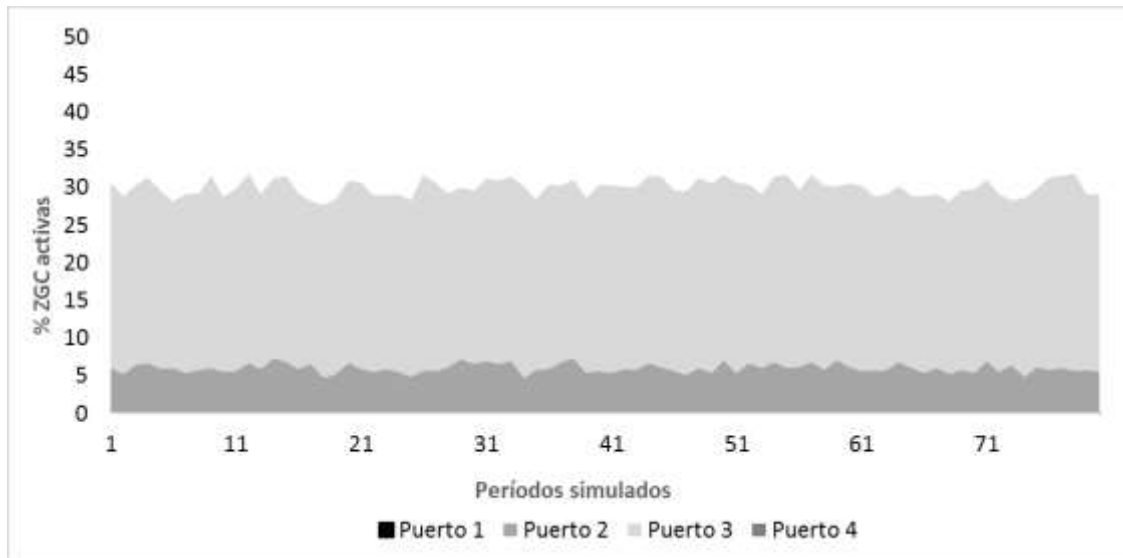


Figura 27: Distribución de ZGC según su elección portuaria sin la existencia del Puerto 4.

En la Tabla 11 se muestra el cambio en la captura de mercado de cada puerto, observándose que en el caso del Puerto 2, ésta varía de 16.5% a 20.0%. Las ZGC que escogen al Puerto 3, en tanto, aumentan de 73.7% a 80%. El Puerto 1 no tiene captura asociada, lo cual puede explicarse por el efecto que genera la vecindad en la elección portuaria. Al haber mayor cantidad de ZGC hacia un puerto, esto genera una mayor probabilidad de tener vecinos que escojan dicho puerto y por lo tanto la decisión de una ZGC que antes podría haber sido atendida por el Puerto 1, ahora cambia su elección al Puerto 2 o 3, al observar la elección de sus vecinos.

	Puerto 1	Puerto 2	Puerto 3	Puerto 4	ZGC activas
Distribución observada de ZGC en últimos 5 años	0.5%	16.5%	73.7%	9.3%	28.6%
Distribución promedio simulada de ZGC sin Puerto 4*	0.0%	20.0%	80.0%	-	30.0%
Desviación estándar de simulación de distribución de ZGC sin Puerto 4*	0.0	1.8	1.8	-	1.1

Tabla 11: Variación en la distribución de ZGC según su elección portuaria sin la existencia del Puerto 4.

*Los valores señalados corresponden al promedio de los 80 períodos simulados.

En la Figura 28, se muestra la reasignación de ZGC por puerto cuando el Puerto 4 no se encuentra disponible. En este escenario 22 de las 69 ZGC cambian su elección, la cual coincide con el puerto más cercano en 16 de ellas. Respecto al total de ZGC, 84% escoge el puerto más cercano, a diferencia de la situación histórica en que sólo lo hacía el 64%.

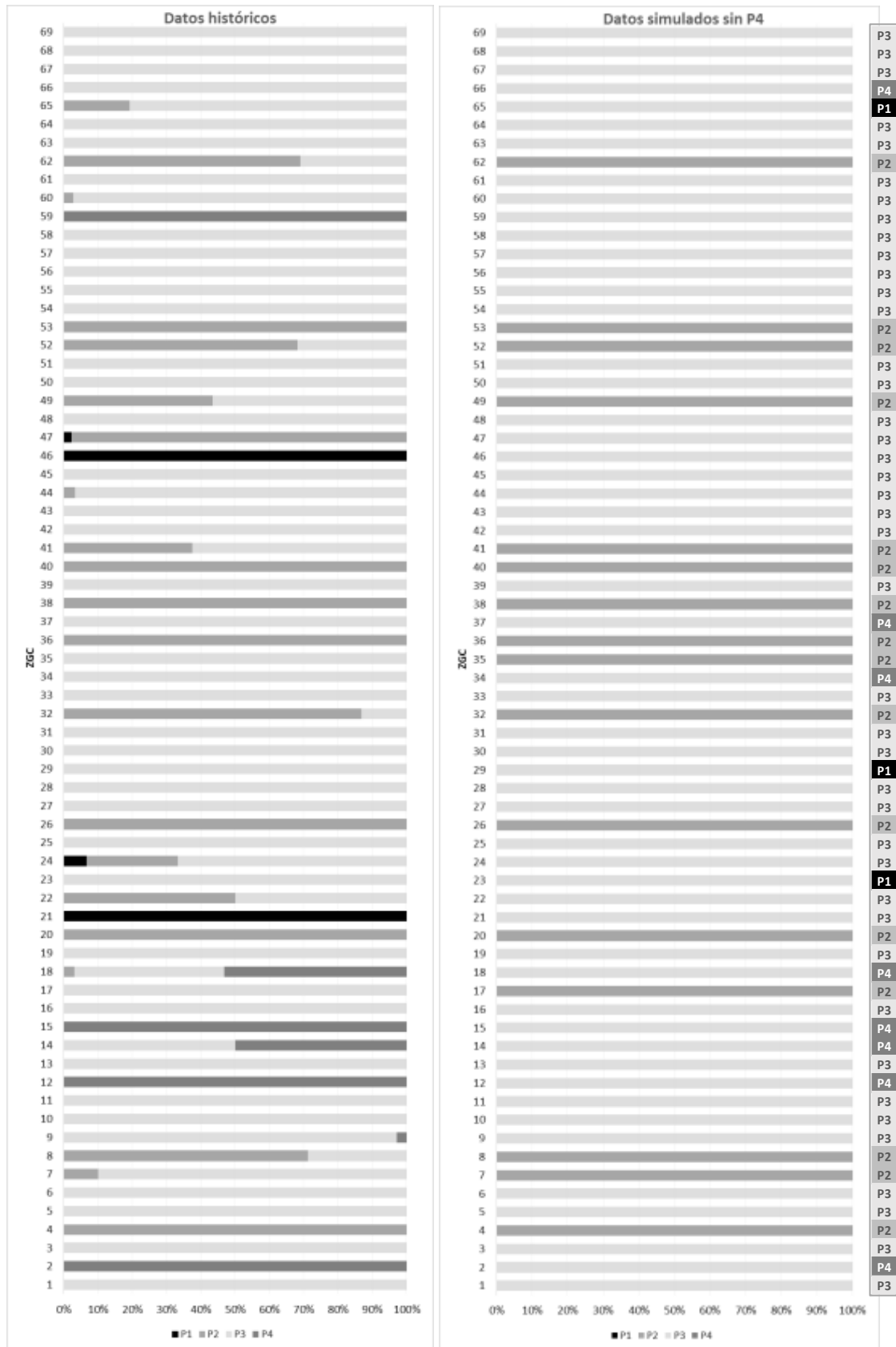


Figura 28: Comparación de la distribución histórica promedio de elección portuaria de cada ZGC con la distribución simulada de elección sin la existencia del Puerto 4.

6.3. Conclusiones

La capacidad predictiva del ACPEP fue evaluada en dos ámbitos. En el primero de ellos, el ACPEP fue aplicado asumiendo una variación en las constantes de atractividad de los puertos, mientras que, en el segundo ejercicio, se aplicó considerando una reducción de las alternativas portuarias.

En el caso de variación de las constantes de atractividad portuaria, se analizó primero el impacto en el comportamiento de las ZGC, que producía igualar la constante del puerto con menor atractividad a la del puerto más atractivo. Los resultados mostraron que, ante iguales niveles de servicio entre dos puertos alternativos, los exportadores preferirían en general al más cercano, existiendo por lo tanto una baja reasignación de ZGC.

Se estudió luego, en cuánto tenía que aumentar la atractividad del puerto menos demandado, con el fin de lograr un aumento significativo de captura de ZGC. En este análisis fue posible observar el fenómeno de interacción espacial y temporal, ya que a pesar de que se generaba un aumento significativo de la atractividad del puerto, el cambio en la elección de las ZGC no era inmediato. Lo anterior se explica por el efecto de la vecindad en la función de elección, ya que se requiere que los vecinos vayan cambiando su elección para que se vea afectada la decisión de una determina ZGC.

El comportamiento descrito concuerda con lo observado en la realidad, ya que un puerto más lejano debe ofrecer condiciones considerablemente mejores que su competencia para generar un cambio significativo en el patrón de captura de demanda. El tiempo que demore en captar más ZGC, dependerá a su vez de las ventajas que observen los exportadores respecto a los costos extras que les genere desplazarse.

El otro ejercicio realizado se enfocó en analizar el impacto que tenía en el hinterland la disminución de las alternativas de elección. Así, se pudo determinar cómo se redistribuían las ZGC entre los puertos restantes. Parte interesante de los resultados es que las ZGC no se reasignan necesariamente al puerto más cercano, mostrando, así como el ACPEP recoge la complejidad asociada a las variables no observadas.



Capítulo 7

CONCLUSIONES FINALES Y LÍNEAS DE INVESTIGACIÓN FUTURA

7. CONCLUSIONES FINALES Y LÍNEAS DE INVESTIGACIÓN FUTURA

7.1. Conclusiones

En esta tesis se desarrolló una herramienta de modelización de la captura de mercado por parte de un nodo portuario que enfrenta competencia por parte de otros nodos dentro de una red de transporte marítimo-terrestre.

Dada la complejidad del proceso de elección portuaria, en el que existen múltiples actores que interactúan entre sí y que ponderan una multiplicidad de aspectos y comportamientos espaciales que configuran una elección dinámica, se desarrolló un modelo basado en un autómata celular probabilístico que busca imitar este proceso de elección.

La complejidad del proceso estudiado quedó a la vista luego de la exhaustiva revisión bibliográfica realizada, de la que se pueden señalar principalmente los siguientes puntos:

- No existe consenso respecto al cual es el actor del proceso que tiene mayor poder de decisión. Esto dependerá de las condiciones que ocurran al momento de la elección y de múltiples factores tales como: tipo de producto, tipo y duración de contratos, tipo de movimiento (importación o exportación), tamaño del exportador/importador, entre otros. Sin embargo, se destacan a navieras,

freight forwarders y exportadores, como los principales actores del proceso, siendo los últimos los más mencionados en la literatura.

- Se reconoce que todos los actores se interrelacionan, generando un proceso que se retroalimenta, y que por lo tanto dificulta su modelización.
- Hay factores considerados en el proceso de elección portuaria, que son independientes del agente tomador de decisión. Entre estos se encuentra la localización (entendida como distancia, tiempo o costo de viaje). Otros factores que se mencionan con recurrencia, son: servicios navieros disponibles, intermodalidad, costos marítimos, costos portuarios, congestión, infraestructura, eficiencia y servicios portuarios.
- Hay efectos espaciales y temporales presentes en la decisión que son reconocidos en la literatura.
- Los modelos de elección discreta han sido la metodología de modelización más utilizada. Sin embargo, la gran mayoría de estas investigaciones no se hace cargo de las características espaciales y temporales, así como tampoco del comportamiento heterocedástico de las alternativas del conjunto de elección portuaria.
- Existe un Modelo Probit Multinomial que incorpora explícitamente efectos espaciales (representados a través de correlación serial y dependencia estadual), temporales (incorporados como dependencia estadual) y una matriz de varianza-covarianza general para el conjunto de elección. Sin embargo, a pesar de sus buenos resultados, su aplicación resultó altamente complicada debido a las dificultades matemáticas de especificación, calibración y estimación de probabilidades de elección.

Dadas las características recién descritas, el desafío fue encontrar un enfoque de modelización que diera cuenta de la complejidad del sistema estudiado, pero que permitiese una calibración y aplicación simple. Estas características son parte de lo que define la teoría de AC.

Los AC son una herramienta muy flexible para la modelización de fenómenos dinámicos, no lineales y complejos, características que fueron evidenciadas en la revisión bibliográfica del proceso de elección portuaria. Además, los AC son muy eficientes

computacionalmente, ya que se basan en reglas simples de decisión que permiten la transición del estado de una célula a otro, en función de la interacción que ésta tenga con sus vecinos. El conjunto de células forma un sistema, que evoluciona de manera compleja.

Para definir un AC, se requieren cuatro componentes:

- Unidades llamadas “células”, que en el caso de elección portuaria correspondieron a Zonas Generadoras de Carga (ZGC)-
- Un número discreto de estados, que en este caso se definió como activo por alguno de los nodos portuarioa del sistema o bien inactivo, dependiendo si la ZGC exportaba carga por alguno de los puertos del conjunto de elección.
- Una vecindad, que fue definida como un círculo de radio R , dentro del cual dos o más ZGC se consideran vecinas.
- Una regla de transición, que indica las condiciones que se deben cumplir para que la ZGC cambie de estado.

Como parte de esta última regla, se definió una función de transición que contenía 3 componentes: una asociada a la interacción espacial de la ZGC con sus vecinos, en función de la elección de puerto realizada por sus vecinos en el período anterior; otra referida al peso de la distancia que implica desplazar la mercancía desde la ZGC a cada puerto del conjunto de elección; y una tercera, correspondiente a una constante de atractividad propia de cada puerto.

Por otra parte, dado lo compleja de la decisión, y asumiendo que un modelador no puede describir completamente el proceso de elección portuaria sólo con un conjunto de atributos, se incorporó una componente estocástica, que explica al menos dos fenómenos: el modelador no tiene información perfecta, y que una ZGC puede activarse/desactivarse por variables exógenas que no están incluidas en la especificación del modelo.

Todas estas componentes definieron el llamado Autómata Celular Probabilístico de Elección Portuaria (ACPEP), el cual fue calibrado con un algoritmo genético.

Para determinar las ventajas de esta estructura de modelización, sobre, por ejemplo, los modelos de elección discreta, se realizó una modificación en la especificación del modelo Probit Multinomial antes señalado. Sin embargo, al incorporar la correlación espacial entre las ZGC, resultó una matriz de varianza covarianza mucho más grande que cualquiera de las que se ha publicado hasta el momento para la calibración de este tipo de modelos. Lo anterior dio cuenta que la estructura espacio-temporal representada por el ACPEP no es factible de incorporar en modelos de elección discreta.

Para probar el ACPEP, éste fue aplicado al caso de exportaciones de vehículos de Brasil, considerándose el período comprendido entre los años 1997 y 2015. Los datos fueron tratados de manera trimestral, por lo que se contó con una serie de 74 períodos.

Las ZGC, en tanto, correspondieron a municipios que habían exportado vehículos en el período antes mencionado por alguno de los puertos considerados como alternativas de elección.

Con el fin de asegurar que las ZGC contaran con el mismo conjunto de alternativas, se escogieron puertos de uso público y con un hinterland común. Los puertos que cumplían estas condiciones fueron: Paranaguá, Santos, San Sebastián y Rio de Janeiro. El puerto de San Sebastián, a pesar de poseer un muy bajo movimiento de vehículos, se decidió mantener con el fin de probar la capacidad del modelo frente a una alternativa que pudiese mejorar sus condiciones de operación.

Así, la instancia de aplicación del ACPEP, quedó definida por 69 ZGC, 4 puertos y 74 períodos de observación.

Una de las características de los datos, es que en promedio sólo un 30% de las ZGC se mantenía activa. Esto en general implica un desafío mayor en la calibración de modelos de elección discreta, ya que estos no responden adecuadamente cuando se tiene un bajo número de observaciones activas. Sin embargo, los AC pueden representar fácilmente esta dinámica.

Respecto a la distribución histórica de las ZGC a los puertos, y concordante con la revisión bibliográfica analizada, se detectó que un porcentaje importante de ZGC escogía el puerto más cercano a ella. Sin embargo, hubo un 36% que no siguió este

comportamiento, dando cuenta así de la existencia de otros factores que afectan este proceso de elección.

En cuanto a la calibración del modelo, se obtuvo un porcentaje de éxito del 73%, siendo superior incluso a muchos modelos de elección discreta que han tratado el proceso de elección portuaria.

Los resultados confirmaron la hipótesis realizada respecto al peso de la interacción espacial entre las ZGC, tanto por el parámetro obtenido que la representa dentro del ACPEP, como por el radio resultante de la vecindad con la cual cada ZGC analiza la decisión de sus vecinos.

Respecto a los parámetros que definen la probabilidad de activación de las ZGC, se observaron valores que indican una relativa resistencia de activación de las ZGC, siendo consistente con lo observado en los datos reales.

Las constantes de atractividad de los puertos, en tanto, tuvieron signo negativo tal como era esperado, ya que la constante del puerto que atraía mayor número de ZGC fue fijada en cero para evitar problemas de identificabilidad. De igual forma, se obtuvo signo correcto para el parámetro de la variable distancia.

Las simulaciones realizadas con el modelo mostraron que el ACPEP representaba adecuadamente la dinámica del sistema, obteniendo participaciones de mercado y porcentajes de ZGC activas similares a las observadas en los datos reales.

Respecto a la capacidad predictiva del ACPEP, esta fue evaluada en dos ámbitos. En el primero de ellos, el ACPEP fue aplicado asumiendo una variación en las constantes de atractividad de los puertos, mientras que, en el segundo ejercicio, se aplicó considerando una reducción de las alternativas portuarias.

En el caso de variación de las constantes de atractividad portuaria, se estudió, por ejemplo, en cuánto tenía que aumentar la atractividad del puerto menos demandado, con el fin de lograr un aumento significativo de captura de ZGC. En ese análisis fue posible observar cómo operaba el fenómeno de interacción espacial y temporal, ya que a pesar de que se generaba un aumento significativo de la atractividad del puerto, el cambio en la elección de las ZGC no era inmediato. Lo anterior se producía

principalmente por el efecto de la vecindad en la función de elección, ya que el cambio en la elección de una ZGC, dependía de que una proporción importante de sus vecinos realizara ese mismo cambio en períodos anteriores.

El comportamiento descrito concuerda con lo observado en la realidad, ya que un puerto más lejano debe ofrecer condiciones considerablemente mejores que su competencia para generar un cambio significativo en el patrón de captura de demanda. El tiempo que demore en captar más ZGC, dependerá a su vez de las ventajas que observen los exportadores respecto a los costos extras que les genere desplazarse.

Otro ejemplo de aplicación del ACPEP fue el análisis del impacto que tenía en el hinterland contar con un puerto menos dentro del set de elección, pudiendo observarse cómo se reasignaban en el tiempo las ZGC en los puertos restantes.

Este último análisis es clave para poder medir, por ejemplo, los efectos en la red de transporte producto de la reasignación de ZGC al incorporarse o cerrarse un nodo portuario.

En resumen, el ACPEP propuesto, tiene importantes implicancias tanto desde el punto de vista teórico como práctico. Desde el punto de vista teórico, este modelo ayuda a determinar patrones complejos (comportamiento caótico, sensibilidad a las condiciones iniciales, etc.) en la dinámica del hinterland producto de la simulación que logra del proceso de elección portuaria con sus características espaciales y temporales, enfoque que hasta ahora había sido abordado de manera parcial y con metodologías de alta complejidad.

En la dimensión práctica, constituye una valiosa herramienta para medir el impacto que tiene sobre el hinterland medidas que varíen los niveles de servicio portuarios, o más aún, medir el impacto de la apertura o cierre de nuevas instalaciones portuarias. Todo esto es información clave para evaluar, por ejemplo, el desempeño de políticas públicas que influyan sobre las condiciones de atractividad de los puertos que poseen un hinterland común.

7.2. Líneas Futuras de Investigación

A partir de los trabajos realizados en esta tesis quedan abiertas diferentes líneas de investigación que se detallan a continuación:

- Incorporar al modelo propuesto una diferenciación entre tamaño de exportadores, con el fin de identificar patrones de comportamiento entre ellos.
- Extender el modelo a una dimensión multiproducto. Esto último requeriría la búsqueda de una forma alternativa de vincular que permita capturar la interacción entre exportadores de diversos productos.
- Incorporar, en función de la información disponible, variables portuarias como, número de servicios navieros disponibles, tarifas, niveles de servicio, etc. que permitan explicar de manera más desagregada el proceso de elección portuaria, facilitando así la evaluación de acciones que repercuten en dichas variables.
- Incorporar al modelo propuesto, restricciones de capacidad del sistema, con el fin de medir el impacto de la saturación de las distintas componentes de la red en la elección portuaria.
- Dado el efecto que tiene la intermodalidad en el tramo terrestre de la cadena de suministro, sería de interés también definir una estructura de autómata celular que considere su efecto la elección portuaria.



REFERENCIAS

REFERENCIAS

- Aguilera, F. (2006). Predicción del crecimiento urbano mediante sistemas de información geográfica y modelos basados en autómatas celulares. *GeoFocus* 6. pp. 81-112.
- Almeida, C., Gleriani, J., Castejon, E. y Soares-Filho, B. (2008). Using neural networks and cellular automata for modelling intra-urban land-use dynamics. *International Journal Of Geographical Information Science* 22 (9). pp. 943-963.
- Anfavea (2016). Anuário da Indústria Automobilística Brasileira. Brasil
- Aronietis, R., Van de Voorde, E. y Vanelslander, T. (2010). Port Competitiveness Determinants of Selected European Ports in the Containerized Cargo Market. Paper presented at IAME 2010.
- Barredo, J., Kasanko, M., McCormick, N. y Lavalle, C. (2003). Modelling dynamic spatial processes: simulation of urban future scenarios through cellular automata. *Landscape and Urban Planning* 64. pp. 145–160.
- Ben-Akiva, M. y Lerman, S.R. (1985). Discrete choice analysis: theory and application to travel demand. The MIT Press, Cambridge, Massachusetts.
- Breukelaar, R. y Bäck, Th. (2005). Using a genetic algorithm to evolve behavior in multidimensional cellular automata: emergence of behavior. Proceeding GECCO '05 Proceedings of the 7th annual conference on Genetic and evolutionary computation. Washington DC, USA.
- Bunch, D. (1991). Estimability in the multinomial probit model. *Transportation Research* 25B. pp. 1-12.
- Chang, Y-T, Lee, S.Y. y Tongzon, J.L. (2008). Port selection factors by shipping lines: Different perspectives between trunk liners and feeder service providers. *Marine Policy* 32 (6). pp 877–885.
- Chopard, B. y Droz, M. (2005). Cellular Automata Modeling of Physical Systems. Cambridge University Press. Cambridge.



- De Langen, P.W. (2007). Port competition and selection in contestable hinterlands: the case of Austria. *European Journal of Transport and Infrastructure Research* 7 (1). pp. 1-14.
- García-Alonso, L. y Sánchez Soriano, J. (2009). Port selection from a hinterland perspective. *Maritime Economics and Logistics* 11. pp 260–269.
- Gardner, M. (1970). Mathematical Games: The fantastic combinations of John Conway's new solitaire game "life". *Scientific American* 223. pp. 120-123.
- Garrido, R. (1997). Analysis of Spatial and Temporal Characteristics of Freight Demand. Ph. D. Dissertation, Department of Civil Engineering, University of Texas at Austin.
- Garrido, R. y Mahmassani, H. (2000). Forecasting freight transportation demand with the space–time multinomial probit model. *Transportation Research Part B: Methodological* 34 (5). pp. 403–418.
- Garrido, R. y Leva, M. (2004). Port of destination and carrier selection for fruit exports: a multi-dimensional space-time multinomial probit model. *Transportation Research Part B* 38 (7). pp. 657-667.
- Goldberg, D. (1989). Genetic Algorithms in Search. Optimization and Machine Learning. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.
- Goles, E. y Martínez, S. (1998). Cellular Automata and Complex Systems. Nonlinear Phenomena and Complex Systems Series. Kluwer Academic.
- Griffeath, D. y Moore, C. (2003). New Constructions in Cellular Automata. Oxford University Press, Oxford.
- Guy, E. y Urli, B. (2006). Port Selection and Multicriteria Analysis: An Application to the Montreal-New York Alternative. *Maritime Economics and Logistics* 8 (2). pp. 169-186.
- Ha, M.S. (2003). A comparison of service quality at major container ports: implications for Korean ports. *Journal of Transport Geography* 11 (2). pp. 131-137.
- Halim, R. A., Kwakkel, J. H. y Tavasszy, L. A. (2016). A strategic model of port-hinterland freight distribution networks. *Transportation Research Part E: Logistics and Transportation Review* 95. pp. 368-384.

- Hausman, J. y Mcfadden, D. (1984). Specification Tests for the Multinomial Logit Model. *Econometrica* 52 (5). pp. 1219-1240.
- Jiang, J., Lee, L.H., Chew, E.P. y Gan, C.C. (2015). Port Connectivity Study: An Analysis Framework from a Global Container Liner Shipping Network Perspective. *Transportation Research Part E* 73. pp. 47-64.
- Kramberger, T., Rupnik, B., Štrubelj, G. y Prah. K. (2015). Port Hinterland Modelling Based on Port Choice. *PROMET – Traffic&Transportation* 27 (3). pp. 195-203.
- Lee, S-Y, Chang, Y-T y Tae-Woo, P. (2010). Container port selection factors: heterogeneity among major market player. *Journal of International Logistics and Trade* 8 (2). pp. 73-90.
- Leva, M. (2002). Caracterización de la Toma de Decisiones de Exportadores Frutícolas: Un Enfoque Espacio-Temporal. Tesis de Magíster, Departamento de Ingeniería de Transporte, Pontificia Universidad Católica de Chile.
- Li, X., Yang, Q. y Liu, X. (2007). *Science in China Series D: Earth Sciences* 50 (12). pp. 1857–1866.
- Lirn, T.C., Thanopoulou, H.A., Beynon, M.J. y Beresford, A.K.C. (2004). An Application of AHP on Transshipment Port Selection: A Global Perspective. *Maritime Economics and Logistics* 6 (1). pp. 70-91.
- McFadden, D. (1974) Conditional Logit Analysis of Qualitative Choice Behaviour. *Frontiers in Econometrics*, In: P. Zarembka (ed.) New York: Academic Press.
- Magala, M. y Sammons, A. (2008). A New Approach to Port Choice Modelling. *Maritime Economics and Logistics* 10. pp. 9–34.
- Malchow, M. y Kanafani, A. (2004). A Disaggregate Analysis of Port Selection. *Transportation Research Part E* 40. pp. 317-337
- Mangan, J., Lalwani, C. y Gardner, B. (2002). Modelling port/ferry choice in RoRo freight transportation. *International Journal of Transport Management* 1. pp. 15-28
- Manski, Ch. y Mcfadden, D. (1981). Structural Analysis of Discrete Data and Econometric Applications. Charles F. and Daniel L. Editors. Cambridge: The MIT Press.

- McMillen, D. (1992). Probit with spatial autocorrelation. *Journal of Regional Science* 32. pp. 335-348.
- Meersman, H., Pauwels, T., Van de Voorde, E., y Vanellander, T. (2008). The relation between port competition and hinterland connections: the case of the 'Iron Rhine' and the 'Betuweroute'. Conference proceedings of the International Forum on Shipping. Ports and Airports (IFSPA).
- Meersman, H., Van de Voorde, E. y Vanellander, T. (2010). Port Competition Revisited. *Review of Business and Economics* 55 (2). pp. 210-233.
- Mitchell, M., Crutchfield, J.P., y Das, R. (1996). Evolving cellular automata with genetic algorithms: A review of recent work. In Proceedings of the First International Conference on Evolutionary Computation and Its Applications (EvCA'96). Russian Academy of Sciences.
- Mohring, H. (1972). Optimization and Scale Economies in Urban Bus Transportation. *American Economic Review*. pp. 591-604.
- Moya, J.M. y Feo-Valero, M. (2016). Port choice in container market: a literature review, *Transport Reviews*. <http://dx.doi.org/10.1080/01441647.2016.1231233>.
- Ng, A. S-F., Sun, D. y Bhattacharjya, J. (2013). Port choice of shipping lines and shippers in Australia. *Asian Geographer* 30 (2). pp. 143-168.
- Nir, A.S., Lin, K. y Liang, G.S. (2003). Port choice behaviour – from the perspective of the shipper. *Maritime policy & Management* 30 (2). pp. 165-173.
- Papola, A. (2004). Some developments on the cross-nested logit model. *Transportation Research Part B: Methodological* 38 (9). pp. 833–851.
- Santé, I., García, A., Miranda, D. y Crecente, R. (2010). Cellular automata models for the simulation of real-world urban processes: A review and analysis. *Landscape and Urban Planning* 96 (2). pp. 108–122.
- Slack, B. (1985). Containerization, inter-port competition and port selection. *Maritime Policy and Management* 12 (4). pp. 293–303.

- Song, D-P, Lyons, A., Li, D. y Sharifi, H. (2016). Modeling port competition from a transport chain perspective. *Transportation Research Part E: Logistics and Transportation Review* 87. pp 75–96.
- Song, D-W y Yeo, K-T. (2004). A competitive Analysis of Chinese Container Ports Using the Analytic Hierarchy Process. *Maritime Economics and Logistics* 6. pp.34–52.
- Steven, A.B. y Corsi, T.M. (2012). Choosing a port: An analysis of containerized imports into the US. *Transportation Research Part E: Logistics and Transportation Review* 48 (4). pp. 881- 895.
- Tiwari, P., Itoh. H. y Doi, M. (2003). Shippers’ Port and Carrier Selection Behaviour in China: A Discrete Choice Analysis. *Maritime Economics and Logistics* 5. pp.23–39.
- Tongzon, J. y Sawant, L. (2007). Port choice in a competitive environment: from the shipping lines’ perspective. *Applied Economics* 39. pp. 477–92.
- Tongzon, J. (2009). Port choice and freight forwarders. *Transportation Research Part E: Logistics and Transportation Review* 45 (1). pp. 186–195.
- Train, K. (2002). Discrete Choice Methods with Simulation. Cambridge University Press. Cambridge.
- Ugboma, C., Ugboma, O., y Ogwude, I.C. (2006). An Analytic Hierarchy Process (AHP) Approach to Port Selection Decisions - Empirical Evidence from Nigerian Ports. *Maritime Economics and Logistics* 8 (3). 251-266.
- Van Canneyt, M., Gartner, M., Heinig, S., Monteiro de Cavalho, F., Ouedraogo, I., Overbeek, D. (2011). Lazarus, the Complete Guide. Blaise Pascal Magazine. ISBN: 978-94-90968-02-1.
- Vázquez, J. y Oliver, J. (2008). Evolución de autómatas celulares utilizando algoritmos genéticos. Universidad de Deusto. Spain
- Vega, L., Cantillo, V., y Arellana, J. (2014). Modelización de la elección de puerto desde una perspectiva desagregada: un análisis para los puertos colombianos. XVII Congreso Panamericano de Ingeniería del tránsito, transporte y logística. Santander (España).



- Veldman, S., García-Alonso, L. y Vallejo-Pinto, J. (2011). Determinants of container port choice in Spain. *Maritime Policy and Management* 38 (5). pp. 509–522.
- Veldman, S., García-Alonso, L. y Vallejo, J. (2013). A port choice model with logit models: A case study for the Spanish container trade. *International Journal of Shipping and Transport Logistics* 5. pp. 373-389.
- Veldman, S., García-Alonso, L. y Liu, M. (2015). Testing port choice models using physical and monetary data: a comparative case study for the Spanish container trades. *Maritime Policy & Management* 3 (4). pp. 495-508.
- Von Neumann, J. y Burks, A. W. (1966). Theory of self-reproducing automata. Urbana, University of Illinois Press.
- Wang, X., Meng, Q. y Miaoc, L. (2016). Delimiting port hinterlands based on intermodal network flows: Model and algorithm. *Transportation Research Part E: Logistics and Transportation Review* 88. pp. 32-51.
- Weigend, G.G., (1958). Some elements in the study of port geography. *Geographical Review* 48. pp. 185-200.
- Wiegmans, B. W., Van Der Hoest, A., y Notteboom, T. E. (2008). Port and terminal selection by deep sea container operators. *Maritime Policy and Management* 35(6). pp. 517–534.
- Williams, H.C.W.L. y Ortúzar, J. de D. (1982). Behavioural theories of dispersion and the misspecification of travel demand models. *Transportation Research* 16 B. pp. 167-219.
- Woo, S-H. (2011). Seaport research: a structured literature review on methodological issues since the 1980s. *Transportation Research Part A: Policy and Practice* 45 (7). pp.667–685.
- Yeo, G-T., Ng, A.K.Y., Lee, P. T-W y Yang, Z. (2014). Modelling port choice in an uncertain environment. *Maritime Policy & Management* 41(3). pp. 251-267.

ANEXOS

Anexo A: Modelización de elección portuaria: Un enfoque de autómatas celulares probabilísticos

RESUMEN

La modelización de la elección de puertos por parte de exportadores de carga, suele realizarse asumiendo que esta elección depende esencialmente de los costos de transporte entre el puerto y el origen o destino final dentro del hinterland. Sin embargo, este proceso de elección está dominado por una multiplicidad de variables y actores, que dificultan su modelización con técnicas estándar como los modelos de elección discreta.

Una de las razones que explican la incapacidad de estos métodos para predecir el comportamiento futuro es la característica de sistema complejo que se observa en este proceso. En efecto, la elección de puerto es el resultado de la interacción de muchos actores a través de reglas relativamente simples pero que en conjunto evolucionan en forma difícil de predecir.

En esta investigación se utiliza una herramienta que ha demostrado ser de una enorme capacidad predictiva y descriptiva de sistemas complejos de distinta naturaleza: la teoría de autómatas celulares (AC). Estos modelos logran captar con facilidad interacciones espacio-temporales que en otros contextos de modelización son extremadamente difíciles de incorporar y calibrar.

En este artículo se presenta un modelo de asignación portuaria basado en un autómata celular probabilístico, que busca imitar la decisión de múltiples actores que interactúan entre sí y que ponderan una multiplicidad de aspectos y comportamientos espaciales que configuran una elección dinámica, sin necesidad de conocer explícitamente a todos los actores o a todos los determinantes de un proceso discreto que cambia en el tiempo.

La especificación del autómata celular diseñado es aplicada a un caso práctico de exportación de automóviles en Brasil y se calibra utilizando un algoritmo genético.

Los resultados obtenidos muestran que el modelo logra replicar con un alto porcentaje de aciertos el comportamiento dinámico de la elección, obteniéndose que el parámetro espacial, asociado a la componente de vecindad del autómata, es significativo al igual que el correspondiente a las características y ubicación de los puertos analizados.

1. INTRODUCCIÓN

1.1 Motivación

Más del 85% de las mercancías movilizadas globalmente utilizan el transporte marítimo como modo principal. Esta característica parece ser una tendencia de largo plazo al observar los tamaños de naves que se utilizan para el transporte marítimo. Estas tendencias han ido de la mano con mejoramientos en la capacidad de transferencia de los puertos y en la profesionalización de todos los servicios requeridos por estas embarcaciones.

Dado el escenario anterior, la planificación portuaria hoy en día requiere de un estudio prospectivo de la utilización de los distintos nodos portuarios de una red para mantener un balance eficiente de capacidad de transferencia, minimizar costos y mantener las externalidades negativas bajo control.

La proyección de demanda por transporte de mercancías, fundamental para la planificación portuaria, es típicamente incorporada en los programas maestros de desarrollo de los puertos. Tradicionalmente, en la práctica se realiza un estudio del mercado potencial del puerto en función de los distintos orígenes/destinos de carga de los puertos considerados como competencia. Así, la demanda de estos puertos se proyecta en base a tasas mundiales de crecimiento del PIB para un cierto horizonte de planificación.

La distribución de mercancías entre los distintos puertos que compiten suele realizarse de acuerdo a un modelo que asume que la elección de puerto depende esencialmente de los costos (generalizado) de transporte de mercancías entre el puerto y el origen o destino final dentro del hinterland. De esta forma se estima un porcentaje de captura que tendría el puerto bajo estudio sobre los flujos totales de los puertos de la red relevante. Estos valores son luego sensibilizados con distintos factores de captura.

Hay otros estudios, en que el enfoque de elección es realizado desde el punto de vista de la naviera y no del exportador de carga. En estos casos, la tarifa y variables de servicio de los puertos, se vuelven las variables fundamentales.

En ambos casos, estas son sólo algunas de las variables que entran en juego en la elección. En efecto, variables tales como la capacidad negociadora, especialización en un tipo de mercancía, contratos de largo plazo, integración comercial, son factores relevantes en la captura del mercado y que son de muy difícil predicción.

1.2 Descripción del problema

El problema a tratar consiste en el desarrollo de un modelo de elección portuaria por parte de los exportadores, en el que se consideran variables vinculadas a los puertos correspondientes al conjunto de elección, así como la influencia que tiene la decisión de un exportador sobre otro que se encuentra espacialmente cercano (interacción espacial).

La hipótesis referente a la influencia que tienen las decisiones de los exportadores entre sí, se respalda en la evidencia empírica, respecto a que exportadores de mayor tamaño influyen en la decisión de puerto de recalada de las naves, generando así condiciones de servicios nuevos y mayor capacidad de negociación con las navieras.

Al gatillarse nuevos servicios navieros, se generan a su vez mejores condiciones de negociación para exportadores más pequeños y por lo tanto, muchos de ellos, seguirán el comportamiento de sus vecinos, escogiendo puertos similares. Lo anterior se acompaña a su vez, con que la ubicación espacial de los exportadores se vincula con ciertas actividades económicas similares y por lo tanto las necesidades portuarias y navieras probablemente serán compartidas.

1.3 Antecedentes

Respecto al comportamiento en la elección portuaria, la mayoría de la literatura menciona aspectos como los costos de transporte hacia/desde el hinterland y la confiabilidad del recorrido total en un esquema de cadena de suministro más que de la elección por sí sola. Hay también estudios que muestran que en la práctica hay varios otros aspectos que determinan la elección portuaria (al menos siete según Chang et al, 2008). Otros estudios realzan la gran variabilidad por tipo de producto en esta elección (Malchow y Kanafani, 2001). Por otra parte, Tongzon (2009) señala que en muchos casos es el *freight forwarder* quien decide el puerto de recalada, en un complicado proceso de dos etapas.

En cuanto a los modelos de elección portuaria, existen estudios con diversas metodologías, tales como redes neuronales, elección discreta, entre otras. La mayoría de los estudios basados en Redes Neuronales Artificiales (RNA) se focalizan en aspectos técnicos del modelo tal como su construcción, al algoritmo de entrenamiento o su capacidad predictiva comparada. En el contexto específico de modelización de demanda portuaria hay algunos ejemplos de aplicación. Al-Deek (2001) compara un modelo de RNA (back-propagation) con uno de regresión lineal para discernir qué método es más eficiente en modelos de generación de demanda por carga portuaria.

Utiliza datos del Puerto de Miami, Florida y concluye que si bien el modelo de RNA es más preciso, necesita una mayor cantidad de datos que podría ser difícil obtener. Rodríguez et al (2013) analizan varios modelos en esta área, concluyendo que, entre los modelos de serie de tiempo, regresión lineal múltiple y RNA, estos últimos presentan la mayor eficacia predictiva. Weiqun y Nuo (2003) utilizan un modelo de RNA para predecir la carga movilizada en el puerto de Dalian, China.

Lam et al (2004) desarrollan modelos de RNA para predecir la carga movilizada por el puerto de Hong Kong. Trabajan con datos históricos (1983-2000) de los movimientos de carga y variables explicativas de su modelo. Obtuvieron buenos resultados al comparar su desempeño con otros modelos alternativos de regresión lineal múltiple típicamente utilizados por el puerto con anterioridad.

En cuanto a los modelos de elección discreta, existe un gran número de enfoques de modelización basados en el modelo Logit Multinomial (por ejemplo, Nir et al., 2003; Magala y Sammons, 2008) pero que tienen el problema serio debido a correlación entre alternativas y/o heteroscedasticidad de las alternativas de elección. En efecto, el modelo logit supone como base fundamental que ninguna alternativa está correlacionada entre sí y además, la varianza de los errores aleatorios del modelo es igual para todos ellos. Ninguna de esas dos características se cumple en el caso portuario, ya que los puertos presentan correlación entre sí ya sea por cercanía (correlación espacial) o por contratos de largo plazo con navieras o agentes exportadores (importadores), que generan correlación temporal en las decisiones; por otra parte, hay alta variación entre las alternativas portuarias lo que hace muy difícil la existencia de homocedasticidad en las varianzas de los términos de error de los modelos.

Hay pocos estudios publicados en revistas especializadas que utilicen un modelo de elección discreta que no requiera del cumplimiento de estas dos restricciones. Uno de ellos resuelve parcialmente el problema, a través de un modelo logit jerárquico, que si bien no puede trabajar correlaciones generales si puede considerar heterocedasticidad (ver Veldman et al., 2013). Tal vez el único modelo que incorpora explícitamente ambos elementos es el de Garrido y Leva (2004). Allí, los autores presentan un modelo probit multinomial que incorpora correlaciones, espaciales, temporales y una matriz de varianza-covarianza general para el conjunto de elecciones. Si bien ese modelo entrega buenos resultados y representa correctamente la interacción y decisiones de agentes, su uso es complicado por las dificultades matemáticas de calibración y estimación de probabilidades de elección, lo que lo hace poco transferible a la práctica de no-expertos en modelización.

Existen otras metodologías más innovadoras, tal como la presentada por Petri, et al (2014) quienes desarrollaron su propia estructura de modelización de demanda por transporte, motivados por la constatación de las diferencias que se constataban al utilizar los modelos más comunes (y aceptados a nivel práctico) con una misma base de datos. Esto implicaría que los resultados serían fuertemente dependientes de la elección del modelo a utilizar y por ende estos autores decidieron desarrollar una estructura de modelización que no tuviese este problema. Para ello se apoyaron en las llamadas Redes Bayesianas, que son grafos que contienen probabilidades condicionales en sus arcos y que permiten realizar inferencia en casos con pocos datos disponibles, alta variabilidad y complejidad (como el caso de la demanda por transporte de carga portuaria). Su estructura utiliza metodologías de minería de datos (*data mining*) en un esquema de dos fases. En la primera fase hay un pre-procesamiento de los datos para discernir qué variables podrían ser candidatas a pertenecer a un modelo predictivo. Esto lo realizan con la ayuda de árboles de decisión que analizan frecuencias de datos para proponer un conjunto de variables explicativas que en una fase posterior alimentan a una red Bayesiana que es calibrada con datos disponibles en la unión Europea. No es claro sin embargo, que esa metodología logre definitivamente evitar el problema de la dependencia de los resultados con la elección del tipo de modelo a utilizar.

Yeo, et al (2014) proponen un modelo conceptual de elección de puertos basado en lógica difusa. En este enfoque, diversas fuentes (y formatos) de datos, tanto objetivos como subjetivos, se combinan para evaluar el desempeño de un puerto, determinado a través de la definición de una “puntuación difusa” definida usando grados de creencia, que se vinculan con el razonamiento probatorio para producir una puntuación de preferencia para establecer criterios de elección de puerto. El método se aplica para analizar la elección de los principales puertos de contenedores del noreste asiático, teniendo como agente de decisión a una compañía naviera. El resultado de esta aplicación es una puntuación de preferencia para elegir un puerto, permitiendo comparar cada puerto ya sea con el conjunto completo de características de elección o con cada una de ellas en particular, lo que permite además establecer medidas de desempeño individual de un mismo puerto respecto a sus propias características en periodos anteriores.

En cuanto a la aplicación de Autómatas Celulares como herramienta de modelización de elección portuaria, no se encontró bibliografía específica al caso estudiado. Sin embargo, tal como se muestra en Goles y Martínez (1998) este tipo de herramienta permite modelar fenómenos complejos y no lineales (como es la demanda por servicios portuarios) con gran eficiencia computacional y uso de información disponible.

Un AC es un modelo matemático del campo de la física computacional, propuesto inicialmente por John von Neumann en la década del 1950 para modelar y describir sistemas complejos como una colección masiva de objetos simples que interactúan localmente unos con otros. Son modelos muy adecuados para estudiar sistemas dinámicos que evolucionan en pasos discretos (ver por ejemplo Griffeath y Moore, 2003; Chopard y Droz, 2005).

1.4 Objetivos

Esta investigación tiene por objetivo desarrollar una nueva herramienta de modelización de captura de mercancías por parte de un nodo portuario que enfrenta competencia por parte de otros nodos dentro de una red de transporte marítimo-terrestre.

Para esto se propone utilizar una herramienta que ha demostrado ser de una enorme capacidad predictiva y descriptiva de fenómenos de distinta naturaleza: la teoría de autómatas celulares (AC). Principalmente, estas herramientas logran captar con facilidad interacciones espacio-temporales que en otros contextos de modelización son extremadamente difíciles de incorporar y calibrar (Garrido y Leva, 2004). Se postula aquí, que la elección portuaria es una decisión de múltiples actores que interactúan entre sí y que ponderan una multiplicidad de aspectos y comportamientos espaciales que configuran una elección dinámica que un AC puede imitar sin necesidad de conocer a todos los actores o a todos los determinantes de un proceso discreto que cambia en el tiempo.

2. DISEÑO DE AUTÓMATA CELULAR PARA ELECCIÓN PORTUARIA

2.1 Estructura y características del Autómata Celular Portuario (ACP)

Considérese un área correspondiente al hinterland de un sistema de m puertos. Se divide el área en un conjunto mutuamente excluyente y colectivamente exhaustivo de n zonas de generación de carga (ZGC). La Figura 1 muestra un esquema de las zonas y puertos que simula el ACP. Las zonas corresponden a las celdas del autómata celular portuario.

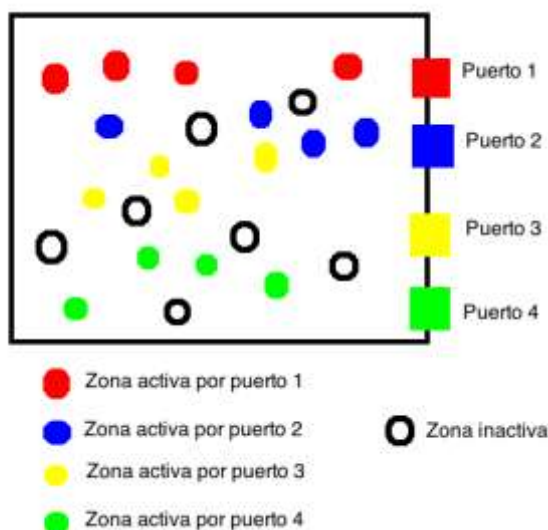


Fig. 1 – Esquema conceptual de las zonas y puertos en el ACP

Cada ZGC actúa como un AC que puede tomar dos estados: activo/inactivo. Una ZGC está activa durante el período de tiempo t si utiliza uno de los puertos disponibles. La condición de actividad/inactividad en un cierto período, dependerá de la distancia de la zona al puerto, características propias del puerto (tarifas portuarias, disponibilidad de infraestructura, tiempos de servicio portuario, capacidad negociadora, especialización portuaria, existencia de contratos de servicio u otros) y el estado de las ZGC vecinas.

Se postula que la función que define la elección portuaria no sólo depende de los atributos antes señalados, sino que además incorpora un elemento estocástico que da cuenta (como en modelos de elección discreta) de al menos dos fenómenos: el modelador desconoce con certeza los atributos de cada puerto y una ZGC puede activarse/desactivarse por elementos exógenos a las variables incluidas en el modelo.

Así, la ZGC i -ésima tiene una función estado $e_i = j$, con j un entero entre 0 y m , donde m representa el número de puertos. Esta función toma el valor 0 cuando la zona está inactiva y j cuando la zona está activa por el puerto j .

2.1.1 Vecindad en el ACP

La vecindad de una ZGC está dada por aquellas ZGC que estén dentro de un círculo de radio R (parámetro a calibrar) como se muestra en la Figura 2.

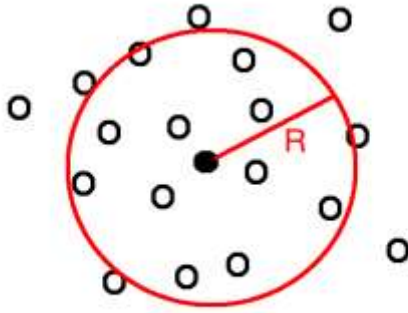


Fig. 2 – Esquema de la vecindad definida en el ACP

2.1.2 Función de Actualización

Una ZGC modifica su estado de acuerdo al siguiente criterio de actualización de un AC probabilístico:

1. Cada zona tiene asociado un número real z comprendido entre $0 \leq z \leq 1$, calibrado por el algoritmo genético. En cada paso del algoritmo, este número z , se reemplaza por otro número elegido de una distribución normal con media igual a z y desviación estándar σ .
2. Se evalúa la función $\Phi_1^i \in \mathbb{R}^m$ sobre la región de vecindad de la ZGC i que se está actualizando. Esta función entrega un vector de m componentes. Cada uno de sus m componentes está formado por la proporción de ZGC dentro de su círculo de vecindad que escogen el puerto j (con $j=1,2,\dots,m$), respecto del total de ZGC activas de la vecindad, esto es:

$$\Phi_1^i = \left(\frac{n_1}{N_i}, \frac{n_2}{N_i}, \dots, \frac{n_m}{N_i} \right) \quad \forall i = 1, 2, \dots, n \quad (1)$$

en que $n_1, n_2, \dots, n_m$ es el número de ZGC activas vecinas que escogieron el puerto 1 a m respectivamente en el período t y N_i es el total de ZGC activas vecinas dentro del círculo de vecindad de la ZGC i .

3. Se evalúa la función $\Phi_2^i \in \mathbb{R}^m$. Esta función tiene como dominio las distancias a los m puertos, y las características propias de cada puerto (como tiempos de espera y tarifas). La expresión final de esta función, se presenta en la Ecuación (2).

$$\Phi_2^i = (\theta_d d_{i1} + \theta_c c_1, \theta_d d_{i2} + \theta_c c_2, \dots, \theta_d d_{ij} + \theta_c c_j) \quad \forall i = 1, 2, \dots, n \quad (2)$$

donde θ_d y θ_c son coeficientes a calibrar (en este ejercicio todos toman el valor 1). d_{ij} representa la distancia desde ZGC $_i$ al puerto j (con $j=1,2,\dots,m$), y c_j representa

los atributos del puerto j desde la perspectiva de los actores en la ZGC $_i$. En el caso de la variable c_j , no se contaba con información histórica para todos los puertos, por lo que fue calibrada como un coeficiente por el algoritmo genético.

4. A los componentes de la función Φ_2^i se les aplica un mapa que convierte su mayor componente en una probabilidad p , al segundo mayor en una probabilidad $p(1-p)$ y así sucesivamente decayendo como ley de potencia hasta $p(1-p)^{m-1}$. A esta nueva función le llamaremos Φ_2^i .
5. Finalmente se obtiene la función presentada en la Ecuación (3).

$$\Phi^i = \alpha \Phi_1^i + (1 - \alpha) \Phi_2^i \quad \forall i = 1, 2, \dots, n \quad (3)$$

con $\Phi^i \in \mathbb{R}^m$. De acuerdo a esta especificación, α representa el peso que tiene en la elección portuaria la componente espacial asociada al estado de las ZGC vecinas y $(1-\alpha)$, el peso de la distancia y las características propias del puerto.

6. Si el número z de la zona es mayor que cierto umbral (T), el estado e_i de cada zona se actualiza y se le asigna el valor del puerto (un entero entre 0 y m) correspondiente a la componente más grande de la función Φ^i .

2.1.3 Parámetros del modelo

De acuerdo a la especificación anterior del modelo, los parámetros a calibrar son:

- El radio que define la vecindad (R).
- El número (α) que representa el peso a los vecinos.
- El umbral (T) para el número z de cada zona.

Adicionalmente el modelo requiere calibrar los siguientes parámetros auxiliares:

- El valor asociado a las características de cada puerto ($c_1, c_2, \dots, c_m$).
- La desviación estándar (σ), que define el ruido en el número z .
- La probabilidad (p) que se utiliza en el método de rango para elegir el puerto.
- El número (z_i) para cada ZGC $_i$ en estudio.

2.2 Calibración del modelo

Los parámetros del modelo fueron calibrados utilizando un algoritmo genético, cuyo procedimiento es el siguiente:

- Se define una población inicial de 100 individuos. Cada uno de estos individuos representa una simulación completa para todo el horizonte de tiempo considerado. El genotipo de cada individuo son los valores de los parámetros del modelo (elegidos al azar en la primera generación). El fenotipo asociado a cada individuo, corresponde a los datos observables involucrados en la dinámica de la actividad portuaria (cantidad de zonas activas en el tiempo por un puerto determinado).
- Se define la siguiente función objetivo:

$$F_0 = \sum_{i,t} f_0(i, t) \quad (4)$$

$$f_0(i, t) = \begin{cases} \gamma \text{ si } (e_i = er_i \neq 0) \\ 1 \text{ si } (e_i = er_i = 0) \\ 0 \text{ en otro caso} \end{cases} \quad \gamma \in \mathbb{R}^+ \quad \forall t \quad (5)$$

donde i representa una ZGC y t el período observado. En esta expresión e_i y er_i representan el estado de una ZGC en la simulación y en los datos reales respectivamente, en cada período observado. El valor de γ permite tener un grado de libertad para derivar a generaciones más adaptadas, privilegiando los aciertos activos (es decir, se privilegia un acierto en una zona activa que el acierto de una zona inactiva). Los 100 individuos se ordenan de acuerdo a la función objetivo. El primero es el individuo con el valor más grande de F_0 .

- Se definen operadores de cruce y mutación para obtener la nueva generación de individuos a partir de la generación pasada, manteniendo el mejor individuo.
- La calidad de cada individuo se evalúa junto con un descriptor de diversidad para mantener una población heterogénea en cada generación.
- Se define un criterio de convergencia que se relaciona con el porcentaje de aciertos en la dinámica del sistema. Este criterio consiste en expresar la función objetivo (Ecuación 4) en porcentaje y detener el algoritmo para un valor determinado (criterio de parada).

3. DISEÑO DEL EXPERIMENTO

3.1 Instancia de aplicación

Son muy pocos los países que disponen de un registro público de los orígenes y destinos de carga de comercio exterior al interior de su país. En el caso de Brasil, existe una base de datos pública que permite obtener la información antes señalada desagregada por tipo de carga, con un nivel de desagregación de hasta 6 dígitos del código arancelario.

Dentro de las cargas principales no graneleras que genera Brasil, se encuentra la exportación de vehículos. Esta carga requiere naves especializadas llamadas RoRo (roll on roll of), las cuales en general tienen una menor prioridad de atención en puertos que no poseen terminales de vehículos especializadas. Esto ha llevado a que marcas como Ford, construyan terminales portuarios privados que eviten los problemas de congestión de los puertos de uso público.

La atención de naves RORO se concentra principalmente en 5 puertos de Brasil, incluyendo en éstos, los terminales privados.

Para probar el modelo propuesto se consideró un conjunto de puertos con *hinterland* común, con el fin de asegurar que las ZGC tuviesen el mismo conjunto de alternativas disponibles.

En este ejemplo de aplicación los puertos escogidos fueron Paranaguá, Santos, Sao Sebastiao y Rio de Janeiro. El puerto de Sao Sebastiao, posee un muy bajo movimiento de vehículos, sin embargo, se decidió mantener en el conjunto de elección, para probar la capacidad del modelo frente a una alternativa que pudiese mejorar sus condiciones de operación.

Las ZGC, en tanto, corresponden a los Municipios que han exportado vehículos por alguno de los puertos seleccionados entre los años 1997 y 2015. Para tener un mayor número de períodos, los datos fueron considerados de manera trimestral, lográndose así 74 observaciones en el tiempo.

En la Figura 3 se presenta la distribución espacial de las 69 ZGC y los 4 puertos considerados.

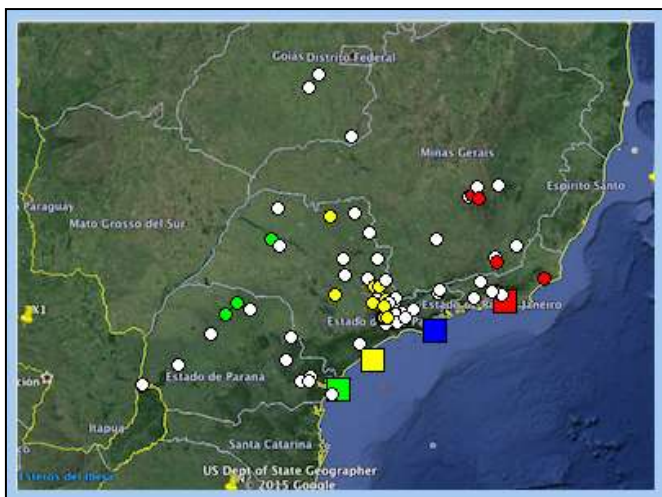


Fig. 3 – Puertos y ZGC consideradas en este trabajo. Los cuadrados representan puertos (un color por puerto) y los círculos representan ZGC. Las zonas rojas están activas por el puerto rojo. Lo mismo para los otros colores. Las zonas blancas están inactivas.

Como se puede observar en la Figura 3, algunas ZGC se encuentran muy concentradas en ciertos sectores, los cuales corresponden a los principales nodos logísticos de la industria automotriz de esta región de Brasil. El porcentaje observado de ZGC activas en cada período, varía entre 12% y 39%, siendo el promedio de los últimos 20 períodos (5años), igual a 29%. La Figura 4, muestra este porcentaje para cada período disponible, así como la curva logarítmica que se ajusta a los datos.

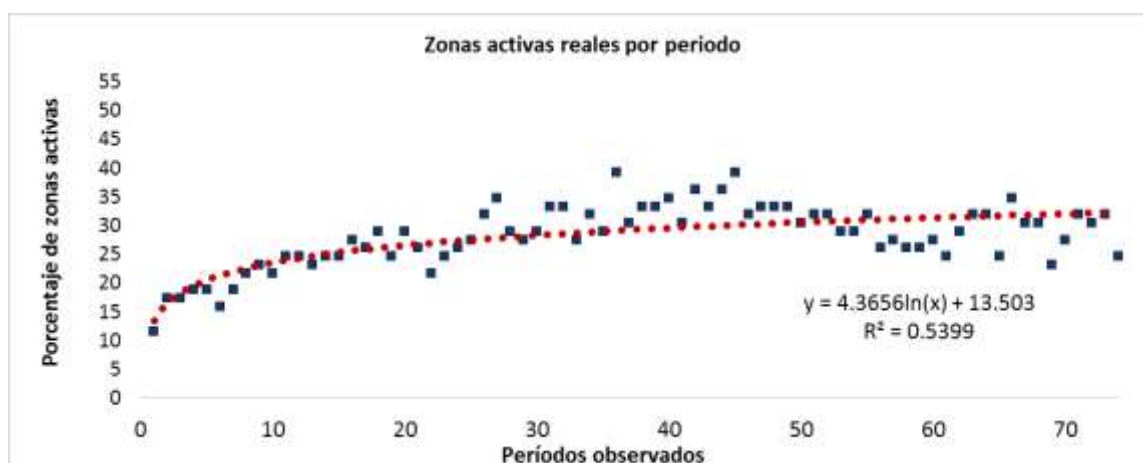


Fig. 4 –Porcentaje de ZGC activas en todos los períodos observados.

3.2 Resultados la calibración

Los parámetros obtenidos del modelo, utilizando el algoritmo antes señalado, se presentan en la Tabla 1.

Parámetro	Valor
R (km)	69
α	0.27
T	0.92

Tabla 1 – Parámetros obtenidos para el Autómata Celular Portuario

Estos resultados se obtuvieron considerando un porcentaje de aciertos de 70% y un γ igual a 4.

Los valores indican que la elección de puerto de una ZGC, se ve influenciada en un 27% por la elección de las ZGC ubicadas dentro de un radio de 69 km de dicha zona. El resto de la elección estaría influenciada por la distancia y características del puerto.

El alto valor obtenido para el parámetro T , correspondiente al umbral, se explica por la baja proporción de zonas activas que se observa en cada período. Esta condición implica una mayor dificultad de activación de una ZGC en el tiempo, la cual se gatilla sólo si se supera el umbral T . Esta característica es fácilmente descrita por un autómata celular, pero no así por un modelo de elección discreta o de regresión, ya que éstos tienden a presentar mayores errores para elecciones poco presentes en la muestra de calibración.

3.3 Comportamiento del modelo como herramienta predictiva

Para evaluar el comportamiento del modelo se realizaron corridas de prueba sin alterar los parámetros asociados a cada puerto. Así, el cambio en la actividad de las zonas y su asignación a cada puerto, se explica principalmente por la componente $\alpha\phi_1$ de la función ϕ (efecto de la actividad de la vecindad de cada zona), y por la aleatoriedad representada por z respecto al umbral calibrado.

El modelo fue ejecutado 20 veces durante 80 períodos (cada período representa un trimestre, por lo que el periodo de predicción equivale a 20 años), con el fin de observar su comportamiento en el largo plazo.

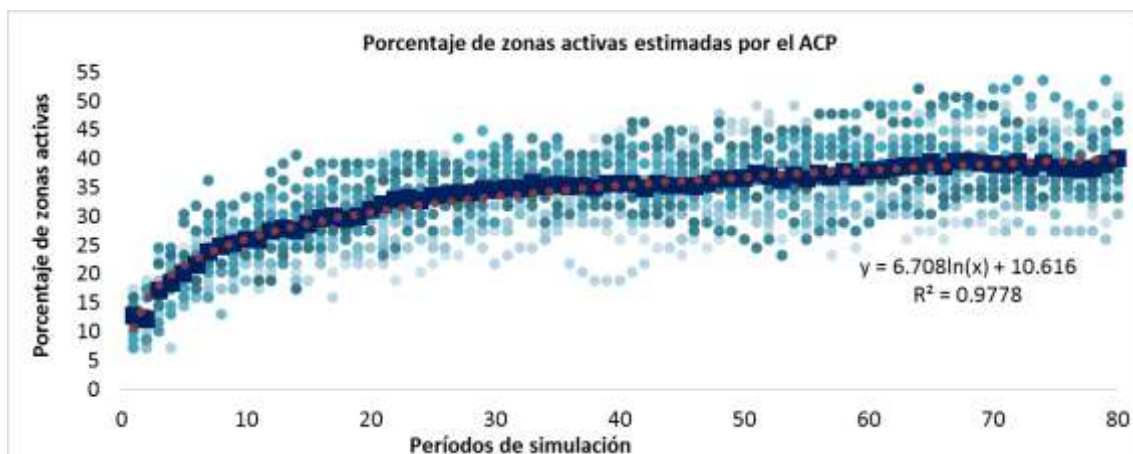


Fig. 5 –Predicción a 20 años de porcentaje de zonas activas

Al comparar el porcentaje promedio de aciertos de cada período (línea azul) con los datos reales de los 74 períodos disponibles (Figura 4), se observa que existe en ambos casos, una tendencia similar en el tiempo. En cuanto al rango sobre el cual se mueve el total de zonas activas, éste se extiende entre 12% y 40% en el promedio del caso simulado y entre 12% y 39% en los datos reales, siendo por tanto de órdenes de magnitud muy similares. Lo anterior indica que el modelo logra representar adecuadamente la dinámica del proceso estudiado.

Un tipo de análisis que se puede realizar con este modelo, tiene relación con las variaciones en las participaciones de mercado, frente a cambios en las características propias de cada puerto. A modo de ejemplo, se estudió el comportamiento que tendría el mercado, frente a dos puertos cercanos, y en los cuales uno de ellos logra ser un 20% más atractivo que su principal competencia. En el caso del modelo, esto significa una variación en la componente c_j de la Ecuación 2.

En la Figura 6 se muestra la participación de mercado en la situación base simulada. Como se puede observar el Puerto 1, no tiene participación de mercado, a pesar de estar espacialmente muy cercano al Puerto 3. Esto se debe a que el Puerto 1 no cuenta con condiciones atractivas para el mercado (infraestructura, accesibilidad, equipos, tarifas, etc.).

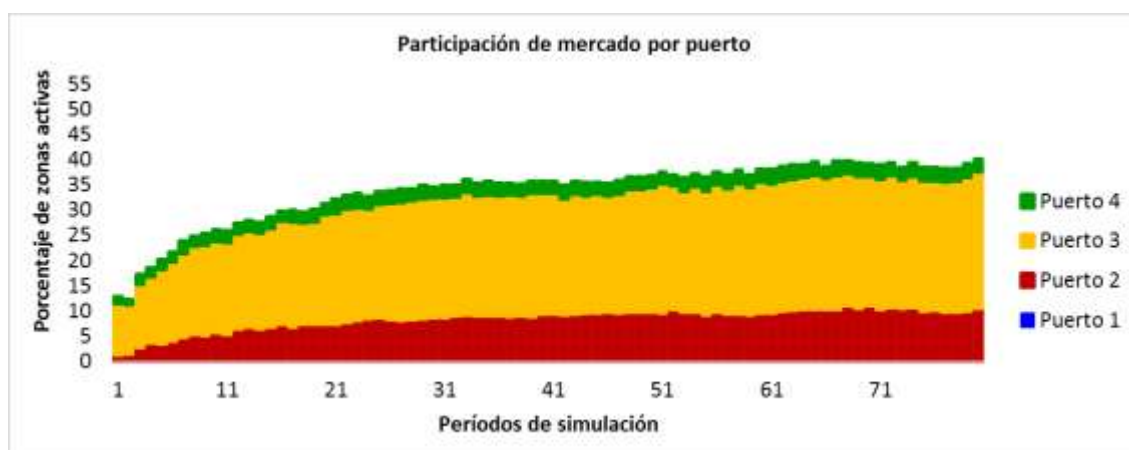


Fig. 6 – Predicción a 20 años de participación de mercado por puerto.

Sin embargo, si asumimos que este puerto lograra ser un 20% más competitivo que su principal competencia (Puerto 3), su participación promedio en el horizonte de tiempo simulado aumentaría a 12%, tal como se muestra en la Figura 7.

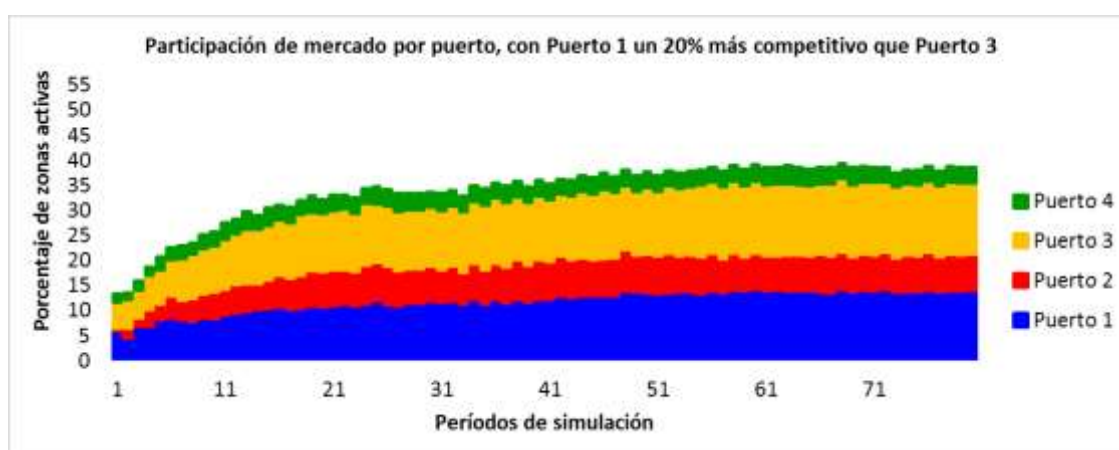


Fig. 7 – Predicción a 20 años de participación de mercado por puerto, considerando al Puerto 1 un 20% más competitivo que el Puerto 3.

Con este mismo ejemplo de aplicación es posible estimar una elasticidad promedio (estable en largo plazo), de la participación de mercado del Puerto 3, respecto a variaciones en la competitividad del Puerto 1. De acuerdo a los resultados obtenidos, la variación de un punto porcentual en la competitividad del Puerto 1, implicaría una variación de la participación de mercado del Puerto 3, en 2,3 puntos. Esta alta elasticidad se debe a dos componentes. Por una parte, la competitividad misma del puerto, y por otra, el efecto espacial de su elección por parte de los vecinos de las zonas ubicadas en el hinterland de ambos puertos.

4. CONCLUSIONES

Se desarrolló un modelo de asignación portuaria basado en un autómata celular probabilístico, que busca imitar la decisión de múltiples actores que interactúan entre sí y que ponderan una multiplicidad de aspectos y comportamientos espaciales que configuran una elección dinámica.

La especificación del autómata celular diseñado fue aplicada a un caso práctico de exportación de automóviles en Brasil y se calibró utilizando un algoritmo genético.

Los resultados indicaron que la elección de puerto de una ZGC, se ve influenciada en un 27% por la elección de las ZGC ubicadas dentro de un radio de 69 km de dicha zona. El resto de la elección estaría influenciada por la distancia y características propias de cada puerto.

Para analizar la capacidad predictiva del modelo, éste fue ejecutado 20 veces durante 80 períodos (equivalente a 20 años). Los resultados mostraron una tendencia equivalente a la presentada por los datos observados, al igual que los rangos de porcentaje de zonas activas, indicando que el modelo logró representar adecuadamente la dinámica del proceso estudiado.

Uno de los usos prácticos de este modelo es la determinación de la elasticidad de la participación de mercado de cada puerto frente a variaciones en las características de éstos. Así por ejemplo, se estimó que para el caso de dos puertos cercanos, como los son Santos y Sao Sebastiao en Brasil, si este último lograra un punto porcentual de competitividad por sobre Santos, lograría captar 2,3 puntos de la participación del mercado de exportación de vehículos.

AGRADECIMIENTOS

Mabel Leva desea agradecer al Grupo de Investigación de Sistemas de Transporte de la Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos de la Universidad de Cantabria, España, por su apoyo y colaboración. Alejandro León y Rodrigo Garrido desean agradecer el apoyo de la Facultad de Ingeniería de la Universidad Diego Portales de Chile.

REFERENCIAS

- AL-DEEK, H.M. (2001). Which method is better for developing freight planning models at seaports: neural networks or multiple regression?. *Journal of the Transportation Research Board* 1763, pp. 90-97.
- CHOPARD, B. y DROZ, M. (2005). *Cellular Automata Modeling of Physical Systems*. Cambridge University Press, Cambridge.
- GARRIDO, R.A. y LEVA, M. (2004). Port of destination and carrier selection for fruit exports: a multi-dimensional space-time multinomial probit model. *Transportation Research Part B* 38 (7), pp. 657-667.
- GOLES, E. y MARTÍNEZ, S. (1998). *Cellular Automata and Complex Systems*. Nonlinear Phenomena and Complex Systems Series, Kluwer Academic.
- GRIFFEATH, D. y MOORE, C. (2003). *New Constructions in Cellular Automata*. Oxford University Press, Oxford.
- LAM, W. H. K, Ng, P. L. P, SEABROOKE, W. y HUI, E. C. M (2004). Forecasts and Reliability Analysis of Port Cargo Throughput in Hong Kong. *Journal of urban Planning and Development* 130 (3), pp. 133–144.
- MAGALA, M. y SAMMONS, A. (2008). A New Approach to Port Choice Modelling. *Maritime Economics & Logistics* 10 (9–34).
- MALCHOW, M. y KANAFANI, A. (2001). A disaggregate analysis of factors influencing port selection. *Maritime Policy & Management: The flagship journal of international shipping and port research* 28 (3).
- NIR, A. S., LIN, K. y LIANG, G. S. (2003). Port choice behaviour--from the perspective of the shipper. *Maritime Policy & Management: The flagship journal of international shipping and port research* 30 (2).
- PETRI, M., FUSCO, G. y PRATELLI, A. (2014). A New Data-Driven Approach to Forecast Freight Transport Demand. *Lecture Notes in Computer Science* 8582, pp. 401-416.
- RODRIGUEZ, T., GONZÁLEZ, N. y SOLER-FLORES, F. (2013). Forecasting models in ports transport systems: Are ANNs Applications the solution?. *Proceedings in EIIC - The 2nd Electronic International Interdisciplinary Conference*, EDIS - Publishing Institution of the University of Zilina, Eslovaquia.

TONGZON, J. (2009). Port choice and freight forwarders. *Transportation Research Part E: Logistics and Transportation Review* 45 (1), pp. 186–195.

VELDMAN, S., GARCIA-ALONSO, L. y VALLEJO, J. (2013). A port choice model with logit models: A case study for the Spanish container trade. *International Journal of Shipping and Transport Logistics* 5(41398), pp. 373-389.

WEIQUN, C. y NUO, W. (2003). Container throughput forecast model for a district port based on neural network. *Proceedings of the Eastern Asia Society for Transportation Studies* 4, pp.254- 258.

YEO, G-T, NG, A.K.Y., LEE, P. T-W y YANG, Z. (2014). Modelling port choice in an uncertain environment. *Maritime Policy & Management* 41(3), pp. 251-267.

Anexo B: A probabilistic cellular automaton to forecast port choice decisions

ABSTRACT

The port choice problem consists in predicting the selection of a port, made by an agent who has alternatives to choose from. Most of the literature has tackled this problem assuming a discrete choice model dependent on the ports' characteristics and agents' attributes. However, in practice the port choice decision depends also on the choices made by other agents as well as decisions made by these agents in the past. There are only a few examples that incorporate the complexity generated by spatio-temporal interactions between agents. However, those modelling structures are rather cumbersome, precluding their use in practical cases. This article presents a new modelling framework to predict port choice decisions, based on the theory of cellular automaton (CA), which is simple in structure and can be quickly calibrated and applied. This framework is a probabilistic CA intended to imitate the decision processes made from multiple shippers that interact with each other. These shippers face similar alternatives of seaports for exporting their products within a certain time span. The port choice here is a dynamic decision that depends on the ports' characteristics and attributes of each shipper at a given time, as well as the decisions made by their neighbours. The outcome of the interaction is a discrete decision that evolves in time according to the dynamics of the system as a whole. The specified CA was applied to the case of vehicle exports from Brazil and the calibration was performed through a genetic algorithm. The results show that the probabilistic CA is able to replicate the historic behaviour of the port choice decisions in the Brazilian vehicle industry, with a high degree of success. The spatial component of the CA turned out to be of major relevance in the dynamic decision process along with the attributes and geographical location of ports.

Keywords: Port choice, cellular automaton, probabilistic cellular automaton, port competition, hinterland formation, discrete choice.

1. INTRODUCTION

More than 85% of the cargo transported globally is moved through the maritime transportation system. This characteristic seems to be a long-run trend, especially when the size of the vessels and the ports infrastructure has risen considerable in the last decade and so has the transference capacity and specialized labour in this industry.

Given this scenario, the port planning processes will need to depend on studies and models to predict (to a certain extent) the load of the different port nodes in a transportation network, in order to balance the transference capacity, cost and time consumption and keeping the negative externalities under control.

The distribution of cargo between the different available competing ports (also called demand capture) is usually analysed from two perspectives: the shippers' viewpoint and the shipping company's viewpoint. The shipper's perspective is that of choosing the port that maximises some utility criterion. For example, the shippers may rely on models whose main explanatory variable is the generalized cost of transporting the cargo between ports and (from) destinations (origins), within their hinterlands. From the perspective of the shipping companies, the choice between alternative ports rely on charges, rates and ports' level of service, as explanatory variables that define the port's efficiency (Tongzon and Sawant, 2007). There are though, additional aspects that condition the port choice. In fact, according to Chou et al. (2008) there are at least six relevant aspects, which are: local cargo volume; terminal handling charge; berth availability; port location; transshipment volume and feeder network.

In both cases, only a subset of the explanatory variables is included into the models, mainly because of difficulties to measure them. In fact, in some cases these variables are economically sensitive for the stakeholders and, most importantly, it is difficult to predict them for the entire planning horizon. For example, key aspects such as bargaining power, specialization in a niche (e.g. refrigerated cargo), long-term contracts, commercial integration, among others, are left out of the models for the difficulties explained above.

The problem to be tackled in this article is that of the prediction of port choice from the shippers' perspective, considering variables related to port performance as well as spatial and temporal interaction, i.e. the influence of one shipper's decision over another that is geographically close and the effect of decisions along the time line.

The hypothesis about the influence of the shipper's decisions over one another is backed up by empirical evidence. In fact, large shippers affect the port choice decision of other shipping companies (Tiwari et al, 2004), due to the higher volumes (Chou et al., 2008). These large volumes generate the proper conditions for new services to flourish and also improve the shippers bargaining position vis-à-vis the shipping companies. With the presence of new shipping services, smaller shippers also improve their bargaining position and consequently smaller shippers will tend to follow the large shippers, most likely their neighbours, to take advantage of the long-term relationship between the large shippers and shipping companies. At the same time, the shippers' location is linked to similar economic activities and consequently there is also an indirect relationship between spatial location and port decisions for both actors.

This article attempts to develop a new modelling framework to describe and predict the behaviour of decision makers in the maritime transport chain, in which various shippers choose a port of origin to export their shipments during a period of time. To attain this main objective, we adapt the structure of a CA model, in order to capture the spatio-temporal interactions that in other modelling contexts are rather difficult to calibrate and operate (Garrido and Leva, 2004).

2. LITERATURE REVIEW

2.1 The Port Choice in Maritime Transportation

The usual explanation for the port choice behavior in the specialized literature is the generalized transportation cost from/to the hinterland to/from the ports and the reliability of the total route within the context of a supply chain instead of the sole port choice. Other studies point out significant differences in the triggers of port choice between different commodities (Malchow and Kanafani, 2010). According to Moya and Feo-Valero (2016), the identification of the decision-maker is one of the main critical issues in port choice modelling. Furthermore, they point out that the decisions of different agents are highly interrelated and inherently iterative. Due to the latter, the determinants of the decision process vary significantly with the characteristics of the decision-maker, its logistics requirements and objectives.

Tongzon (2009) studies the port choice from the freight forwarders' viewpoint; finding out that the main factors in this decision are port efficiency, shipping frequency, adequate infrastructure and location. Magala and Sammons (2008), studied the port choice problem incorporating the concept of supply chain, pointing out that shippers select a set of logistics services, in which the port itself is only part of the whole chain.

There are authors who studied the port choice problem from the shipping lines perspective, where the port charges and level of service at the port become the fundamental decision variables (see Tongzon y Sawant, 2007). However, according to Chou et al. (2008), there are at least six relevant decision variables: local cargo volume, port's charges, berth availability, port's location, transferred volume and feeder connection.

Other authors have studied the problem of port choice considering various types of actors. For example, Song and Yeo (2004) identified 73 elements that contribute to the global competitiveness of the main ports in China, considering the opinions of ship-owners, shippers, shipping companies, terminal operators, researchers and scholars. After this broad analysis Song and Yeo (2004) concluded that location was the most relevant variable as far as competitiveness is concerned.

Lee et al (2010), identify and evaluate the main determinants of the port choice from the carriers, shippers and terminal operator's perspectives. They identify 38 attributes for each of the seven analyzed categories (port location, port characteristics, port's operation conditions, hinterland characteristics, ships characteristics, among others). The above-mentioned stakeholders evaluated these attributes, finding considerable variability in the assessment of each attribute by each stakeholder, and only one attribute that was commonly evaluated as relevant to all the stakeholders: port accessibility. However, further analysis of the shippers and containers' terminal operators, revealed that they had a larger number of common attributes and also giving a larger weight to "port's operation conditions" than that given by the shippers.

2.2 Port Choice Modeling

The common family of port choice studies is discrete choice models (Manski and McFadden, 1981). These models assume the following: there are rational agents that face the selection of a single choice among a choice set formed by all the alternatives available to her. Each alternative can be conveniently described by its own attributes. The selecting agent has a utility function (a score) that depends on the attributes of each alternative, some individual characteristics plus a random component that explains deviations from the rational behavior (given the available explanatory variables). Therefore, the agent chooses a certain alternative according to the probability that the random utility function is maximized when selecting that particular alternative.

One of the most popular discrete choice models based on random utility theory is the Multinomial Logit (MNL, see Hausman and McFadden, 1984), which requires two conditions: all the available alternatives have to be uncorrelated between them, and the random component of the utility function has to be homoscedastic (i.e. the variance of all the alternatives' utility functions have to be equal). Examples of application of the MNL to the port choice problem are found in Veldman et al., 2015, Tiwari et al, 2004 and Malchow and Kanafani, 2004.

These MNL models show a good performance in terms of nominal prediction power. However, it is likely that the reported performance was highly biased due to the violation of MNL requirements (homoscedasticity and independence between alternatives) to produce unbiased and consistent estimators. Indeed, the choice set (i.e. their available alternatives) had significant correlation and heteroscedasticity. Those conditions are commonly found in the port choice problem. In fact, the spatial proximity of different port terminals and the type of long-term contracts established between shippers and carriers that define ex-ante the port terminals to be used within the duration of the contract.

These effects create both spatial and temporal correlation that is not easy to deal with in the context of discrete choice models.

In addition, the maritime conditions, the ports infrastructure and the interface with the land transportation system (to name a few) impose large variability in the port services, which creates the ideal conditions for a heteroscedastic behavior when considered as alternatives within a choice set. As a result, there are only a few examples of published studies in specialized journals that deal with more robust models that allow the violation of the MNL restrictions. One of these attempts is the modelling framework proposed by Veldman et al (2013). This approach partially solves the problem of heteroscedasticity through the use of the Hierarchical Logit model (HL, see Papola, 2004). However, the use of HL models avoids only a small fraction of the possible correlation/heteroscedasticity problem. The latter is due to the rigid structure of the variance-covariance matrix that the HL imposes.

To the best of the authors' knowledge, the only modeling approach that deals simultaneously with a general structure that allows alternatives with a general random component's specification is Garrido and Leva (2004). In that study the authors present a general Multinomial Probit model (MNP, see Garrido and Mahmassani, 2000). The MNP model put forward by those authors incorporates spatial and temporal correlations between alternatives and a general structure for the variance-covariance matrix. The model gives a good tool for forecasting port choice as well as representing the interactions between agents in a coherent way. However, there are complications with the calibration process (numerical stability, multiple maxima for the likelihood function, large computational effort, among others) that makes it difficult to be used in practice or by non-econometric experts.

There are other methodologies developed more recently to address this very same problem. One of them is the approach proposed by Yeo, et al (2014). They propose a port choice model based on fuzzy logic. In this approach the authors combine various data sources and formats, both objective and subjective, to assess the performance of a given port, determined through a "fuzzy score" using degrees of belief to come up with a criterion to establish the port choice. This modelling approach was applied to the choice of container ports in northeast Asia. The outcome of the model is a preference score that allows the comparison of different ports either vis-à-vis other ports or the system as a whole. However, this methodology is strongly dependent on the opinion of the stakeholders considered in the sample, which can easily change in line with their preferences at that particular time.

2.3 Cellular Automaton Modelling

The CA seems to be a direct tool to deal with a phenomenon that is dynamic, non-linear and complex (Goles and Martínez, 1998) with significant computational efficiency and with reasonable data needs. The CA can be defined, for our purposes, as a mathematical model belonging to the field of computational physics. Proposed originally by John von Neumann during the 1950's to simulate and describe complex systems as a massive collection of simple objects that interact with each other locally. These simulation models proved to be especially adequate to study dynamic systems that evolve in discrete steps (see for example Griffeth and Moore, 2003; Chopard and Droz, 2005).

The simulation structure is based on rather simple interactions among shippers and between shippers and ports to predict the port selection. The port choice emerges as a result of the dynamic interactions with feedback between these decision-makers. In spite of the simplicity of these interactions, the proposed methodology captures the complex behavior of the system and predicts the port choice evolution with high accuracy. In addition, the modelling framework is simple enough to be quickly calibrated and applied to real cases.

This tool would be of great value for methodological and practical reasons alike. From the theoretical perspective, this approach could be a tool for searching complex patterns (chaotic behavior, sensitivity to initial conditions, etc.) in the dynamics of market share formation, as well as identifying attributes that influence a port's market share. On the practical side, this approach is useful to establish not only the current market share of a port but also its hinterland and the variables that affect its boundaries within a competitive ports' system.

3. DESIGN OF A CELLULAR AUTOMATON FOR PORT CHOICE

3.1 A Simulation Structure for the Port Choice Problem with a Cellular Automaton Model

Consider a given spatial region that serves as a shared hinterland for a network formed by m ports. We divide this region into n mutually exclusive and collectively exhaustive freight generation zones (FGZ). Each FGZ can have $m+1$ states of activation: one for each available port and one state of inactivity, meaning that if a FGZ "has chosen" port i as a port of call, then its current state is i , and if the FGZ has not chosen any port then it is not active, hence its state is Inactive. Thus, the i -th FGZ has a state function as follows: $e_i = j$, with $j=0,1,2,...,m$ where 0 means inactivity and j means that i -th FGZ has chosen port j .

Figure 1 shows a scheme of the FGZ and ports on which a port choice cellular automaton (CA) model will operate. Each FGZ, represented as a circle, corresponds to a single cell of the CA. Thus, each FGZ acts like a CA that is either active or inactive. A given FGZ is active during

discrete time t if it sends a shipment to an available port. The port choice decision of that FGZ at time t depends on the distance to the port, the operational characteristics of the port (e.g. tariffs, availability of dedicated space for incoming vehicles, additional services, expected waiting times, transfer capacity, specialization, long term contracts, bargaining position, among many others). Our novel approach assumes that the condition of activity at time t also depends on the state of its neighbouring FGZ.

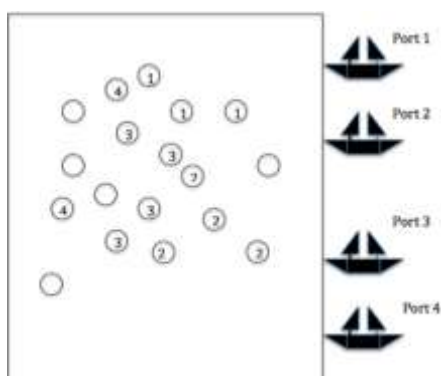


Fig. 1 – Conceptual scheme of Freight Generation Zones and Ports within a shared hinterland.

In addition, this modelling framework assumes that the port choice not only depends on the attributes mentioned above, but also incorporates a stochastic element that would explain (similar to the assumption of random utility theory) at least two phenomena. First, the modeller ignores the actual attributes of each port at any given time. Secondly, a FGZ could become active/inactive triggered by exogenous variables not included in the model specification. Thus, our modelling approach will be called Probabilistic Port Choice Cellular Automaton (PPCCA).

3.1.1 The Neighborhood Concept in the PPCCA Model

There are several ways to interpret the concept of neighbourhood in CA. For example, Von Neumann introduced the concept of neighbourhood in a rectangular grid as the cells sharing a common edge (four neighbours in a squared grid). Other classical concept is that of Moore neighbourhood, in which a neighbour is each one of the eight cells surrounding a given cell in a squared grid. In the port choice case, the FGZ not necessarily are ordered in a Manhattan metric array.

Consequently, the concept of regular grids does not fit the actual spatial pattern. Instead, we incorporate the concept of radial neighbourhood, in which two cells are R -neighbours if they both are circumscribed in a circle of radius R , as shown in Figure 2.

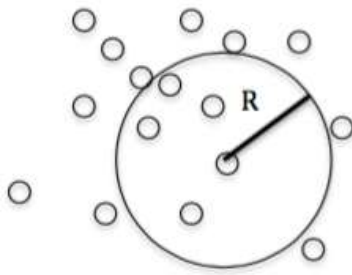


Fig. 2 – R -neighbours in the PPCCA are those FGZ within the circle.

3.1.2 Transition Function

In the PPCCA model each cell corresponds to a FGZ that transits from its current state to the next one (in discrete time steps) according to a stochastic transition function. Probabilistic CAs update their state according to a rule that follows a certain probability distribution. Accordingly, they belong to the category of discrete-time random dynamical systems. Incorporating stochasticity in the spatial interaction between cells allows these models to identify complex behaviours that may emerge from the system.

The PPCCA will update the state of each FGZ, as a probabilistic CA, according to the following rules:

- Each zone has a real number associated with it: $z \in \mathbb{R} \wedge 0 \leq z \leq 1$. This number is calibrated by a genetic algorithm (that will search for the optimal combination of parameters that best fit a set of observed data). At each simulation step, z is replaced by another real number, drawn from a normally distributed sample whose mean value is z and its standard deviation is σ (also calibrated by the genetic algorithm).
- Each FGZ i has a function $\Phi_1^i \in \mathbb{R}^m$ whose domain is the R -neighbourhood of the i -th FGZ. This function's outcome gives a vector of m components. Each component is formed by the ratio between the number of FGZ within the R -neighbourhood that choose the port j , relative to the total number of active FGZ within the same R -neighbourhood. i.e.:

$$\Phi_1^i = \left(\frac{n_1}{N_i}, \frac{n_2}{N_i}, \dots, \frac{n_m}{N_i} \right) \quad \forall i = 1, 2, \dots, n \quad (1)$$

where $n_1, n_2, \dots, n_m$ is the number of FGZs that chose the port 1 to m respectively in a given time step, and N_i is the total number of active zones within the R -neighbourhood of i -th FGZ.

- Each FGZ i has a function $\Phi_2^i \in \mathbb{R}^m$ whose domain is the set of distances to the m ports.

$$\Phi_2^i = (\theta_d d_{i1} + \theta_c c_1, \theta_d d_{i2} + \theta_c c_2, \dots, \theta_d d_{ij} + \theta_c c_j) \quad \forall i = 1, 2, \dots, n \quad (2)$$

where θ_d and θ_c are coefficients to be calibrated by the genetic algorithm. d_{ij} represents the distance from the i -th FGZ to the port j (with $j=1,2,\dots,m$), and c_j represents an aggregated *proxy* of the various costs involved in the choice of port j (note that these coefficients may implicitly include the costs associated to the carrier selection). As the c_j coefficients represent some generalised cost function, hence smaller values are preferred to larger ones (they can also be interpreted as a proxy for the *disutility* of choosing port j).

- The function Φ_2^i is mapped as follows: the smallest component of Φ_2^i is assigned a probability p , the second smallest component is assigned a probability $p(1-p)$ and so forth decreasing as a geometric law until reaching $p(1-p)^{m-1}$. The new mapped function, that contains the probability values (instead of distances) for each component, is renamed $\Phi_2'^i$.
- Finally, a compound function is defined as follows:

$$\Phi^i = \alpha \Phi_1^i + (1 - \alpha) \Phi_2'^i \quad \forall i = 1, 2, \dots, n \quad (3)$$

with $\Phi^i \in \mathbb{R}^m$. In this specification, the parameter α represents the influence of the R -neighbours' behaviour on the port choice decision, i.e., the spatial component of the decision, associated to the FGZ R -neighbours' interaction, and $(1-\alpha)$, the influence of the distance and other systematic attributes related to the "attractiveness" of a port, on the port choice decision. The attractiveness here has opposite sign to the coefficients c_j

- If $T \leq z$, (where T is an activation threshold calibrated by the genetic algorithm) then the state e_i of each FGZ is updated with a value between 0 and m , representing the choice (0 corresponds to inactivity) corresponding to the highest component of the function in expression (3).

3.1.3 Model Parameters

According to the specification described above, the following parameters need to be calibrated before the model can simulate the dynamics of the port choice decisions:

- R : The radius of neighborhood for all the FGZs.
- α : The weight of the neighbours' port choice decision
- T : Activation threshold

Additionally, the model requires the calibration of the following parameters:

- The value associated to the overall costs of each port ($c_1, c_2, \dots, c_m$).
- The standard deviation (σ) for the whole area and (z_i) for each FGZ i .
- The probability (p) is used in a range manner to add variability to the system's dynamics. Thus, the "best candidate" will not always be chosen but instead it will have the highest probability to be chosen.

3.2 Model Calibration

The parameters described above were calibrated with the aid of a genetic algorithm (Goldberg, 1989), implemented in Lazarus programming language (Van Canneyt et al, 2011). The genetic algorithm runs as follows:

- An initial population of 100 individuals is generated. Each one of these individuals represent a set of values for the PPCCA's parameters. The genotype of each individual is formed by the values of the unknown parameters (randomly selected at the beginning of the optimization process).
- The phenotype associated to each individual corresponds to the observed data reflecting the port choice dynamics, i.e., active FGZ for each port at every time step during the whole sample. For each phenotype, various indicators can be computed to measure the goodness of fit between modelled choice and actual data.
- The selection of the optimal parameters' values must be done in terms of the best fit to the observed data. Therefore, we defined a score to be optimized as follows:

$$F_0 = \sum_{i,t} f_0(i, t) \quad (4)$$

$$f_0(i, t) = \begin{cases} \gamma & \text{if } (e_i = er_i \neq 0) \\ 1 & \text{if } (e_i = er_i = 0) \\ 0 & \text{otherwise} \end{cases} \quad \gamma \in \mathbb{R}^+ \quad \forall t \quad (5)$$

where i represents a FGZ and t is the observed time period. The best set of values is such that F_0 is maximum. In fact, each time the set of values yields a successful assignment (i.e. the simulated port choice coincides with the observed data at a given time period), the value of $f_0(i, t)$ increases. In this expression, e_i and er_t represent the state of a FGZ in the simulation and the observed dataset respectively at each time step. γ is an auxiliary parameter that allows a degree of freedom to induce the search towards more adapted generations. The greater the value of γ the greater the relevance given to an active success, i.e., a realization in which the simulation predicts an active FGZ, which is also active in the observed sample.

- Once the algorithm has evaluated the objective function F_0 , the 100 individuals are arranged in a descending order. Then, crossover operators are defined between the top 30% of the list and the rest of the population.
- To incorporate diversity in the next generations we defined mutation operators. The gene to be mutated is chosen randomly.
- The quality of each individual is evaluated along a diversity descriptor, to guarantee a heterogeneous population at each generation. Thus, at each future generation the “best” individual of the previous generation is preserved.
- A convergence criterion is defined, in accordance with the percentage of success between simulated and observed individuals. This percentage is an input from the modeller. The algorithm stops when a pre-specified threshold is reached.

4. EXPERIMENTAL DESIGN

4.1 Application Instance

Only a few countries have a public archive with detailed origin-destination data for international commerce. Brazil is one of those countries. In fact, they have publicly available datasets with freight transportation data disaggregated by type of commodity up to a 6-digit customs code (<http://www.aliceweb2.mdic.gov.br/>).

For this reason, we have decided to apply the modeling framework to actual data obtained from this source. Among the non-bulk commodities available in the Brazilian database we have chosen the Roll-on Roll-off (RORO) market, which is very active in Brazil, especially for vehicles exports. There are more than 30 brands with over 65 plants located in the South,

Southeast, Center-West, Northeast and North of Brazil²⁵. The RORO vessels operate mainly in five Brazilian seaports (both private and public terminals).

To apply and test the developed methodology, we selected a set of ports with public terminals and common hinterland, in order to control that all the FGZ had the same available choice set. The selected ports were Paranaguá, Santos, Sao Sebastiao and Rio de Janeiro. The Port of Sao Sebastiao moves a significantly lower volume of cargo compared to its competitors; however, it was included in the choice set to test the ability of the modeling framework to respond to changes in the ports' conditions and attributes when an improvement has been made (e.g. investment in capacity).

The FGZ correspond to *Municipios* which are administrative geographic units in which the whole country is divided. The selected sample contains records of vehicle exports through any of the above mentioned ports since 1997 until 2015. The temporal resolution is a trimester, obtaining thus a total of 74 observations in a time series.

The proportion of active FGZ at each time interval ranges between 12% to 39%, and within the last 20 periods (five years of data) the average is 29%. Figure 3 shows the percentage of active FGZ at each period in the sample, as well as a smooth curve fitted to the data.

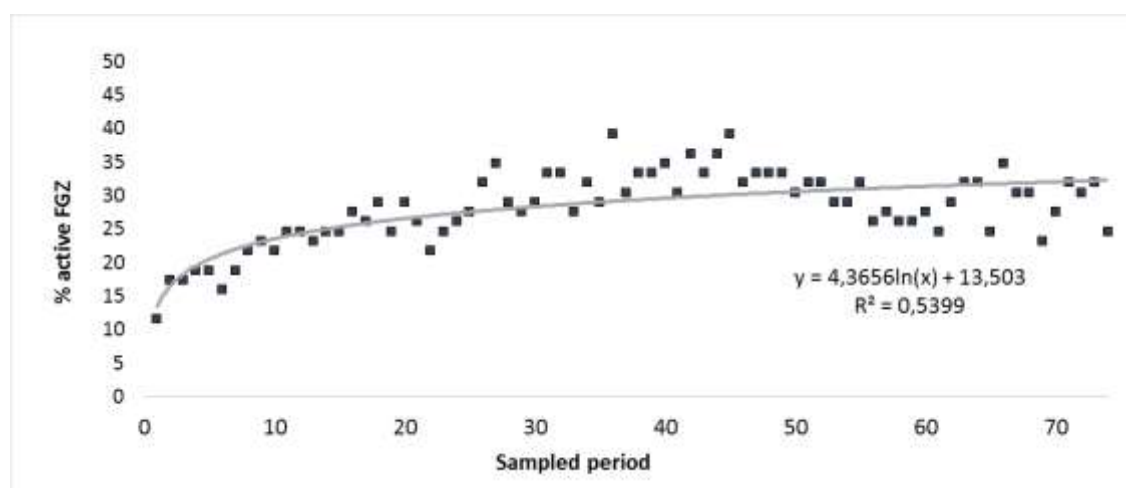


Fig. 3 –Percentage of active FGZ in each sampled period

Regarding the historical market share of each port, Figure 4 shows Port 1 with a very low share, which is explained by its accessibility from the entire area (only 4% of the FGZs are closer to this port than to its competitors) as well as its lack of attractive market conditions (infrastructure, equipment, charges, etc.). On the other hand, Port 3 is the most attractive

²⁵ Source: Brazilian Automotive Industry Yearbook 2016. It includes only companies associated to ANFAVEA (Brazilian Automotive Industry Association)

from the location perspective; 63% of the FGZs are closer to it, followed by Port 2 with a 23% and Port 4 with 12%.

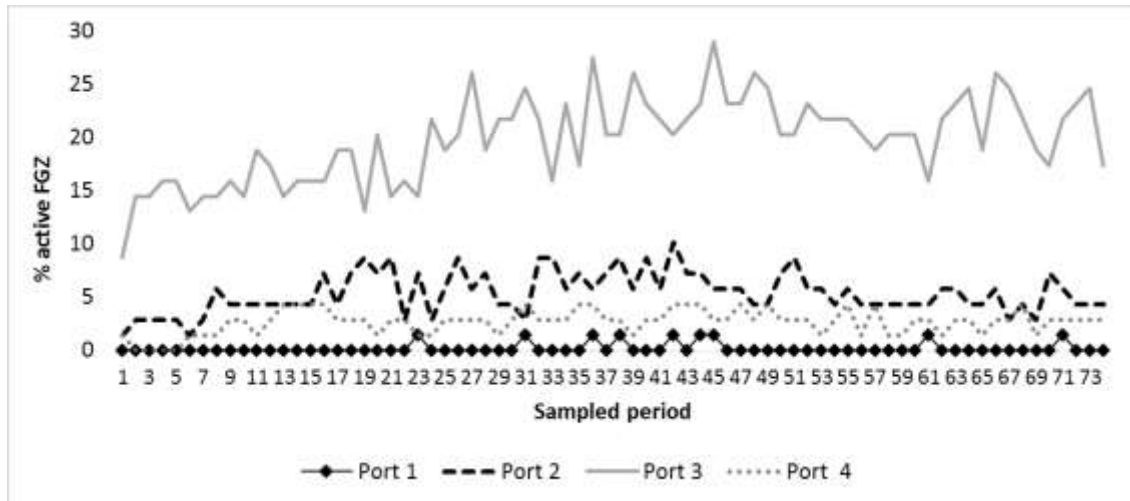


Fig. 4 –Historical market share of each port in the choice set

4-2 Calibration Results

The PPCCA parameters were estimated through the genetic algorithm implementation explained in Section 3.2. The obtained parameters are: $R = 69\text{km}$; $\alpha = 0.27$; $T = 0.92$. As stated in Section 2.2, the pre-specified threshold acts as stopping criterion for the algorithm. In this case that criterion was 70%, with was the highest value found for this sample. This prediction success is considered high when compared to other transportation demand models calibrated with actual data. In the case of discrete choice models, the proper way to establish their goodness of fit is by comparing the ratio of log-likelihood of an only-constants model and the log-likelihood of a given specification. That ratio is named ρ^2 . The best published models reached values of ρ^2 around 0.5 for the port choice problem. Table 1 shows the goodness of fit reached by some discrete choice models applied to the port choice problem.

Table 1: Goodness of Fit for Discrete Choice Models Applied to Port Choice

Author	Model	ρ^2
--------	-------	----------

Tiwari et al. (2004)	MNL	0.57
Malchow and Kanafani, (2004)	MNL	0.54
Veldman et al. (2015)	MNL	0.507

The percentage of success reached by the PPCCA was computed using a calibration parameter that assigns four times more weight to the success of an active FGZ than to an inactive one, i.e. the value of γ in section 3.2 takes the value 4.

The obtained values show that the port choice in a FGZ is influenced by the choice of its R -neighbours with a weight of 27% (neighbors within a circle of 69km of the FGZ). The remaining portion of the selection score is then explained by the distance to the port and the port's characteristics.

Note that the threshold T was calibrated obtaining a high value (0.92 and the maximum possible value was 1.0). This value is explained by the low proportion of active FGZs in the whole sample; about 30% of the FGZ remain active during the sample time span. This characteristic of the dynamic system is easily captured by a CA model vis-à-vis a discrete choice model or a regression model, because those modelling approaches tend to present larger errors for agents with lower activity (or presence) within the calibration sample. The latter is due to the fact that the error term variance in discrete choice models is independent of the value of the choice probability. Thus, the error magnitude is larger when compared to alternatives with small market shares than it is to larger market shares. Consequently, those ports with significantly smaller market shares are subjected to error components that could be much larger than their systematic component

4.3 Model's performance as a forecasting tool

To evaluate the model's performance as a forecasting tool we carried out several runs of the model keeping the ports' attributes fixed (i.e. $(1 - \alpha)\Phi_2^i$ remained constant in expression (3)). Thus, the change in the FGZ's activity is explained mainly by the term $\alpha\Phi_1^i$ in the Φ^i function (the spatial effect of the R -neighborhood) and the randomness added by the threshold condition for activation.

The model was executed 20 times during 80 periods (each period representing one trimester), i.e. an equivalent to 20 years' simulation, in order to observe the long run behavior of the system's evolution.

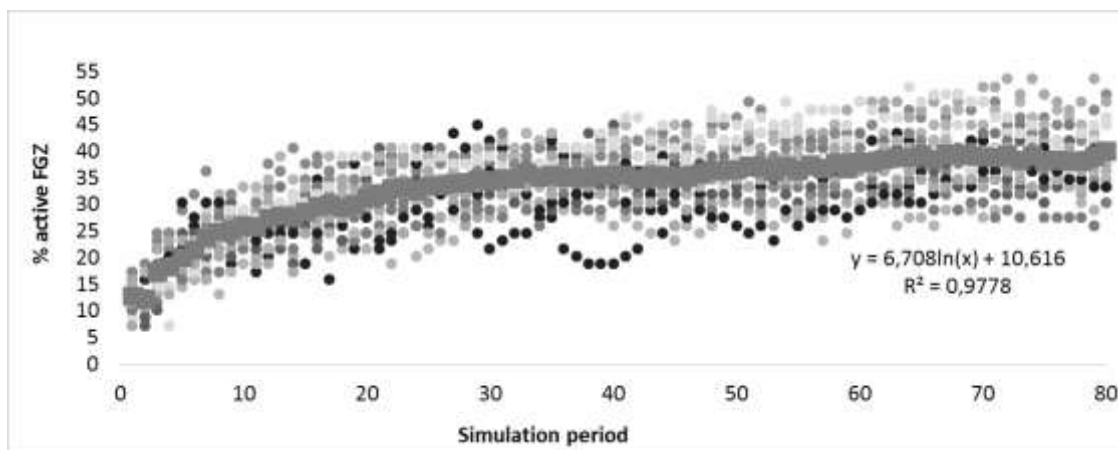


Fig. 5 –Forecasting of active FGZ for a 20 years' simulation horizon

When comparing the average percentage of successful forecasts at each period (solid line in Figure 5) with the actual data for the 74 available periods (Figure 3), it can be seen that the trend in time is similar in both cases. The number of active FGZs varies between 12% and 40% in the simulated dynamics (i.e. the CA evolution) whereas the actual number of active FGZ lies between 12% and 39%, which is a good indicator of the accuracy of the calibrated cellular automaton in terms of capturing the system dynamics.

We studied the case of two ports close to each other (in terms of distance) when one of them increases its competitiveness by a 20% compared to its best competitor (i.e., the disutility of Port j was increased 20%). In the model, the latter would mean to decrease c_j by 20% in equation (2). Figure 6 shows the market share in the baseline simulation. As can be seen, Port 1 does not capture any FGZs even though it is located close to the Port 3. The explanation for this is that Port 1 does not have sufficiently attractive conditions compared to its competitors (infrastructure, accessibility, equipment, charges, etc.).

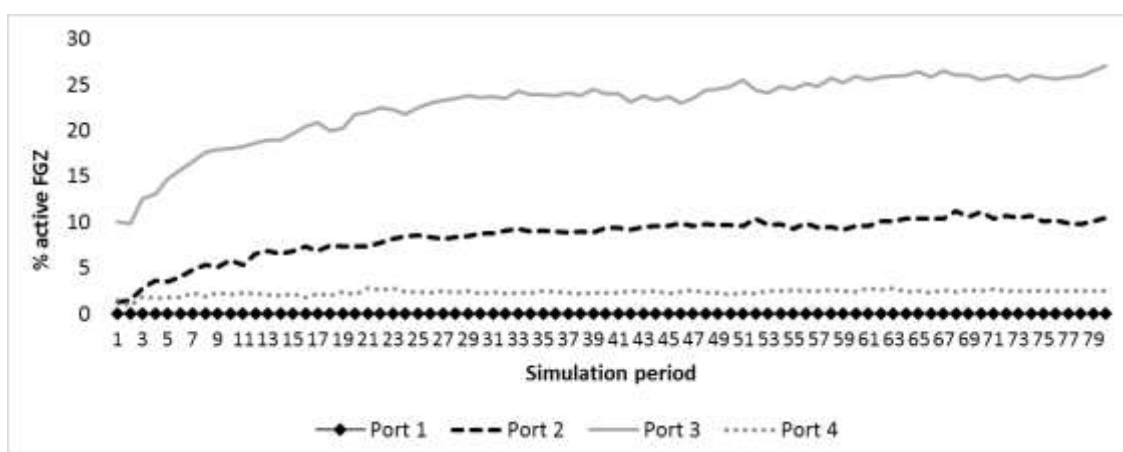


Fig. 7 –Market share forecast per port in a 20 years' span

However, if Port 1 would manage to increase its attractiveness (or decrease its overall costs, represented by c_1) by 20% (compared to its closest competitor Port 3), its market share would increase by 12% in the simulated horizon, as shown in Figure 7.

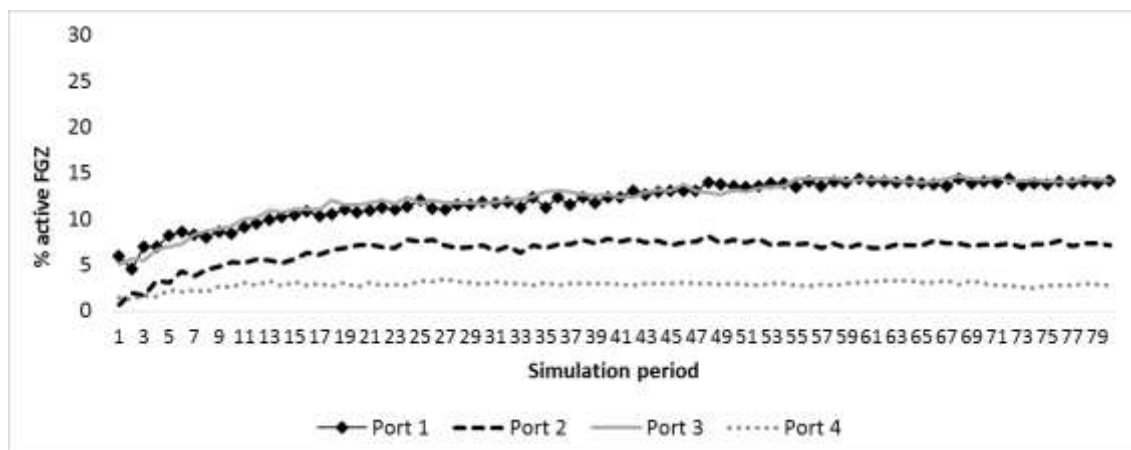


Fig. 7 – Market share forecast per port in a 20 years' span with Port 1's attractiveness increased by 20% respect to Port 3.

With this very same example is possible to estimate the average elasticity for the market share of Port 3 with respect to variation in the attractiveness of Port 1 (represented as lower overall costs). According to the obtained values, a 1 point variation in the attractiveness of Port 1 would trigger a 2.3 points of variation in the market share of Port 3. This high elasticity is due to two components. On one hand the attractiveness of the port itself, and on the other hand the spatial effect produced in the rest of the FGZs within the hinterland of both ports.

5. CONCLUSIONS

In this article, we developed a new modelling structure to solve the port choice problem, based on a probabilistic port choice cellular automaton (PPCCA) that attempts to imitate the decision of multiple spatially located actors that interact with each other. These actors base their decision on various aspects such as their distance to each port, the attractiveness of each port as well as the spatial component of its own neighbourhood, in which the choice made by a FGZ's R -neighbours (those within a circle of radius R) affects the decision of the FGZ itself.

This modelling framework applies only in cases where the decision makers interact between each other spatially through simple rules, but generating a complex system's behavior. Thus, for cases where the decision makers are other than the exporter (e.g. importers), it would be necessary to search for transition rules that capture the interaction between the various agents involved in the choice process.

The specification of the PPCCA was applied to a practical case of vehicle exports in Brazil. The parameters' calibration was done through a specially adapted genetic algorithm.

The results showed that the choice of a port by a given FGZ (i.e. the actors located within it) is strongly influenced (27%) by the choice made by its *R*-neighbours within a circle of 69km from the given FGZ. The rest of the attributes influencing this decision rely on the distance from the FGZ to the ports as well as the particular characteristics of each port.

To test the forecasting ability of the proposed modelling framework, the system's behavior was simulated by executing the PPCCA 20 runs during 80 periods (i.e. a 20 years horizon). The results showed that the simulated behavior of the FGZs capture the dynamics of the actual data with high accuracy, both in the trends and the percentage of active FGZs at each period.

One of the practical uses of this framework is the estimation of elasticity of market shares with respect to any of the attributes involved in the decision process. In particular, the elasticity of market share with respect to the ports characteristics is a value of interest for many stakeholders. As an example, we estimated the value of such elasticity comparing two close ports: Santos and Sao Sebastiao in Brazil. In fact, if Sao Sebastiao would increase its attractiveness by 1 percentage point (over that of Santos), it would increase its market share by 2.3 points in the Brazilian vehicle export industry.

As a future line of research, the modelling framework presented in this article can be modified to represent not only the port choice but also the exported volume and the corresponding flows in the transport network. Another extension could be the inclusion of multiproduct structures. The latter would require the search of an alternative way to represent the concept of neighbourhood, capturing the interaction between exporters of various products, e.g. exporters of the same product could have more influence on their peers than closer neighbours that export a different product. In regard with the case presented in this article, the simulation could be improved with the inclusion of additional explanatory variates that may play a significant role in explaining the system's behavior. For example, costs and times at the ports and for the deep-sea journey, could allow a more accurate description of the overall costs. However, the problem of measuring those values for each alternative and period represents a major challenge.

REFERENCES

- Chou C-C., Chu, C-W. and Liang, G-S (2008). A modified regression model for forecasting the volumes of Taiwan's import containers. *Mathematical and Computer Modelling* 47 (9–10), pp. 797–807.
- Chopard, B. and Droz, M. (2005). *Cellular Automata Modeling of Physical Systems*. Cambridge University Press, Cambridge.
- Garrido, R.A. and Leva, M. (2004). Port of destination and carrier selection for fruit exports: a multi-dimensional space-time multinomial probit model. *Transportation Research Part B* 38 (7), pp. 657-667.
- Garrido, R. and Mahmassani, H. (2000) Forecasting freight transportation demand with the space–time multinomial probit model. *Transportation Research Part B: Methodological* Volume 34, Issue 5, June 2000. Pages 403–418.
- Goldberg, D. (1989) *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA...
- Goles, E. and Martínez, S. (1998). *Cellular Automata and Complex Systems*. Nonlinear Phenomena and Complex Systems Series, Kluwer Academic.
- Griffeath, D. and Moore C. (2003). *New Constructions in Cellular Automata*. Oxford University Press, Oxford.
- Hausman, J. and Mcfadden, D. (1984) Specification Tests for the Multinomial Logit Model. *Econometrica* 52 (5), pp. 1219-1240.
- Lam, W. H. K, Ng, P. L. P, Seabrooke, W. and Hui, E. C. M (2004). Forecasts and Reliability Analysis of Port Cargo Throughput in Hong Kong. *Journal of urban Planning and Development* 130 (3), pp. 133–144.
- Lee, S-Y, Chang, Y-T and Tae-Woo, P. (2010). Container port selection factors: heterogeneity among major market player. *Journal of International Logistics and Trade* 8 (2), pp.73-90.
- Magala, M. and Sammons, A. (2008). A New Approach to Port Choice Modelling. *Maritime Economics & Logistics* 10 (9–34).
- Malchow, M. and Kanafani, A. (2010). A disaggregate analysis of factors influencing port selection *Maritime Policy and Management*, 28, 265-277.
- Malchow, M. and Kanafani, A. (2004). A disaggregate analysis of port selection. *Transportation Research Part E* 40, pp. 317–337.

Manski, Ch. and Mcfadden, D. (1981) Structural Analysis of Discrete Data and Econometric Applications. Charles F. and Daniel L. Editors. Cambridge: The MIT Press, 1981.

Moya, J.M. and Feo-Valero, M. (2016): Port choice in container market: a literature review, Transport Reviews: <http://dx.doi.org/10.1080/01441647.2016.1231233>.

Papola, A. (2004) Some developments on the cross-nested logit model. Transportation Research Part B: Methodological, Volume 38, Issue 9, November 2004, Pages 833–851

Song, D-W and Yeo, K-T. (2004) A competitive Analysis of Chinese Container Ports Using the Analytic Hierarchy Process. Maritime Economics & Logistics.6, pp.34–52.

Tiwari, P., Itoh, H. and Doi, M. (2004). Shippers' Port and Carrier Selection Behaviour in China: A Discrete Choice Analysis. Maritime Economics & Logistics 5, pp.23–39.

Tongzon, J. (2009). Port choice and freight forwarders. Transportation Research Part E: Logistics and Transportation Review 45 (1), pp. 186–195.

Tongzon, J. and Sawant, L. (2007). Port choice in a competitive environment: from the shipping lines' perspective. Applied Economics 39, pp 477–92.

Van Canneyt, M., Gartner, M., Heinig, S., Monteiro de Cavalho, F., Ouedraogo, I., Overbeek, D. (2011) Lazarus, the Complete Guide. Blaise Pascal Magazine. ISBN: 978-94-90968-02-1

Veldman, S., Garcia-Alonso, L. and Vallejo, J. (2013). A port choice model with logit models: A case study for the Spanish container trade. International Journal of Shipping and Transport Logistics 5(41398), pp. 373-389.

Veldman, S., Garcia-Alonso, L. and Liu, M. (2015): Testing port choice models using physical and monetary data: a comparative case study for the Spanish container trades. Maritime Policy & Management 3 (4), pp. 495-508.

Yeo, G-T, Ng, A.K.Y., Lee, P. T-W and Yang, Z. (2014). Modelling port choice in an uncertain environment. Maritime Policy & Management 41(3), pp. 251-267.

Anexo C: Port hinterland formation: A Probabilistic Cellular Automaton Approach

ABSTRACT

This article presents a new methodology to define the hinterland of a maritime port, from the exporters' viewpoint. The main assumption is that hinterland grows as the aggregation of zones from which the exporters send their freight to the port. This methodology relies on the theory of cellular automaton (CA). This CA includes a stochastic component to allow for an accurate representation of spatio-temporal aspects that influence the decision makers. The CA is calibrated through a genetic algorithm. This methodology also allows the estimation of port's demand capture as a time-dependent variable after a change in the ports' level of service. This framework was applied to the automobile exports in Brazil, from 1997 to 2015, with a 73% of prediction success.

Keywords: Port choice; port hinterland formation; cellular automaton; probabilistic cellular automaton; choice model; port market share.

1. INTRODUCTION

1.1 Motivation

Demand forecasting for transportation services remain fundamental for port planning strategies. These forecasts are the main input for a port's master plan to justify large investments. The state of practice is to study the potential of markets within a hinterland for various origins and destinations. A port's demand emerges from a spatial body that encloses the commercial activity: its hinterland. The concept of hinterland is an elusive one. Weigend [31] was one of the first scholars to use this concept for freight transportation. Weigend defined a hinterland as an "organized and developed land space which is connected by a port by means of transport lines, and which receives or ships goods through that port". Weigend also pointed out the existence of various hinterlands, defined by the transported goods. The aggregation of the geographical units from/to which traders decided to transport their goods forms the hinterland.

The agents involved in the trading process evaluate various aspects, from different points of view, and make the port choice decision. That decision defines the demand captured by a port as well as its hinterland.

Note that the hinterland plays a key role in the volume of trade through a given port. The port captures demand within its hinterland bounds. These bounds are dynamic and change in time according to the conditions offered by competing ports and network

effects derived from the proximity of traders who choose the same port. The latter is an expression of the Mohring effect [18], in which the presence of other users of a shipping line increases the probability that a line calls at a given port.

Many actors participate in the port choice process. These actors have various viewpoints and they consider different attributes to make their decisions. These attributes and viewpoints can be translated into variables, dependencies, correlations, trade-offs, etc. which form a complex system whose behavior can be observed as a simple result: for a given zone, within a time interval, a port is chosen to trade. Thus, the aggregation of the zones from/to which the shipments depart/end will define the hinterland for exports/imports of a given port. In this article, we analyze the hinterland for exports, i.e. the port choice from the shippers' perspective.

The market share of competing ports depends on several factors such as the frequency of services, charges, accessibility, throughput, etc. However, most of these variables have commercial value to the port operators and consequently are hard to obtain. Consequently, the published articles consider only a few of these variables. In addition, many relevant variables need to be predicted themselves before being used in a forecasting model. The latter makes their use almost impossible in practical cases. The complexities in the freight transport system make forecasting a very difficult task, in part because of the network effects both in time and in space.

Forecasting is complicated also because commodities are available in many different places around the globe with different prices and lead-times, generating a large set of alternatives to choose from. To simplify the econometric aspects associated with these problems, standard forecasting models tend to neglect relevant interactions, such as shippers/carriers bargaining power, long-term contracts, and vertical integration, among others. The effect of neglecting these interactions is not fully understood yet.

In this article, we develop a modeling framework that incorporates complex interactions between actors where space and time are explicitly considered. The main objective of this article is to present a new modeling framework to identify a hinterland in terms of the choice of port made by shippers within a ports' disputed region. Having such a tool would be of great relevance for methodological and practical reasons alike. For example, from the theoretical perspective, such a model could aid in the search of complex patterns (chaotic behavior under specific initial conditions) in the dynamics of hinterland shapes, as well as identifying attributes that trigger such behavior.

On the applications' field, this framework is useful to establish not only the current hinterland of a port but also the variables that may increase or decrease that hinterland within a competitive ports' system.

We developed a methodology that not only relies on the well-known explicative variables, such as location and level of service of a port (see Refs., [13, 15, 24; 26]), but also rests on the hypothesis that the shippers' decision is directly affected by spatial and temporal interactions. The influence of some shippers in the decision of other shippers within their neighborhood is the root of these spatial interactions. As per the time interactions, they appear as a dynamic result of the diffusion effect of the neighbors' decisions throughout the network of decision makers. In fact, there is empirical evidence of the neighbors' effect in the field of international freight transportation. Large shippers do affect the decision of ocean carriers, which decide to call at certain ports where the volume of these shippers makes that call an attractive option (see, Refs., [1,24]). Once an ocean carrier decides to call on a given port, smaller shippers can benefit from the new available capacity at lower prices (Mohring effect). When the number of new shippers reaches a certain level in that port, their bargaining power increases and hence a spatial interaction between large, medium and small size shipper emerges. That virtuous circle can also disappear if, for example, a larger shipper decides to take its business elsewhere.

These interactions are extremely difficult to incorporate in standard port choice modeling tools [5]. In order to capture the complex spatio-temporal interactions, we build a cellular automaton model, which is simple in structure and can be quickly calibrated and operated.

2. LITERATURE REVIEW

The objective of this Chapter is to present the most relevant literature related to our topic of interest. We start with the general concept of hinterland, followed by three sections describing modeling aspects of the choice of ports from the various actors' perspectives. The last two sections present methodological developments from the modeling perspective in the field of port choice.

2.1 Port Choice from the shippers' viewpoint

From the shippers' viewpoint, the location of a port is one of the most commonly mentioned variables in the specialized literature. Other relevant variables are the type of commodity, transportation services (door-to-door), among others (see [15]).

Magala and Sammons [14], analyze the importance of the supply chain characteristics in the port choice. Even though the ports are considered an integral part of the supply chain, instead of isolated nodes, Tongzon [26] indicates that in many cases (especially in

developing countries), the ports are not fully integrated in the supply chains; consequently, they play different roles depending on their location around the globe. Tongzon [26] describes three types of shippers: those with long-term contracts with ocean carriers, independent shippers and those who rely on freight forwarders. He studied the latter, identifying the main factors that trigger the port choice decision: port efficiency, carriers' frequency, infrastructure quality and location.

Steven and Corsi [23], divide the port choice process into two categories: large and smaller shippers. Large shippers give the port choice responsibility to the carriers while the smaller size shippers make the port choice decision themselves. These researchers found that large shippers are more sensitive to factors that may improve the delivery time, while the smaller shippers are more concerned with direct transportation costs.

Halim et al. [9], point out the relevance of the costs associated to the hinterland with respect to the total costs, even though the travelled distance may represent only 10% of the total origin-destination distance. They also point out the importance of the port connectivity with its hinterland, which is a key variable for the port choice decision along with the optimal routes followed from the shipper's location to the port. These authors developed a strategic freight distribution model, based on total logistics cost and service level requirements. The attractiveness of a port increases with the availability of transshipment services, which often enhances both the capacity and speed of cargo processing. Jiang et al. [11] studied these aspects and developed two mathematical programming models for minimizing transportation time and maximizing transportation capacity for a network of ports. The authors pointed out the key role of alliances as critical elements in maritime transportation, especially for port connectivity (which is a driver for selection). Although they did not consider the effect of alliances in their study, they acknowledge the potential impact of the agreements between partners within the alliances (long-term contracts).

2.2 Port Choice from the carriers' viewpoint

Tongzon and Sawant [25], studied the port choice from the perspective of the shipping lines based on a revealed-preferences approach. They found that the fundamental decision variables are port charges, as well as a wide range of port services. They also pointed out that two ports with overlapping hinterlands, operating at high levels of efficiency would need lower charges, provision of new value added services and adequate infrastructure for gaining a competitive advantage over the other. Other authors (e.g., [1]) found evidence that at least six variables are significantly relevant in the port decision process. These variables are local cargo volume, port charges, berth availability, port location, transferred volume and feeder connectivity.

2.3 Port Choice from other viewpoints

Various researchers have studied the problem of port choice when several agents are involved in the process. Song and Yeo [22] presented a methodology based on opinions of ship-owners, shippers, shipping companies, terminal operators, as well as researchers and scholars. They managed to identify 73 characteristics that enable global competitiveness in the main Chinese ports. Among these characteristics, they pointed out that location was the single most relevant variable as far as ports' competition is concerned.

Lirn et al. [13] studied the variables that influence the port choice process from the carriers and ports operators' viewpoint. They applied the Analytical Hierarchy Process (AHP) method to 47 selected relevant service attributes established from a literature review. They categorized these attributes within four criteria: physical and technical infrastructure, geographical location, management, and terminal cost. The results of the AHP analysis showed that both container carriers and port service providers have similar attributes (as far as port choice is concerned), but with different weights each. These attributes are the following: handling cost, proximity to main navigation routes, proximity to import/export areas, infrastructure condition, and feeder network.

Lee et al. [12], studied the main drivers of port choice from the viewpoints of various actors: shipping companies, shippers, and terminal operators. In their study, they identified 38 attributes within seven categories (port location and characteristics, port's operating conditions, hinterland characteristics, among others). Each actor assessed the attributes from their own perspective. While most of the attributes' importance presented considerable variation between actors, all of them considered "port accessibility" as a relevant attribute. Shippers and terminal operators shared most of the attributes' assessment, but terminal operators considered the attribute "port's operating conditions" significantly more relevant than shippers did.

Recently, Song et al. [21], analyzed ports competitiveness, involving hinterland shipments and transshipments. They studied the corresponding supply chain, considering a logistics-cost function that incorporates port handling charges, deep-sea transportation cost, hinterland transportation cost, and feeder service cost. Their method consisted in applying a game cost model between two competitive ports and one ocean carrier offering multiple shipping services.

2.4 Port choice modeling frameworks

Discrete choice modeling [16], stands as the preferred family of mathematical models to study port choice processes. This modeling framework assumes a set of rational agents with perfect information that must decide which one of various available ports would be the optimal choice, given a set of attributes of ports and agents. This approach assumes that a finite set of attributes can completely describe the available ports. Each agent has a *random utility function* (RUF) that depends on the ports' attributes, agent-specific characteristics as well as a stochastic component that explains possible deviations from rational behavior. Thus, each agent chooses a port according to a probability distribution, found by maximizing the RUF.

The assumption of the RUF's stochastic component defines the functional form of the resulting model. The most common assumption for this component leads to the well-known Multinomial Logit model (MNL, see [10]). However, the assumptions that lead to this simple structure are strong: there is no correlation between available alternatives, and the RUF's stochastic component must be homoscedastic (i.e. the variance of all the alternatives must be equal).

Veldman et al. [29], Tiwari et al. [24] and Malchow and Kanafani [15] present various examples of this methodology (applied to the port choice problem). While some of these models show a good performance in terms of prediction power, is likely that the reported performance is significantly biased. Indeed, the latter occurs because the cases studied with these models do not fulfill the most important MNL requirements outlined above.

Therefore, the econometrics followed to estimate/calibrate their parameters would typically produce biased and inconsistent estimators. Indeed, the choice set in the case of ports, bears a high correlation and heteroscedasticity among the different competing ports. The latter occurs because of the presence of various aspects, such as spatial proximity between terminals, long-term contracts between ocean carriers, ports and/or large shippers, etc. The solution to these practical inconveniences for the MNL would be to use a more general set of assumptions for the stochastic components.

Veldman et al. [28], partly tackle this problem. In fact, these authors solve only a portion of these problems by using a Hierarchical logit model (HL, see [20]). Nevertheless, the HL models can only take care of a small fraction of the correlation/heteroscedasticity problem. The latter is due to the rigid structure of the variance-covariance matrix that the HL imposes.

A novel approach has recently been published [30], where the hinterland of a given port is estimated probabilistically. In that study the authors combine a discrete choice model with the shippers' spatial location, and developed an algorithm that identifies the boundaries of the hinterland as a function of the port choice probabilities.

Yeo et al. [32], proposed a port choice model based on fuzzy logic. The authors combined different data sources and formats, both objective and subjective, to assess the general performance of a port, measured by a "fuzzy score". Their model incorporates degrees of belief to conform a criterion linked to the port choice decision.

The model's result is a preference score that allows the comparison of ports performance, either against other competing ports or with the system as a whole. It also allows comparing a port with itself in previous periods. The authors studied the choice of container ports in northeast Asia with this approach. This methodology however, is highly dependent on the opinions of the interviewed stakeholders. However, these stakeholders may change their opinions for multiple reasons (political bias, strategic behavior, etc.). The latter gives this method a subjective dependence. This dependence weakens the models' capacity to guide decision makers in infrastructure investment or public policy design.

To the knowledge of the authors, the modeling methodology proposed by Garrido and Leva [5], is the only discrete choice structure capable of incorporating alternatives with a general interaction in time and space. In that study, the authors presented a rather general Multinomial Probit model (MNP, see [4]). That MNP model incorporates spatial and temporal correlations between alternatives and a general structure for the variance-covariance matrix that could be adapted to the ports choice case. Nevertheless, the resulting model was rather cumbersome and computationally demanding; the model does not have analytical form and consequently it must be estimated/calibrated using numerical integration in high dimensions. These complications make the application of this model very difficult in practice.

2.5 Cellular automaton modelling

Cellular automaton (CA) are mathematical models (usually related to computational physics) that date back to the 1950's, when first introduced by John von Neumann. A standard CA is a mathematical entity formed by composed of four components: multiple units called "cells", a number of discrete states (e.g. active or inactive), a neighborhood, and transition rules. Each cell interacts locally, and its next state depends on a function of the current state of itself and its neighbors. The CA usually operates on a grid of a known shape. The set of cells form a system, which evolves in a complex manner. This

modeling approach is especially suitable for processes where the immediate neighbors have an influence on the cell, such as urban systems. It has been successfully applied for modeling complex, self-organized and emergent systems.

Take for instance the famous “game of life”, created by John Conway [3]. This game consists of a collection of cells located in an infinite two-dimensional orthogonal grid. These cells can live, die or reproduce. The system of these cells can form numerous patterns during the game. These patterns strongly depend on the cells’ initial conditions. Each cell interacts with its eight neighbors (cells that are horizontally, vertically, or diagonally adjacent). At each time step, the cells’ states may change, according to the following rules:

- Living cells with less than two living neighbors die, as if caused by solitude.
- Living cells with two or three living neighbors live on to the next time step.
- Living cells with more than three living neighbors die, as if by saturation.
- Living cells with exactly three living neighbors become living cells, as if by reproduction.

The initial state of each cell constitutes the seed of the system. Applying these rules simultaneously to all the cells in a given stage creates a new generation. All the births and deaths occur simultaneously. Further generations result from applying these rules repeatedly. This example has all the ingredients that a typical CA contains: a set of homogeneous cells, a set of states for these cells, a spatial arrangement, a definition of neighborhood and a transition rule for the states. These ingredients have proven to be a very flexible modeling tool for dynamic or complex systems, and they are especially efficient to study dynamic systems that evolve in discrete steps (see for example, Refs., [2,8]).

To our knowledge, the specialized literature does not show articles that incorporate the theory of CA to study the hinterland/port choice problem. However, the dynamics found in the hinterland determination problem seems to be a process that CA could replicate efficiently. In fact, a CA specification can incorporate the actors involved in the port choice process as well as the feedback observed in their decisions.

The CA theory is especially adequate to deal with dynamic, non-linear and complex phenomena [7]. Applying CA theory to this problem would allow the modelers to specify computationally efficient models that are also less data-intensive (data are always scarce as a public resource in this field).

3. DESIGN OF A CELLULAR AUTOMATON FOR HINTERLAND DEFINITION

3.1 Structure and characteristics of a hinterland definition probabilistic cellular automaton (HDPCA)

The main problem that we aim to solve is that of identifying the geographical sites that conform a port's hinterland. The aggregation of the port choice decisions of various exporters in a given region will determine this hinterland. Therefore, we need to design a CA structure capable of predicting the choice of a port of call by a set of shippers. These shippers, spread within a bounded region, share a significant portion of the hinterland.

We assume that various ports compete for the hinterland and all of these ports are available to the shippers. We divide this region into n mutually exclusive and collectively exhaustive freight generation zones (FGZ). There are m ports and each FGZ can take one of $m+1$ possible states of activation: one for each available port and one state of inactivity. Thus, if a FGZ "has chosen" port i (with $i = 1$ to m) as a port of call, then its current state is i , and if the FGZ has not chosen any port then it is inactive, i.e. its current state is 0 , which means that no exports departed from that FGZ to that particular port at this stage.

Figure 1 shows a scheme of the FGZ and ports on which a *Hinterland Definition Probabilistic Cellular Automaton* (HDPCA) model will operate. Each circle represents a FGZ, which corresponds to a single cell of the HDPCA, which can be either active or inactive. A given FGZ is active during discrete time t if it uses an available port to export.

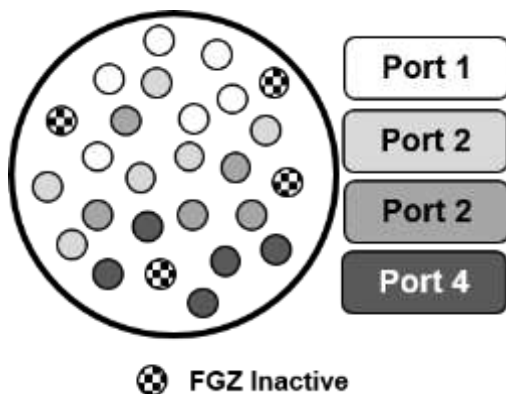


Fig. 1. Conceptual scheme of Freight Generation Zones and Ports within a shared hinterland.

The condition of activity at time t depends on two aspects: the operational characteristics, and the spatial interaction. The operational characteristics are attributes that conform the utility derived from choosing a given port. Examples of these attributes are the distance to each available port, charges, connectivity, transfer capacity, long-term contracts, bargaining position, among many others. As per the spatial interaction, the states of the neighboring FGZ will affect the state that a given cell takes. Thus, the i -th FGZ has a state function defined as follows:

$$e_i = \begin{cases} j, & \text{with } j = 1, 2, \dots, m \\ 0 & \text{otherwise} \end{cases}$$

where 0 means inactivity and j means that i -th FGZ has chosen port j .

As in any other complex system, none of the actors can describe the port choice process completely at a given time. Therefore, various aspects of this complex decision are unknown to a modeler, even though she might know many of the attributes that actors evaluate before their choice. To circumvent this problem of incomplete knowledge, we incorporate a stochastic component within the decision framework. This stochastic component does not imply that the port is randomly chosen by the incumbents. Instead, it indicates that in the absence of perfect information by the modeler, the port choice cannot be deterministically defined by the incomplete set of attributes available to the modeler.

This assumption would explain at least two phenomena that otherwise may seem erroneous from the modeler's perspective: first, an exporter may choose a port that is not his best choice according to the measured attributes. Second, two "identical" exporters (from the attributes point of view) may choose different ports.

Both phenomena can be described as rational decisions, when the modeler knows only a portion of the attributes that trigger the port choice process. Thus, a FGZ can be active/inactive as a results of the interaction of exogenous variables not included in the deterministic model specification.

Typical CA models need some fixed rule to create a new generation in discrete steps. In this case, the rule is a transition function that determines the new state of each cell as well as the states of the cells in its neighborhood, in terms of the current state of the cell itself and the states of other spatially correlated cells (neighborhood).

3.1.1 The neighborhood concept in the HDPCA model

Setting the concept of neighborhood in CA translates into defining *ex-ante* the spatial influence among cells. Two of the most commonly found examples of neighborhood are those defined by Von Neumann and that of Griffeath and Moore [8]. The former considers, in a rectangular grid, that neighboring cells share a common edge (four neighbors in a squared grid), while the latter considers as a neighbor each one of the eight cells surrounding a given cell in a squared grid. For the port choice problem, the FGZ are not necessarily ordered in a Manhattan metric array. Consequently, to fit the actual spatial pattern observed in practice, we incorporate the concept of radial neighborhood, in which two cells are neighbors of radius R if both are circumscribed within a circle of radius R , as shown in Figure 2.

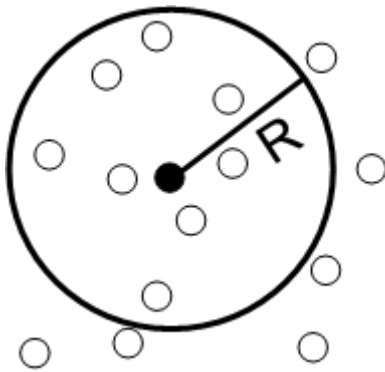


Fig. 2. Radial neighborhood of radius R for the HDPCA

3.1.2 Transition function

The cells in the HDPCA model correspond to FGZ that transit from their current state to the next according to a stochastic transition rule. In general, stochastic CAs update their state according to a stochastic rule derived from a certain probability distribution. Incorporating stochasticity in the spatial interaction among cells, even under simple updating rules, allow these models to replicate emergent properties at system's level.

The HDPCA will update the state of each FGZ according to the following rules:

1. Each FGZ has a probability (μ) to activate itself. This probability is found by calibration through a genetic algorithm (see section 3.1.3 below). To impose an activation with probability μ , at each step the HDPCA draws a real number z from a $(0,1)$ uniform distribution, and compares z against μ . If $z \leq \mu$ then the FGZ becomes active (and it does not activate otherwise).

2. The i -th FGZ has an associated function $\Phi^i \in \mathbb{R}^m$ whose domain is the R -neighborhood of the i -th FGZ. This function's outcome is a vector of m components.

$$\Phi^i = \alpha \Phi_1^i + \beta(d_{i1}, d_{i2}, \dots, d_{im}) + (\gamma_1, \gamma_2, \dots, \gamma_m) \quad \forall i = 1, 2, \dots, n \quad (1)$$

where d_{ij} represents the distance from the i -th FGZ to the port j (with $j=1, 2, \dots, m$) and the constant parameter γ_j represents other systematic attributes related to the attractiveness of a port. Φ_1^i is an m -dimensional vector, formed by the ratio between the number of FGZ within the R -neighborhood that choose the port j , relative to the total number of active FGZ within the same R -neighborhood. i.e.:

$$\Phi_1^i = \left(\frac{n_1}{N_i}, \frac{n_2}{N_i}, \dots, \frac{n_m}{N_i} \right) \quad \forall i = 1, 2, \dots, n \quad (2)$$

where $n_1, n_2, \dots, n_m$ is the number of FGZs that chose the port 1 to m respectively in a given time step, and N_i is the total number of active zones within the R -neighborhood of the i -th FGZ.

According to the specification (1), the parameter α represents the influence of the neighbors' behavior on the port choice decision of the i -th FGZ, i.e., the spatial component of that decision. The parameter β represents the influence of the distance on the port choice decision of the i -th FGZ. β should take a negative value when calibrated, because it weighs a disutility (the longer the distance the lower the utility), whereas α should take a positive value, as it weighs a proxy of the ports' market share.

3. Each component of the vector function Φ^i represents the utility or attractiveness of choosing the i -th port. These components are mapped as follows: its biggest component is assigned a probability p , its second biggest component is assigned a probability $p(1-p)$ and so forth decreasing as a geometric law until reaching $p(1-p)^{m-1}$. This mapping avoids the fact that the port with best market share and lowest distance is always chosen. Instead, the highest utility is assigned a probability of being chosen.
4. The state e_i of the i -th FGZ is updated with a value between 1 and m , representing the outcome of a choice given the probability's described on rule in the above paragraph.

3.1.3 A genetic algorithm for model calibration

The HDPCA described above relies on the specification of a set of parameters to be of practical use. The calibration process, performed through a genetic algorithm (see Refs. [6,17]), is described in this section.

The data set corresponds to a series of records for the automobile industry in Brazil. The series span from 1996 to 2015, divided into trimesters. Thus, the dataset contains 76 time periods. The calibration considers the four main ports for automobile exports. The region was divided into 69 freight generation zones (FGZ). The Hinterland Definition Probabilistic Cellular Automaton model (HDPCA) contains various parameters to be calibrated before its implementation in a practical case (see section 3.1). These parameters are the following:

- R : The radius of neighborhood for all the FGZs.
- α : The weight of the neighbors' port choice decision.
- β : The weight of the distance from FHZ to the port.
- γ : The systematic preference component (m-dimensional vector).
- μ : The activation probability for the FGZ (m-dimensional vector).

To calibrate these parameters' values, we designed a genetic algorithm implementation that runs as follows:

1. We define an initial population of 100 individuals. Each one of these individuals represents a set of values for the HDPCA's parameters. These initial values are randomly selected for the first generation.

$$I_1 = (R_1, \alpha_1, \beta_1, \gamma_1, \mu_1, p_{1-1}, p_{1-2}, p_{1-3}, p_{1-4}, z_{1-1}, z_{1-2}, \dots, z_{1-69})$$

$$I_2 = (R_2, \alpha_2, \beta_2, \gamma_2, \mu_2, p_{2-1}, p_{2-2}, p_{2-3}, p_{2-4}, z_{2-1}, z_{2-2}, \dots, z_{2-69})$$

...

$$I_{100} = (R_{100}, \alpha_{100}, \beta_{100}, \gamma_{100}, \mu_{100}, p_{100-1}, p_{100-2}, p_{100-3}, p_{100-4}, z_{100-1}, z_{100-2}, \dots, z_{100-69})$$

2. The phenotype associated to each individual corresponds to the observable outcome: number of active FGZ, chosen port for each FGZ, inactive FGZ, market share for each port, etc. For each phenotype, various indicators can be computed to measure the goodness of fit between modeled choice and actual data.

3. The selection of the optimal parameters' values must be done in terms of the best fit to the actual data. Therefore, we defined a score to be optimized as follows:

$$F_0 = \sum_{i,t} f_0(i, t) \quad (3)$$

$$f_0(i, t) = \begin{cases} \delta & \text{if } (e_i = er_i \neq 0) \\ 1 & \text{if } (e_i = er_i = 0) \\ 0 & \text{otherwise} \end{cases} \quad \delta \in \mathbb{R}^+ \quad \forall t \quad (4)$$

4. The best set of values is such that F_0 is maximum. In fact, each time the set of values yields a successful assignment (i.e. the simulated port choice was actually chosen by a given FGZ at a given time period), the value of $f_0(i, t)$ increases.
5. The variables e_i and er_i in expression (4) represent the state of the i -th FGZ in the current simulation and the actual record in the database respectively. Therefore, when both variables are equal, then the HDPCA has successfully replicated a port choice decision.
6. δ is an auxiliary parameter that allows an extra degree of freedom to privilege those individuals that are more successful to replicate active FGZ instead of inactive ones, i.e., if the model correctly predicts an active FGZ, its success is valued δ times more than a success in the prediction of inactivity of a FGZ.
7. Once the model has evaluated the objective function F_0 , the 100 individuals are arranged in a descending order. Then, crossover operators are defined between the top 30% of the list and the rest of the population. For example, Figure 3 shows the crossover between two "parents" (individuals 1 and 2) and the "son" who survives to the next generation. The point where the genetic code is cut-off is randomly chosen. Thus, Figure 3 shows only one possibility among many others (e.g. a crossover could include the combination of all the parents genetic code).

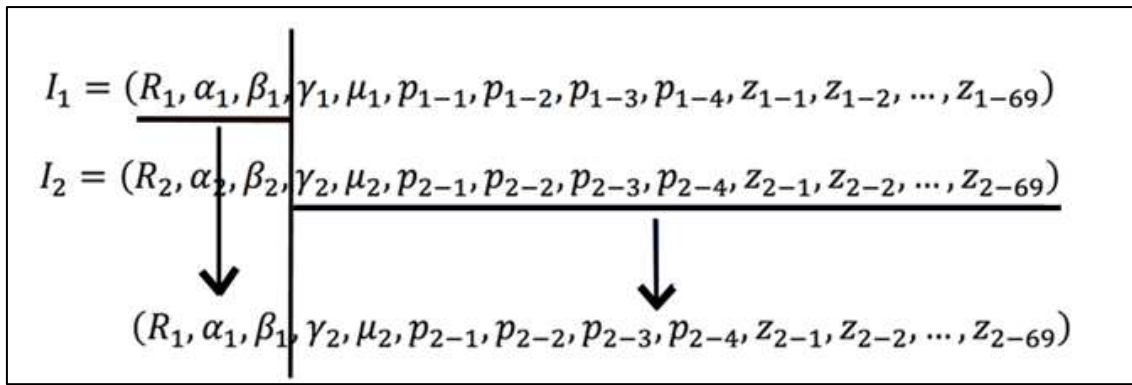


Fig. 3. Schematic crossover between individuals 1 and 2, belonging to generation t , and their offspring for the next generation ($t+1$).

8. To incorporate diversity in the next generations we defined mutation operators. Figure 4 shows the schematic idea of this mutation process. The gene to be mutated is chosen randomly.

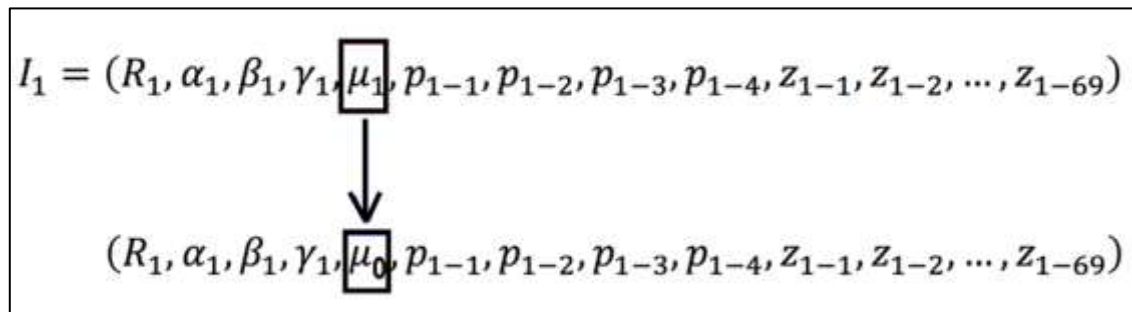


Fig. 4. Description of the mutation operator. This example shows individual 1 from a generation t and the mutant individual in generation $t+1$.

9. The quality of each individual is evaluated along a diversity descriptor, to guarantee a heterogeneous population at each generation. Thus, at each future generation the “best” individual of the previous generation is preserved.
10. Convergence criterion: the genetic algorithm stops when a given percentage of success is reached during the evolution. This percentage is an input from the modeler.

4. EXPERIMENTAL DESIGN

4.1 Application instance

The methodology described in the previous sections uses three data sets. First, a set of exporters spatially located in a certain region. Second, a list of the ports chosen by each exporter at each time interval. Third, a set of ports' attributes that proxy their level of service. There are not many countries that keep track of this information. One exception is Brazil, which has a publicly available dataset with the needed information at disaggregate level (<http://www.aliceweb2.mdic.gov.br/>).

Among the data available in Brazil, the roll-on roll-off (RORO) market is one of the most dynamic within the exporters. For that reason, we have decided to study the case of automobile exports. In this industry, the RORO vessels operate mainly in four Brazilian seaports. Those ports compete for (approximately) the same hinterland, i.e., the four ports are available to all of the exporters located in that hinterland. These ports are the following: Paranaguá, Santos, Sao Sebastiao and Rio de Janeiro. In this case, the FGZ correspond to *municípios*, i.e. administrative geographic divisions.

The dataset contains records for automobile exports from 1997 to 2015 through these four ports. The temporal division is the trimester, thus generating a time series of 74 observations. The spatial data accounts for 69 FGZ in the region of analysis. These FGZ are not uniformly distributed in the whole region. In fact, there are clusters corresponding to logistics nodes for the Brazilian automobile industry. At each trimester, a FGZ can be “on” or “off” depending on their exporting status. The percentage of active FGZ per period is depicted by the scattered triangles in Figure 5, while the smooth line shows a fitted curve to the data.

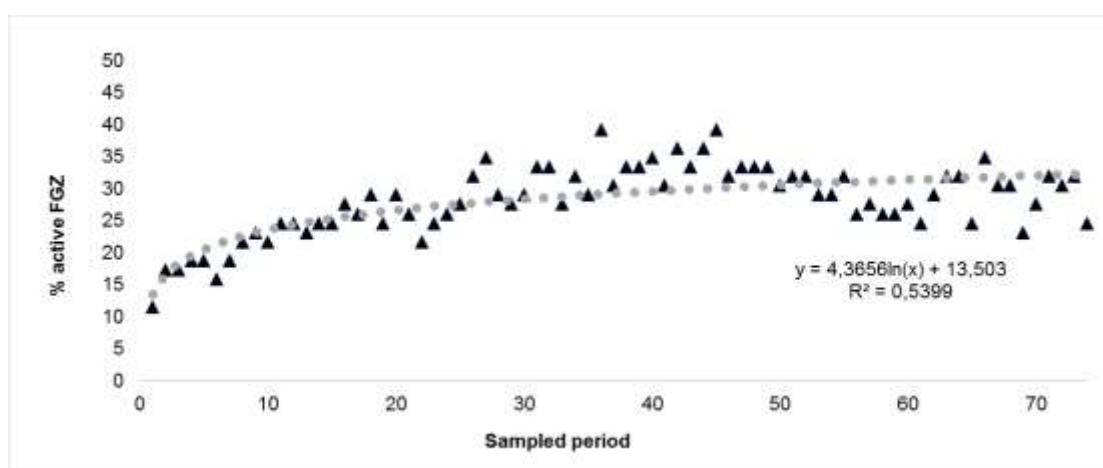


Fig. 5. Percentage of active FGZ at each sampled period

Figure 6 shows the historical market share of each port; it can be seen that Port 1 (Sao Sebastiao) has a very low share compared to its competitors, in spite of being the closest to the highest market share (Santos). The latter can be explained for the lack of attractiveness of Port 1 in terms of level of service attributes (e.g. infrastructure, accessibility, technology, charges, etc.).

Regarding the location of the FGZ and their distance to the competing ports, 63% of them are closer to Port 3 (Santos), 23% are closer to Port 2 (Rio de Janeiro), 12% are closer to Port 4 (Paranaguá), and only 4% are closer to Port 1 (Sao Sebastiao). These proportions follow a similar trend to the market shares of each port, shown in Figure 6 (i.e. 71% Port 3, 19% Port 2, 10% Port 4, 1% Port 1). The latter is an indication of the role that distance plays in the exporters decisions.

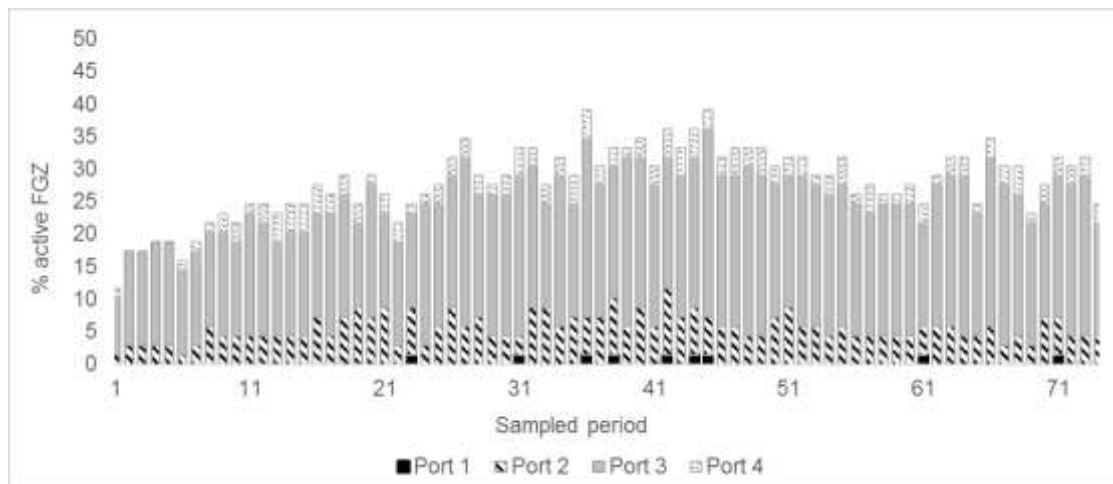


Fig. 6. Historical market share of each port in the choice set

4.2 Calibration Results

The HDPCA parameters were estimated through a genetic algorithm (see Section 3.1.3). The calibrated values are presented in Table 1. From the econometric perspective, it is worth mentioning that as γ_i is a constant parameter (i.e. fixed for each port) there is a maximum of $(m-1)$ identifiable values for them. Consequently, three of these parameters are estimated. Hence, the value of γ_3 was set at 0 (with no loss of generality). For the same econometric reason, the value of the probability p was fixed (set at 0.75).

Table 1

Estimated Parameters for the Hinterland Definition Cellular Automaton Model

Parameter	Value				
<i>Radius</i> (km)	267	μ_{27}	0.004	μ_{60}	0.89
α	828	μ_{28}	0.01	μ_{61}	0.88
β	-12	μ_{29}	0.01	μ_{62}	0.66
γ_1	-818	μ_{30}	0.07	μ_{63}	0.04
γ_2	-417	μ_{31}	0.82	μ_{64}	0.93
γ_4	-297	μ_{32}	0.39	μ_{65}	0.99
μ_1	0.17	μ_{33}	0.55	μ_{66}	0.01
μ_2	0.06	μ_{34}	0.55	μ_{67}	0.03
μ_3	0.56	μ_{35}	0.05	μ_{68}	0.01
μ_4	0.27	μ_{36}	0.09	μ_{69}	0.37
μ_5	0.05	μ_{37}	0.15		
μ_6	0.02	μ_{38}	0.17		
μ_7	0.00	μ_{39}	0.06		
μ_8	0.76	μ_{40}	0.21		
μ_9	0.08	μ_{41}	0.12		
μ_{10}	0.01	μ_{42}	0.13		
μ_{11}	0.08	μ_{43}	0.84		
μ_{12}	0.03	μ_{44}	0.22		
μ_{13}	0.37	μ_{45}	0.55		
μ_{14}	0.05	μ_{46}	0.17		
μ_{15}	0.01	μ_{47}	0.54		
μ_{16}	0.52	μ_{48}	0.04		
μ_{17}	0.17	μ_{49}	0.07		
μ_{18}	0.95	μ_{50}	0.06		
μ_{19}	0.53	μ_{51}	0.42		
μ_{20}	0.04	μ_{52}	0.97		
μ_{21}	0.05	μ_{53}	0.09		
μ_{22}	0.09	μ_{54}	0.16		
μ_{23}	0.13	μ_{55}	0.34		
μ_{24}	0.27	μ_{56}	0.13		
μ_{25}	0.82	μ_{57}	0.83		
μ_{26}	0.02	μ_{58}	0.83		
		μ_{59}	0.15		

Usually, the calibration processes have standard measures of success. For example, the success of the least squares method in linear regression is measured by the coefficient of determination R^2 . In this case, there is no standard indicator to gauge the success of the model's calibration. Thus, we used the *weighted percentage of success* (WPS) defined as a weighted proportion of success in predicting the chosen port at a given time for each FGZ.

We chose a weighted measure instead of a simple proportion because it is easier for a model to assign parameters' values that replicate inactive FGZ (i.e. both the model and the observed data coincide in that the FGZ is "off"). Indeed, the majority of the FGZ are inactive during any single period, and therefore targeting inactivity only, could yield a high proportion of success, to the expense of a lower success in predicting active FGZ.

Consequently, the WPS was computed assigning a weight of 1.0 to an inactive success, while a weight of 4.0 was assigned to an active success. The value of this weight is designed by δ as a coefficient for calibration.

The genetic algorithm found an optimal neighborhood radius of 267 km. Within this radius, 83% of the FGZ have five or more neighbors. The latter is an indication of a clustering effect, which should be reflected by the value of the spatial parameter α . Indeed, the optimal α value was positive and high, indicating a strong influence of the neighbors' decisions on the individual decision at any given FGZ.

The genetic algorithm reached a 73% of WPS in the calibration process. This is a high degree of success considering that there are 69 FGZ simultaneously making decisions during 74 trimesters. This rate of success reached by this methodology seems high when compared to the success of other econometric models in the same field, calibrated with actual data.

Nevertheless, the comparison between alternative models (such as discrete choice models) must be done indirectly, because those models do not share the same statistics (to measure their goodness of fit) with other modeling structures. In fact, the natural candidates to compare the predictive performance are the discrete choice specifications.

These models use an indirect measure of their goodness-of-fit/forecasting performance based on the proportion of likelihood gained by the specified model in contrast with a constant model that reproduce always the fixed market share (named "constants-only model").

The latter is expressed as the ratio of the log-likelihood at the optimal solution and the log-likelihood of the constants only model, and it is referred to as rho squared (ρ^2); a ρ^2 value of 1.0 would mean that the logit specification captures the total likelihood of the sample (a proxy for perfect fit), whereas a ρ^2 value of 0 would mean that no additional explanation is found by using the logit specification instead of a constant-only model.

Table 2 shows the best results found in the literature for the logit model's goodness of fit in this field. As can be seen, most of the values are near 0.5, with the only exception of the Nir et al's [19] model with a $\rho^2=0.645$. It is interesting to note that even the hierarchical logit model (HL), with a more general specification, does not reach the best performance. These results give an indication that the HDPCA calibration present high rate of success when predicting the outcome of the chosen sample.

Table 2

Goodness of fit of various logit models applied to the port choice problem.

Author	Model	ρ^2
Nir et al. [19]	MNL	0.645
Tiwari et al. [24]	MNL	0.57
Malchow and Kanafani [15]	MNL	0.54
Vega et al. [27]	HL	0.568
Veldman et al. [28]	MNL	0.377
Veldman et al. [29]	MNL	0.507

Figure 7 shows a histogram of the FGZ's activation probability. It can be seen that only 29% of them have a probability higher than 0.5, indicating a relative resistance of those FGZ to become active at each period. The latter is also reflected in the actual data where about 30% of the FGZ remain active during the whole sampled time span.

This dynamic characteristic is easily represented by the HDPCA, but it would not be so in the case of standard modeling approaches, such as discrete choice or regression modeling, because those approaches tend to produce larger errors for agents with low level of activity, expressed as a low presence in the calibration samples.

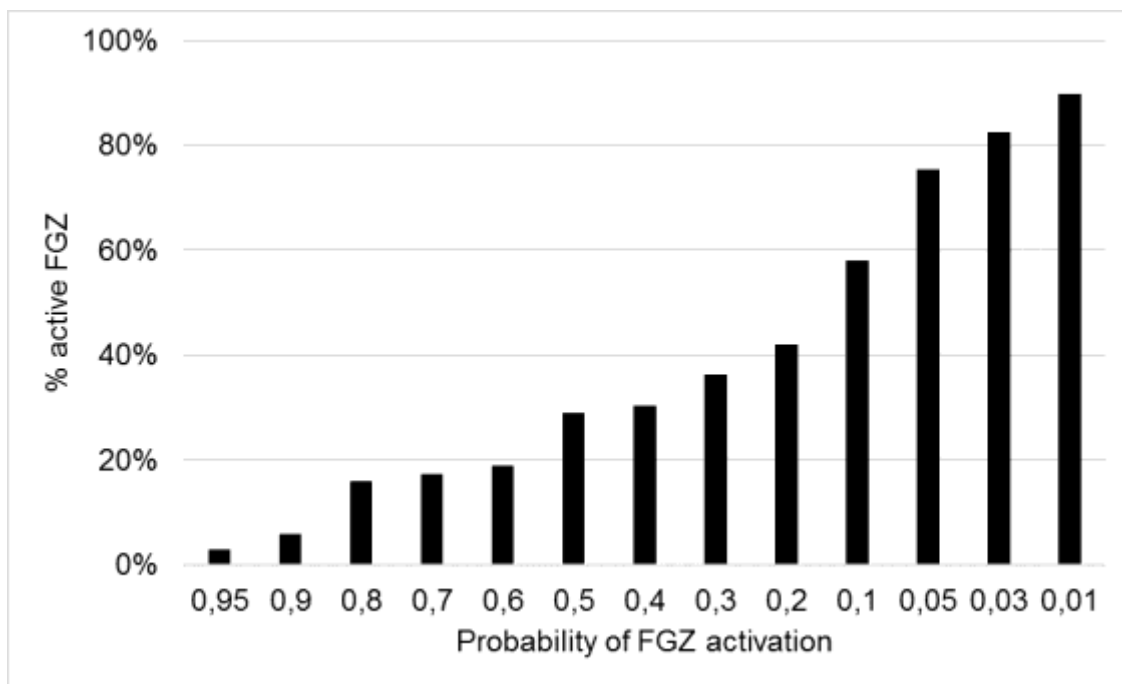


Fig. 7. FGZ's Activation Probability

4.3 Measuring the HDPCA's Performance

To evaluate the model's performance as a forecasting tool we carried out several runs of the model. The change in the FGZ's activity is explained mainly by the term $\alpha\Phi_1$ in the Φ function (the spatial effect of the neighborhood) and the randomness added by the threshold condition for activation. The simulation periods were set to 80 trimesters, and within each period the model was ran 20 times, in order to capture the long-run emergent behavior of the system.

Figure 8 shows the percentage of active FGZ for each simulation period (trimesters). The scattered dots show the actual simulation results and the solid line shows the average over the 20 runs per period.

The dynamic behavior depicted in Figure 8 shows temporal stability, in a scenario where the ports' level of service remains unchanged. The average number of active FGZ is 29.8% versus 28.1% in the observed data. In addition, the dynamics of the process is accurately captured by the HDPCA. Indeed, in the latest 10 years the difference between the actual percentage of active FGZ and those obtained from the simulator is only 0.9 points (29.9% from the HDPCA versus 30.8% in the observed data).

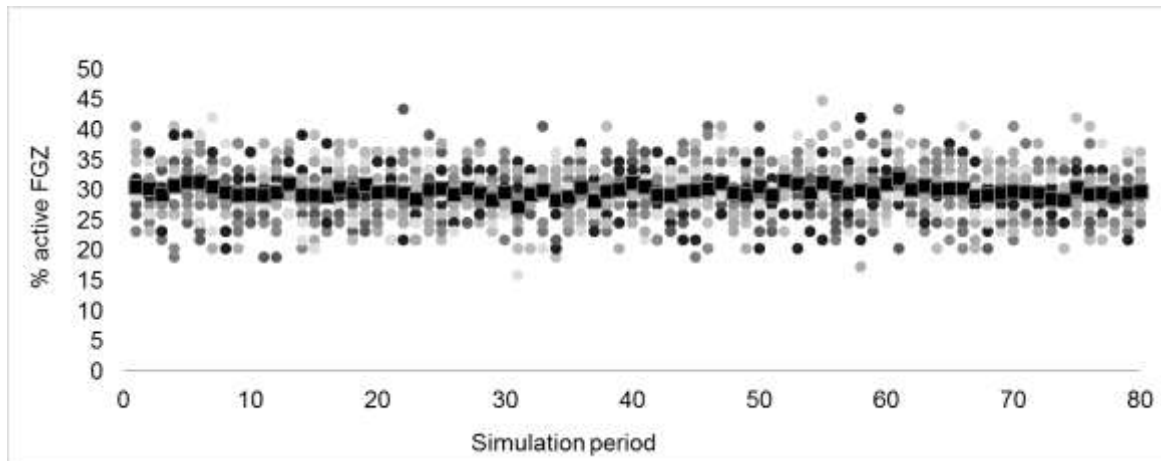


Fig. 8. Forecasting of active FGZ for a 20 years' simulation horizon

The modeling tool also showed accurate market share estimations. The greatest variation with respect to the observed data was found in Port 3 with a difference of 1.5 percentage points, which is considered highly accurate in the practical field. Table 3 shows the results for the four ports.

Table 3

Market shares with respect to the total number of FGZ.

	Port 1	Port 2	Port 3	Port 4	Active FGZ
Simulation	0%	5.9%	21.3%	2.6%	29.8%
Real data	0.2%	5.4%	19.8%	2.7%	28.1%

4.4 Model's performance as a predictive tool

One of the many analyses that can be performed with a tool like the HDPCA, is the forecasting of market shares as a result of a change in the level of service of one or more ports. In this particular model, the ports' characteristics are represented by the parameters γ_i . Changing the values of γ_i is equivalent to alter the attributes that give each port the level of service that is not related to distance. A better level of service is reflected in a greater value of its parameter γ_i .

As an example, we studied the behavior of the whole system of ports if the level of service of ports 1 and 3 were equal. Both ports share the same hinterland (they are approximately 160 km apart from each other), and consequently if the exporters faced the same level of service, it would be expected that Port 1 would increase its market share. When this scenario was simulated, the increase for Port 1 was relatively small (only 1.6%).

The distance between the FGZ and both ports explains the latter. In fact, on average, the FGZ are 100 km closer to Port 3 than to Port 1. This result shows that before equal level of service between two alternative ports, exporters would still prefer the closer one (in terms of distance). Therefore, it is worth to study the trade-off between level of service and distance. In terms of the modeling tool, how much larger should the value of the parameter γ_i be (for the port with lower share) in order to reach a similar market share between both ports.

Figure 9 shows the change in the captured demand by each port as a function of an increasing value of γ_i . It can be observed that when the value of γ_i approaches 1,250, the demand captured by Port 1 increases considerably. However, this increase starts after about 130 periods, reaching later a point in which it surpasses the capture of Port 3 (after 495 periods). The explanation of this dynamic behavior is that when the expression (1) is evaluated for ports 1 and 3, the difference in the value of the term $\beta(d_{i1}, d_{i2}, \dots, d_{im})$ between both ports reaches about the same value of the differences in their corresponding γ_i values, i.e. 1,250. The latter occurs because there are 43 FGZ that are located approximately 100 km closer to port 3 than to Port 1. This fact makes that the value of β (found to be -12 in the calibration) multiplied by the difference in distances amounts approximately 1,250.

Note however, that the timing of the stable demand capture is far from being immediate. The delay in reaching this point is due to the effect of the neighborhood in the choice function in expression (1), i.e. $\alpha\Phi_1^i$, expressing the effect of a spatio-temporal inertia in the choice process. Thus, if Port 1 would improve its level of service in such a way that $\gamma_1 \geq 1400$ it should reach the demand capture of Port 3 in 10 periods (less than three years).

The behavior describe above, is in accordance with what is observed in actual cases, because when two ports share the same hinterland, the farthest port must offer level of service conditions that are considerably better than its competitor to generate a significant change in the demand capture pattern. The lag taken by the improved port to reach the stable demand capture will depend on the advantages expected by the exporters in relation to the direct cost of traveling to a farther location with their cargo. Thus, the better the level of service improvement in one port, the faster the change would be in the port choice by the exporters and its neighbors.

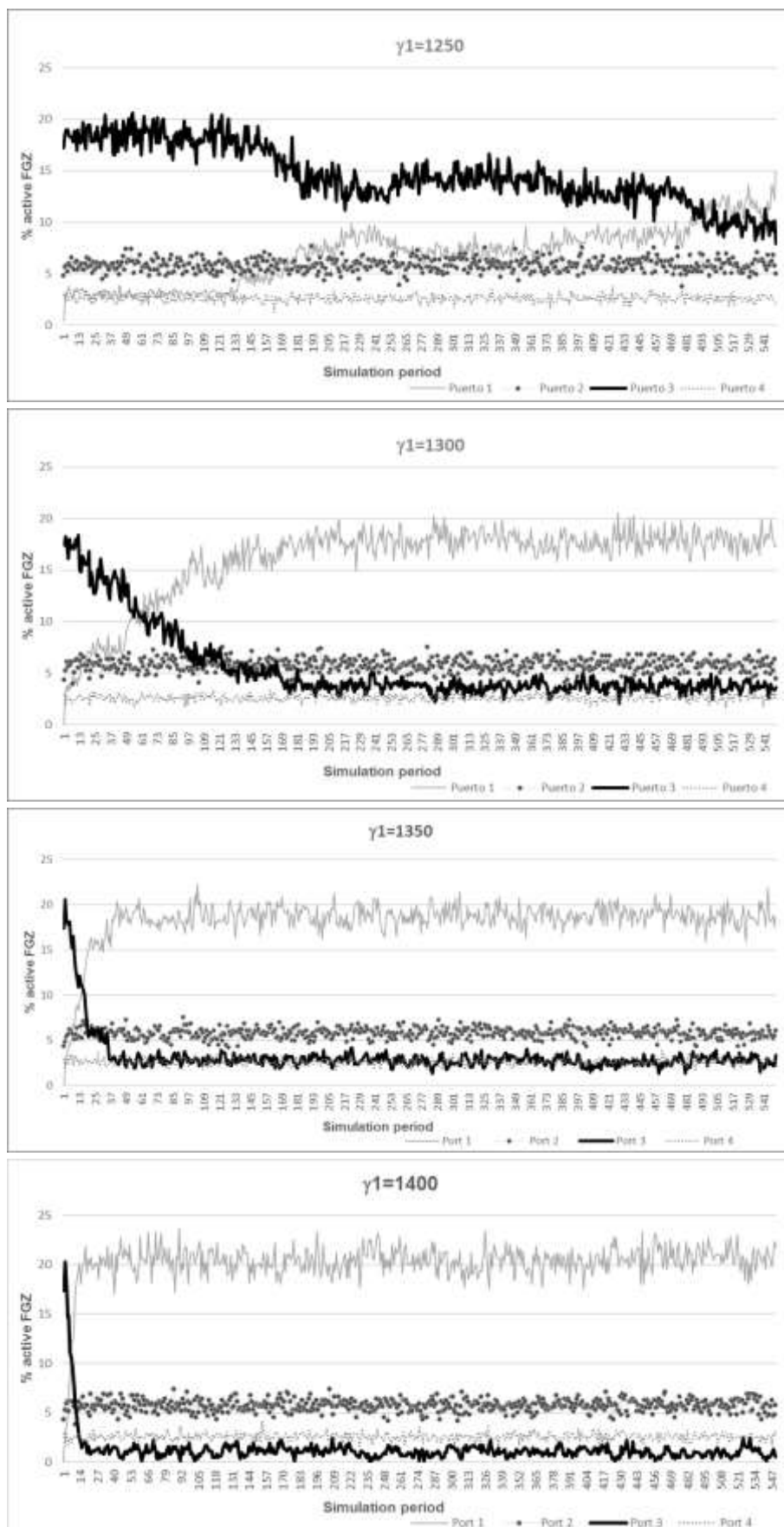


Fig. 9. Market share of each port with $\gamma_1 = 1250, 1300, 1350$ y 1400

5. CONCLUSIONS

In this article, we developed a new methodology to simulate the process of hinterland formation. In this case the hinterland is defined from the shippers' viewpoint. This methodology is formed on the hypothesis that for a given shipper, the process of port choice is dynamic and depends not only on the level of service of the ports and the attributes of the shipper, but also on the decision of other shippers within a certain spatial neighborhood.

These aspects of the choice process were represented in a mathematical expression that incorporates the decision of neighbors, the trip cost from the shipper's location to the port (proxied by the distance) and a free parameter that accounts for the attributes that make a port more desirable with respect to its competitors. In addition, this methodology assumes that the modeler does not have perfect information about the ports' characteristics nor the shippers' attributes and consequently it includes a stochastic component that accounts for the lack of information that may affect the choice process. This expression is then incorporated into a scheme of a probabilistic cellular automaton whose transition function is given by the expression described above.

The shippers are located within a certain region of interest, which is divided into several spatial units that are mutually exclusive and collectively exhaustive, called freight generation zones (FGZ). Thus, shippers located in a given FGZ can be either active or inactive. If there is transport between that FGZ to any port, then the FGZ is active, and it is inactive otherwise. Each FGZ can be active through any of the available ports; hence if n ports were available, a FGZ can be in $n+1$ possible states, where the state i ($i \neq 0$) indicates that exports are being sent from the FGZ to the i -th port and the case $i=0$ indicates inactivity.

The resulting modeling tool is called the Hinterland Definition Probabilistic Cellular Automaton (HDPCA) and is intended to simulate the port choice decision process made by perhaps several individuals acting together in order to select the port that best suits their particular case. This decision is the result of a complex system of interactions and previous decisions that influence the decision of other actors both in time and space. The HDPCA allows the modeler to analyze not only the choice that would take place at any given FGZ but also to study the evolution of the collective behavior that occurs at any time step as a dynamic spatio-temporal process.

The HDPCA framework was specified and calibrated using a genetic algorithm implementation. The calibrated model was applied to a real case for the automobile exports industry in Brazil.

The model specification was calibrated using a sample that contained data for automobile exports from 1997 to 2015 through the main ports of call, temporally divided into trimesters, generating a time series of 74 observations. The spatial data accounted for 69 FGZ in the region of analysis. The calibrated model reached a high degree of coherence between the predicted versus the actual choice in the sample (73% of successful prediction).

The results of the application of the HDPCA indicated that in the case of automobile exports from Brazil, the port choice decision is significantly influenced by the decision of other exporters in a neighborhood of 267 km of radius.

The results showed also that port choice was influenced by the distance from each FGZ to the selected port (which acts as a proxy of the trip cost) as well as the intrinsic characteristics of each port (proxied by a free parameter).

To gauge the robustness of the results given by the HDPCA, the process was simulated for 80 trimesters (in order to capture the long-run emergent behavior of the system), and within each trimester the model was ran 20 times. The results showed that the modeling tool is able to adequately represent the dynamics of the port choice process, replicating accurately not only the aggregate market share but also the captured demand by each port.

The presented methodology is especially adequate to simulate the individual and collective behavior about the demand capture within a contested hinterland. The methodology also allows the simulation of a system of competing ports when their attributes change, as well as the dynamic process followed by the shippers. The latter is possible due to the neighbor's effect which accounts both for spatial and temporal inertia.

REFERENCES

- [1] Chang, Y-T, Lee, S.Y. and Tongzon, J.L., Port selection factors by shipping lines: Different perspectives between trunk liners and feeder service providers, *Marine Policy* 32 (2008), 877–885.
- [2] Chopard, B. and Droz, M. *Cellular Automata Modeling of Physical Systems*, Cambridge University Press, Cambridge, England, 2005.
- [3] Gardner, M., *Mathematical Games: The fantastic combinations of John Conway's new solitaire game "life"*, *Scientific American* 223, 1970, 120-123.

- [4] Garrido, R. and Mahmassani, H., Forecasting freight transportation demand with the space–time multinomial probit model, *Transportation Research Part B: Methodological* Volume 34 (2000), 403–418.
- [5] Garrido, R.A. and Leva, M., Port of destination and carrier selection for fruit exports: a multi-dimensional space-time multinomial probit model, *Transportation Research Part B* 38 (2004), 657-667.
- [6] Goldberg, D., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Longman Publishing Co., Boston, MA, USA, 1989.
- [7] Goles, E. and Martínez, S., *Cellular Automata and Complex Systems*, Nonlinear Phenomena and Complex Systems Series, Kluwer, 1998.
- [8] Griffeath, D. and Moore, C., *New Constructions in Cellular Automata*, Oxford University Press, Oxford, 2003.
- [9] Halim, R. A., Kwakkel, J. H. and Tavasszy, L. A., A strategic model of port-hinterland freight distribution networks, *Transportation Research Part E: Logistics and Transportation Review* 95 (2016), 368-384.
- [10] Hausman, J. and Mcfadden, D., Specification Tests for the Multinomial Logit Model, *Econometrica* 52 (1984), 1219-1240.
- [11] Jiang, J., Lee, L.H., Chew, E.P. and Gan, C.C., Port Connectivity Study: An Analysis Framework from a Global Container Liner Shipping Network Perspective, *Transportation Research Part E* 73 (2015), 47-64.
- [12] Lee, S-Y, Chang, Y-T and Tae-Woo, P., Container port selection factors: heterogeneity among major market player, *Journal of International Logistics and Trade* 8 (2010), 73-90.
- [13] Lirn, T.C., Thanopoulou, H.A., Beynon, M.J. and Beresford, A.K.C., An Application of AHP on Transshipment Port Selection: A Global Perspective, *Maritime Economics & Logistics* 6 (2004), 70-91.
- [14] Magala, M. and Sammons, A., A New Approach to Port Choice Modelling. *Maritime Economics & Logistics* 10 (2008), 9–34.
- [15] Malchow, M. and Kanafani, A., A disaggregate analysis of factors influencing port selection, *Maritime policy & Management* 28 (2004), 265-277.

-
- [16] Manski, Ch. and Mcfadden, D., *Structural Analysis of Discrete Data and Econometric Applications*, Cambridge: The MIT Press, London, England, 1981.
- [17] Mitchell, M., Crutchfield, J.P., and Das, R., *Evolving cellular automata with genetic algorithms: A review of recent work*, In *Proceedings of the First International Conference on Evolutionary Computation and Its Applications (EvCA'96)*, Russian Academy of Sciences, 1996.
- [18] Mohring, H., *Optimization and Scale Economies in Urban Bus Transportation*, *American Economic Review* (1972), 591-604.
- [19] Nir, A.S., Lin, K. y Liang, G.S., *Port choice behaviour – from the perspective of the shipper*. *Maritime policy & Management* 30 (2003), 165-173.
- [20] Papola, A., *Some developments on the cross-nested logit model*, *Transportation Research Part B: Methodological* 38 (2004), 833–851
- [21] Song, D-P, Lyons, A., Li, D. and Sharifi, H., *Modeling port competition from a transport chain perspective*, *Transportation Research Part E: Logistics and Transportation Review* 87 (2016), 75–96.
- [22] Song, D-W and Yeo, K-T., *A competitive Analysis of Chinese Container Ports Using the Analytic Hierarchy Process*, *Maritime Economics & Logistics*.6, (2004), 34–52.
- [23] Steven, A.B. and Corsi, T.M., *Choosing a port: An analysis of containerized imports into the US*. *Transportation Research Part E: Logistics and transportation review* 48 (2012), 881- 895.
- [24] Tiwari, P., Itoh, H. and Doi, M., *Shippers' Port and Carrier Selection Behaviour in China: A Discrete Choice Analysis*, *Maritime Economics & Logistics* 5 (2004), 23–39.
- [25] Tongzon, J. and Sawant, L., *Port choice in a competitive environment: from the shipping lines' perspective*. *Applied Economics* 39 (2007), 477–92.
- [26] Tongzon, J., *Port choice and freight forwarders*. *Transportation Research Part E: Logistics and Transportation Review* 45 (2009), 186–195.
- [27] Vega, L., Cantillo, V., y Arellana, J., *Modelización de la elección de puerto desde una perspectiva desagregada: un análisis para los puertos colombianos*, XVII Congreso Panamericano de Ingeniería del tránsito, transporte y logística. Santander, 2014.

- [28] Veldman, S., Garcia-Alonso, L. and Vallejo, J., A port choice model with logit models: A case study for the Spanish container trade, *International Journal of Shipping and Transport Logistics* 5 (2013), 373-389.
- [29] Veldman, S., Garcia-Alonso, L. and Liu, M., Testing port choice models using physical and monetary data: a comparative case study for the Spanish container trades, *Maritime Policy & Management* 3 (2015), 495-508
- [30] Wang, X., Meng, Q., and Miaoc, L., Delimiting port hinterlands based on intermodal network flows: Model and algorithm, *Transportation Research Part E: Logistics and Transportation Review* 88 (2016), 32-51.
- [31] Weigend, GG. Some elements in the study of port geography, *Geographical Review* 48 (1958), 185-200.
- [32] Yeo, G-T, Ng, A.K.Y., Lee, P. T-W and Yang, Z., Modelling port choice in an uncertain environment, *Maritime Policy & Management* 41(2014), 251-267.

Anexo D: Port of destination and carrier selection for fruit exports: a multi-dimensional space-time multinomial probit model

ABSTRACT

This paper studies the selection of carrier and destination port for Chilean fruit exporters. This double selection is analyzed as a stochastic choice process with time and space interactions. A space-time error structure was specified within a multinomial probit model, considering serial correlation, spatial autocorrelation, and state dependence. The modeling approach was successfully applied to the case of grape exporters from Chile to the USA. The results validated the behavioral hypothesis for this process, i.e. there is significant state dependence, serial and spatial correlation in the choice of carrier and destination port, which should be considered when discrete choice models are used to predict export flows.

1.-INTRODUCTION

Fruit exports are an important growing sector among Chilean shippers. However, in spite of its importance, there is a lack of methodological analyses of the numerous decisions involved in this economic activity. This is partly due to the complexity of the process, the many agents and choices involved, and the lack of information and modeling tools.

US market is the main destination for fresh fruits Chilean exports; grapes being the most important among them. All these fruit exports are sold under the *free consignment sales* scheme, in which the consignee gets a fixed rate for selling the fruit, and the revenues (or losses) are entirely assumed by the exporter.

The choice of destination port is imbedded in a complicated business process, much broader than the pure transportation decisions scope. Hence it is not possible to isolate this choice from the rest of the exportation activities and the decisions of the various agents involved.

Among the multiple factors that affect the choice of destination port, aspects like accessibility, land use, infrastructure availability and other external features of the port surroundings suggest the existence of spatial interactions between the available alternatives to the shipper. In addition, the presence of business habits, long term investments and specialization issues suggest the existence of temporal effects in the choice process. As for the carrier choice, only temporal effects seem to be relevant due mainly to long term shipper-carrier relationships.

The rest of this paper is organized as follows; in section two econometric derivation of the model is presented; section three presents results from the application of the spatial-time MNP model to the joint choice of destination port and carrier for Chilean grape exports to the USA. Finally, section four presents the main conclusions.

2.-ECONOMETRIC MODEL DERIVATION

Including temporal and spatial interactions within a discrete choice model is a task that can only be performed over a flexible and general model. This characteristic precludes the use of the most common (and simple) discrete choice models available in commercial econometric software: multinomial logit and hierarchical logit. The choice of modeling structure was an advanced multinomial probit. The only previous study that included both spatial and temporal effects in a probit structure is (Garrido, 1997). This paper enhances the capabilities of that modeling structure.

Let's consider a choice situation involving P ports, N carriers and T time periods. The spatial-temporal choice set is represented by the vector $C=\{pnt\}$, with $P \times N \times T$ elements. The attributes of every C element depend on the port specific characteristics, carrier specific characteristics, and the joint port-carrier pair characteristics, as well as the time interval in which the selection of a given port-carrier combination is made.

According to the random utility theory, a pnt alternative is selected if its utility is the highest among all available alternatives. The utility structure is defined as follows:

$$U(p, n, t) = V(p, n, t) + \delta(p, n, t) \quad (2.1)$$

where $V(p, n, t)$ is the representative (deterministic) utility and $\delta(p, n, t)$ is the error term (stochastic) for the pnt alternative.

The probability of choosing the port p and carrier n during the period t is given by:

$$\begin{aligned} P(p, n, t) &= P\{U(p, n, t) \geq U(q, m, t), \forall qm \in C\} \\ P(p, n, t) &= P\{V(p, n, t) + \delta(p, n, t) \geq V(q, m, t) + \delta(q, m, t), \forall qm \in C\} \end{aligned} \quad (2.2)$$

To define $V(p, n, t)$ and $\delta(p, n, t)$, we have used the multidimensional model notation proposed by Williams and Ortúzar (1982):

$$V(p, n, t) = V_{pt} + V_{nt} + V_{pnt} \quad (2.3)$$

$$\delta(p, n, t) = \varepsilon_{pt} + \varphi_{nt} + \eta_{pnt} \quad (2.4)$$

where V_{pt} , V_{nt} , V_{pnt} are the specific representative utilities for the port (p), carrier (n) and port-carrier pair (pn) in period t , and ε_{pt} , φ_{nt} , η_{pnt} are the stochastic errors for each of them in the same period. Errors are considered to be independent of each other.

Spatial and temporal effects are included within a general error structure considering: (a) ε_{pt} , the error component associated to a port in a given time interval, is the result of a dynamic stochastic process represented by an AR(1) (first order autoregressive process) and it is spatially correlated to the other ports error terms; (b) φ_{nt} and η_{pnt} , the error components associated to a given carrier and a given port-carrier pair respectively, are the result of stochastic independent processes represented by an AR(1).

We incorporated spatial and serial correlation in the ε_{pt} expression, through the error structure of the spatial-temporal probit model developed by Garrido (1997). Accordingly: $\varepsilon_t = \rho W \varepsilon_t + \mu_t$, where ρ is a spatial interaction parameter, W is the $P \times P$ dimension contiguity matrix with elements given by:

$$w_{ij} = \begin{cases} 1, & \text{if } i \text{ and } j \text{ are neighbor ports} \\ 0, & \text{otherwise} \end{cases} \quad (2.5)$$

and μ_t represents an AR(1) process. The error term may be expressed as:

$$\varepsilon_t = (I - \rho W)^{-1} \mu_t \quad (2.6)$$

where I is a $P \times P$ identity matrix. Expression (2-6) is defined only if the expression $(I - \rho W)^{-1}$ is not singular.

The μ_t stochastic process, is represented by an AR(1) model:

$$\mu_t = \alpha \mu_{t-1} + \lambda_t \quad (2.7)$$

where α is a scalar parameter and λ_t is an independently distributed (ID) random vector, such that $\lambda_t \sim N(0, \Omega)$. This vector represents a white noise process and its components have the following moments:

$$\begin{aligned} E(\lambda_{pt}) &= 0 \\ E(\lambda_{pt} \lambda_{qs}) &= \begin{cases} \omega_p^2 & \text{if } p = q, t = s \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (2.8)$$

$$\text{i.e. } \Omega = \begin{bmatrix} \omega_1^2 & 0 & \cdots & 0 \\ 0 & \omega_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \omega_p^2 \end{bmatrix}.$$

Considering the time lag operator L , as $L\mu_t = \mu_{t-1}$, μ_t may be expressed from (2-7) as follows:

$$\mu_t = C(L)\lambda_t, \quad C(L) = \left(\frac{1}{1-\alpha L} \right) \quad (2.9)$$

Expanding $C(L)$ through an infinite series (considering $|\alpha| < 1$), the following expression is obtained:

$$\begin{aligned} C(L) &= \left(\frac{1}{1-\alpha L} \right) = 1 + \alpha L + \alpha^2 L^2 + \cdots + \alpha^i L^i + \cdots \\ C(L) &= \sum_{j=0}^{\infty} \alpha^j L^j \end{aligned} \quad (2.10)$$

Therefore μ_t can be written as follows:

$$\mu_t = \sum_{j=0}^{\infty} \alpha^j \lambda_{t-j} \quad (2.11)$$

Replacing (2-11) into (2-6) we obtained:

$$\varepsilon_t = (I - \rho W)^{-1} \sum_{j=0}^{\infty} \alpha^j \lambda_{t-j} \quad (2.12)$$

The temporal characteristics of the carrier error component (φ_{nt}) and the one for the port-carrier pair (η_{pnt}), are written in the form of an AR(1) process.

Thus, the specific error associated to the carrier becomes:

$$\varphi_t = \theta \varphi_{t-1} + \xi_t \quad (2.13)$$

where θ is a scalar parameter and ξ_t is an ID random vector, such that $\xi_t \sim N(0, \Sigma)$. This vector represents a white noise process, analogous to that of λ_t :

$$\begin{aligned}
E(\xi_{nt}) &= 0 \\
E(\xi_{nt} \xi_{ms}) &= \begin{cases} \sigma_n^2 & \text{if } n = m, t = s \\ 0 & \text{otherwise} \end{cases}
\end{aligned} \tag{2.14}$$

$$\text{i.e., } \Sigma = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_N^2 \end{bmatrix}.$$

The case of φ_t is also analogous to that of λ_t , hence it can be expressed as follows:

$$\varphi_t = \sum_{j=0}^{\infty} \theta^j \xi_{t-j} \tag{2.15}$$

Similarly, the specific error component for the port-carrier pair was also assumed to follow an AR(1), such that:

$$\eta_t = \phi \eta_{t-1} + \chi_t \tag{2.16}$$

where ϕ is a scalar parameter and χ_t is an ID random vector of white noise, such that $\chi_t \sim N(0, \Phi)$. The moments are analogous to λ_t as well:

$$\begin{aligned}
E(\chi_{pnt}) &= 0 \\
E(\chi_{pnt} \chi_{qms}) &= \begin{cases} \kappa_{pn}^2 & \text{if } p = q, n = m, t = s \\ 0 & \text{otherwise} \end{cases}
\end{aligned} \tag{2.17}$$

$$\text{i.e., } \Phi = \begin{bmatrix} \kappa_{11}^2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \kappa_{12}^2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & 0 & \vdots & \vdots \\ \vdots & \vdots & 0 & \kappa_{np}^2 & 0 & 0 \\ 0 & 0 & \vdots & 0 & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \kappa_{NP}^2 \end{bmatrix}.$$

The case of the specific error component of the port-carrier pair is totally analogous to the previous two:

$$\eta_t = \sum_{j=0}^{\infty} \phi^j \chi_{t-j} \tag{2.18}$$

Thus, having ε_t , φ_t , η_t , the error term associated to choosing port p and carrier n during the period t , is expressed as follows:

$$\delta(t) = (I - \rho W)^{-1} \sum_{j=0}^{\infty} \alpha^j \lambda_{t-j} + \sum_{j=0}^{\infty} \theta^j \xi_{t-j} + \sum_{j=0}^{\infty} \phi^j \chi_{t-j} \quad (2.19)$$

Statistical properties of $\delta(t)$ depend on the density functions of the white noise vectors λ_t , ξ_t y χ_t . In this case, we assumed normality, therefore only the first and second moments of $\delta(t)$ have to be estimated to fully describe the probability density function.

The expected value of $\delta(t)$ is given by:

$$\begin{aligned} E(\delta(t)) &= E \left((I - \rho W)^{-1} \sum_{j=0}^{\infty} \alpha^j \lambda_{t-j} + \sum_{j=0}^{\infty} \theta^j \xi_{t-j} + \sum_{j=0}^{\infty} \phi^j \chi_{t-j} \right) \\ &= (I - \rho W)^{-1} \sum_{j=0}^{\infty} \alpha^j E(\lambda_{t-j}) + \sum_{j=0}^{\infty} \theta^j E(\xi_{t-j}) + \sum_{j=0}^{\infty} \phi^j E(\chi_{t-j}) \\ E(\delta(t)) &= 0 \end{aligned} \quad (2.20)$$

The independence between the error terms ε_t , φ_t y η_t implies that the covariance matrix for $\delta(t)$ is the summation of their corresponding covariance matrices.

The covariance matrix of ε_t for two different intervals is given by:

$$\begin{aligned} E(\varepsilon_t \varepsilon'_{t-r}) &= E \left[\left((I - \rho W)^{-1} \sum_{j=0}^{\infty} \alpha^j \lambda_{t-j} \right) \left((I - \rho W)^{-1} \sum_{j=0}^{\infty} \alpha^j \lambda_{t-j-r} \right)' \right] \\ &= (I - \rho W)^{-1} E \left[\sum_{j=0}^{\infty} \alpha^j \lambda_{t-j} \sum_{j=0}^{\infty} \alpha^j \lambda'_{t-j-r} \right] (I - \rho W)^{-1'} \end{aligned} \quad (2.21)$$

where r is the time lag between both intervals.

Considering that only the product of contemporary elements takes values different from zero, expression (2-21) can be written as follows:

$$\begin{aligned} E(\varepsilon_t \varepsilon'_{t-r}) &= (I - \rho W)^{-1} \sum_{j=0}^{\infty} \alpha^{j+r} \alpha^j E[\lambda^2_{t-j-r}] (I - \rho W)^{-1'} \\ &= \sum_{j=0}^{\infty} \alpha^{j+r} \alpha^j (I - \rho W)^{-1} \Omega (I - \rho W)^{-1'} \end{aligned} \quad (2.22)$$

where the first term, $\sum_{j=0}^{\infty} \alpha^{j+r} \alpha^j$ is equivalent to:

$$\sum_{j=0}^{\infty} \alpha^{j+r} \alpha^j = \alpha^r \sum_{j=0}^{\infty} \alpha^{2j} = \frac{\alpha^r}{1-\alpha^2} \quad (2.23)$$

Replacing (2-23) into (2-22) we obtain:

$$E(\varepsilon_t \varepsilon'_{t-r}) = \frac{\alpha^r}{1-\alpha^2} (I - \rho W)^{-1} \Omega (I - \rho W)^{-1'} \quad (2.24)$$

which is the covariance matrix of the specific error associated to the port. $P+2$ of its parameters are unknown.

By following a similar procedure, the covariance matrix of φ_t between two different intervals is given by:

$$\begin{aligned} E(\varphi_t \varphi'_{t-r}) &= E \left[\left(\sum_{j=0}^{\infty} \theta^j \xi_{t-j} \right) \left(\sum_{j=0}^{\infty} \theta^j \xi_{t-j-r} \right)' \right] \\ &= \sum_{j=0}^{\infty} \theta^{j+r} \theta^j E[\xi_{t-j-r}^2] \\ &= \sum_{j=0}^{\infty} \theta^{j+r} \theta^j \Sigma \end{aligned} \quad (2.25)$$

Expression (2-25) can be rewritten as follows:

$$E(\varphi_t \varphi'_{t-r}) = \frac{\theta^r}{1-\theta^2} \Sigma \quad (2.26)$$

which corresponds to the covariance matrix of the specific error associated to the carrier, with $N + 1$ unknown parameters.

Finally, the covariance matrix of the specific error associated to the port – carrier pair is obtained in analogous form:

$$E(\eta_t \eta'_{t-r}) = \frac{\phi^r}{1-\phi^2} \Phi \quad (2.27)$$

which has $P \times N + 1$ unknowns.

Once the covariance matrices of ε_t , φ_t and η_t , are known, the second moment of $\delta(t)$ is determined by:

$$\begin{aligned} E(\delta_t \delta'_{t-r}) &= E(\varepsilon_t \varepsilon'_{t-r}) + E(\varphi_t \varphi'_{t-r}) + E(\eta_t \eta'_{t-r}) \\ &= \frac{\alpha^r}{1-\alpha^2} (I - \rho W)^{-1} \Omega (I - \rho W)^{-1'} + \frac{\theta^r}{1-\theta^2} \Sigma + \frac{\phi^r}{1-\phi^2} \Phi \end{aligned} \quad (2.28)$$

This is a $P \times N$ dimensional matrix with $(P + 2) + (N + 1) + (P \times N + 1)$ unknown parameters.

Only $\left\lceil \frac{(N \times P \times T - 1) N \times P \times T}{2} \right\rceil - 1$ of them are identifiable (Bunch, 1991). Besides, it is

important to notice that α , θ , ϕ and ρ are constrained to be lower than 1.0 in order to meet the stationarity conditions of the temporal parameters and stability of the spatial process.

The specification of the representative utility function (the deterministic portion in expression 2-3) includes attributes for each element: port, carrier, and the combination port-carrier. For V_{pt} and V_{nt} the habits or inertial behavior, are represented by the inclusion of the choice made in the previous period (see Train, 2002).

Therefore,

$$V_{pt} = \sum_{k=1}^K \beta_k X_{kpt} + \beta_{K+1} Y_{p(t-1)} \quad (2.29)$$

where β is a $K+1$ dimension vector; its K first elements are the attributes of the port p (X_{kpt}) and the $K+1$ element corresponds to the choice in the previous period, defined as:

$$Y_{p(t-1)} = \begin{cases} 1 & \text{if port } p \text{ was chosen in } (t-1) \\ 0 & \text{otherwise} \end{cases} \quad (2.30)$$

The sign of the β_{K+1} parameter indicates the satisfaction with the port selected in the previous period.

The representative utility associated to the carrier has a similar expression to that of V_{pt} :

$$V_{nt} = \sum_{h=1}^H \gamma_h X_{hnt} + \gamma_{H+1} Z_{n(t-1)} \quad (2.31)$$

where γ is a $H+1$ dimension vector. Its H first elements are the attributes of the carrier n (X_{hnt}) and the $H+1$ element corresponds to the choice made in the previous period, defined by:

$$Z_{n(t-1)} = \begin{cases} 1 & \text{if carrier } n \text{ was chosen in } (t-1) \\ 0 & \text{otherwise} \end{cases} \quad (2.32)$$

Similarly, the sign of the γ_{H+1} , parameter indicates the satisfaction with the carrier chosen in the previous period.

Finally, the representative utility associated to the port-carrier pair is simply an additive and linear in the parameters function, defined as:

$$V_{pnt} = \sum_{g=1}^G v_g W_{gpnt} \quad (2.33)$$

where, v is a G dimensional vector of attribute parameters associated to the port-carrier pair pn .

Thus, the representative utility $V(p,n,t)$ is formed by:

$$\begin{aligned} V(p,n,t) &= V_{pt} + V_{nt} + V_{pnt} \\ &= \sum_{k=1}^K \beta_k X_{kpt} + \beta_{K+1} Y_{p(t-1)} + \sum_{h=1}^H \gamma_h V_{hpt} + \gamma_{H+1} Z_{n(t-1)} + \sum_{g=1}^G v_g W_{gpnt} \end{aligned} \quad (2.34)$$

This expression, along with (2-19) define the utility function of choosing the port p and carrier n during period t , within the spatial-temporal MNP model (STMNP).

The parameters of the STMNP model are estimated through a simulated maximum likelihood (SML) method. The GHK (Bolduc, 1999) probability simulator was used for this purpose.

The likelihood function for the STMNP model can be expressed as follows:

$$L(X, \theta) = \prod_{e=1}^E \prod_{p=1}^P \prod_{n=1}^N \prod_{t=1}^T P_{epnt}(X_{epnt}, \theta)^{\tau_{epnt}} \quad (2.35)$$

where X represents the attributes matrix, θ represents the unknown parameters vector, E is the number of exporters in the sample, P is the number of ports, N is the number of

carriers and T is the number of time periods to be considered. τ_{epnt} takes the value 1 if the e -th exporter chose the port p and carrier n to in the period t , and it takes the value 0 otherwise.

Due to the nonlinear characteristics of this likelihood function, several local optima may exist. Therefore, the algorithm may converge to different solutions depending on the starting points of the optimization algorithm. Then, it is possible to find a better likelihood value when a constraint is added to the specification (basically when the algorithm finds a better local optimum due to the new constraint). The latter never occurs with simpler models like the MNL with a unique optimum.

The GHK simulator was programmed in a C++ code, along with the BFGS quasi-Newton search method. This code also incorporates a low-discrepancy random number generator to simulate choice probabilities (Leva *et al*, 2002).

3.-APPLICATION

The model described in section 2 was calibrated using actual data about the choice of destination port and ocean carrier for Chilean grape exports to the USA. The calibration database contains information about all grape shipments between 1997 and 2001. The choice set is formed by the ports of Philadelphia (Pennsylvania) and Wilmington (Delaware), and the main carriers: CSAV (also known as Chilean Line) and Lauritzen Reefers. A third alternative was also considered as “other carriers”, representing other minor carriers with common features among them. The export peak season for grapes is from February to April. Within it, three periods were defined, as the first week of every month. The models were specified as follows.

First model

The choice set is formed by the ports of Wilmington and Philadelphia, and the carriers CSAV and Lauritzen. Considering the proximity between both ports, the contiguity matrix was defined as:

$$W = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad (3.1)$$

Covariance matrices Ω , Σ y Φ were assumed to be homoskedastic.

Since there are 12 alternatives (2 ports x 2 carriers x 3 periods), the covariance matrix of the process has the following structure:

$$\begin{bmatrix} \frac{1}{(1-\alpha^2)}\Theta + \frac{1}{(1-\theta^2)}\Sigma + \frac{1}{(1-\phi^2)}\Phi & \frac{\alpha}{(1-\alpha^2)}\Theta + \frac{\theta}{(1-\theta^2)}\Sigma + \frac{\phi}{(1-\phi^2)}\Phi & \frac{\alpha^2}{(1-\alpha^2)}\Theta + \frac{\theta^2}{(1-\theta^2)}\Sigma + \frac{\phi^2}{(1-\phi^2)}\Phi \\ \frac{\alpha}{(1-\alpha^2)}\Theta + \frac{\theta}{(1-\theta^2)}\Sigma + \frac{\phi}{(1-\phi^2)}\Phi & \frac{1}{(1-\alpha^2)}\Theta + \frac{1}{(1-\theta^2)}\Sigma + \frac{1}{(1-\phi^2)}\Phi & \frac{\alpha}{(1-\alpha^2)}\Theta + \frac{\theta}{(1-\theta^2)}\Sigma + \frac{\phi}{(1-\phi^2)}\Phi \\ \frac{\alpha^2}{(1-\alpha^2)}\Theta + \frac{\theta^2}{(1-\theta^2)}\Sigma + \frac{\phi^2}{(1-\phi^2)}\Phi & \frac{\alpha}{(1-\alpha^2)}\Theta + \frac{\theta}{(1-\theta^2)}\Sigma + \frac{\phi}{(1-\phi^2)}\Phi & \frac{1}{(1-\alpha^2)}\Theta + \frac{1}{(1-\theta^2)}\Sigma + \frac{1}{(1-\phi^2)}\Phi \end{bmatrix}$$

where Θ corresponds to $(I - \rho W)^{-1} \Omega (I - \rho W)^{-1}$, in expression (2.28).

The parameters to be estimated from this matrix are α , θ , ϕ and ρ (all constrained to be lower than one) and ω , σ y κ , the standard deviations of the specific matrices associated to the choice of port, carrier and port-carrier pair, respectively.

An important feature of this structure is that spatial and temporal effects depend on their respective covariance matrices. Thus, if the standard deviation parameter of any of these matrices were not significant and hence eliminated, the parameter of the corresponding effect would be directly eliminated along.

The attributes of the representative utility function are: choice of port in the previous period ($Y_{p(t-1)}$); choice of carrier in the previous period ($Z_{n(t-1)}$); and unit transportation cost expressed by the freight rate (C). The specified representative utility function is given by:

$$\begin{aligned} V(p, n, t) &= V_{pt} + V_{nt} + V_{pnt} \\ &= \beta_1 Y_{p(t-1)} + \beta_2 Z_{n(t-1)} + \beta_3 C \end{aligned} \quad (3.1)$$

Considering the 7 parameters in the covariance matrix and the 3 parameters in the utility function, there are 10 parameters to be estimated. For this model, 616 observations were available.

Second model

This model considers the same alternatives as the first model, but adding a third carrier, called “other carriers” grouping smaller size carriers with lower frequencies than the previous two. The contiguity matrix is the same as in the first model. Similarly, the covariance matrices, Ω , Σ and Φ were considered to be homoscedastic.

The structure of the covariance matrix is similar to that in the first model, but larger, as

there are 18 alternatives instead of 12. The representative utility function is also similar. 830 observations were used to calibrate 10 parameters.

Calibration results are presented in Tables 1 and 2, The most significant parameters estimated are those associated to the previous selections of port (β_1) and carrier (β_2), due to their magnitude and significance, demonstrating the existence of inertia in the choice behavior. The positive sign of these parameters indicates a high degree of satisfaction with the previous decisions. The temporal parameters associated to the covariance matrices of the port (α), carrier (θ) and port-carrier pair (ϕ) demonstrates the presence of considerably high serial correlation, ranging from 0.600 to 0.711 in the first model, and from 0.582 to 0.684 in the second model. In both cases, all of them were positive and highly significant.

The value of the spatial parameter (ρ) was 0.512 for the first model and 0.394 for the second model, being more significant in the first model. Its positive value may be explained for the proximity and the similitude between them that makes any external spatial factor to have similar effects on both (for instance, access to the ground transport network in the hinterland).

The remaining parameters, except for that associated to freight rate, were all very significant in the first model. The second model presented high significance for all the parameter except for the freight rate and the standard deviation of the port and carrier selection. Even though the freight rate parameter was not significant, it was kept in the specification because it served to discriminate availability for each observation (not all the shipments had available all the options). The insignificance of this parameter was expected in advance by the modelers. In fact, various decision makers in this industry were interviewed before collecting the data, and they mentioned this attribute as one of the less relevant in their decisions.

The inclusion of the third carrier ("other carriers"), had only a minor effect on the relevant parameters. It lowered the magnitude and significance of the standard deviation of the port and carrier. However, no direct comparison may be made as both models have different sample sizes and choice sets.

4.- CONCLUSIONS

We conclude that the choice of destination port and carrier in the context on Chilean grape exports to the USA is highly influenced by temporal effects, both at the level of deterministic inertia and as a stochastic effect. The latter appeared in the error terms as a first order autoregressive stochastic process, and the former within the representative utility function as state dependence. On the other hand, the spatial effect had also a significant impact in the decision process, and it was represented as spatial autocorrelation.

It is also worth noting the development of a new spatial–temporal structure for MNP models, as a methodological contribution to discrete choice modeling, since it may be used for modeling many choice processes having spatial and/or temporal effects.

REFERENCES

BOLDUC, D. (1999). A practical technique to estimate multinomial probit models in transportation. **Transportation Research, Vol. 33B**, 63-79.

BOLDUC, D., FORTIN, B. and FOURNIER, M. (1996). The impact of incentive policies on the practice location of doctors: a multinomial probit analysis. **Journal of Labor Economics, Vol. 14**, 703 – 732.

BUNCH, D. (1991). Estimability in the multinomial probit model. **Transportation Research, Vol. 25B**, 1-12.

GARRIDO, R.A. (1997). **Analysis of Spatial and Temporal Characteristics of Freight Demand**. Ph. D. Dissertation, Department of Civil Engineering, University of Texas at Austin.

LEVA, M. (2002). **Caracterización de la Toma de Decisiones de Exportadores Frutícolas: Un Enfoque Espacio-Temporal**. Tesis de Magíster, Departamento de Ingeniería de Transporte, Pontificia Universidad Católica de Chile. **(In Spanish)**

LEVA, M., SILVA, M. and GARRIDO, R.A. (2002). Código para estimación de modelos probit multinomiales con simulador GHK y métodos MC y LDMC. Documento de Trabajo Nº 84, Departamento de Ingeniería de Transporte, Pontificia Universidad Católica de Chile.

McMILLEN, D. (1992). Probit with spatial autocorrelation. **Journal of Regional Science**, Vol. 32, 335-348.

SILVA, M. (2002). **Secuencias de Baja Discrepancia para Estimación de Modelos Logit Mixto**. Tesis de Magíster, Departamento de Ingeniería de Transporte, Pontificia Universidad Católica de Chile.

TRAIN, K. (2002). **Discrete Choice Methods with Simulation**. Cambridge University Press, Cambridge.

WILLIAMS, H.C.W.L. and ORTUZAR, J. de D. (1982). Behavioural theories of dispersion and the misspecification of travel demand models. **Transportation Research**, Vol. 16 B, 167-219.

Table 1: First model

	Parameter	Value	t- statistic	t- statistic (difference from 1)
Previous period choice of port	β_1	4.945	5.044	-
Previous period choice of carrier	β_2	3.946	6.884	-
Transportation cost	β_3	-0.0003	0.042	-
Port temporal parameter	α	0.711	5.065	2.056
Spatial parameter	ρ	0.512	4.231	4.032
Port standard deviation	ω	0.184	2.538	-
Carrier temporal parameter	θ	0.706	5.016	2.086
Carrier standard deviation	σ	0.145	2.269	
Port – carrier temporal parameter	ϕ	0.600	5.336	3.548
Port – carrier standard deviation	κ	1.500	4.395	-
Log- likelihood	$L(\theta)$	-212.899	-	-
Average log-likelihood	$L(\theta)/N$	-0.345	-	-

Table 2: Second model

	Parameter	Value	t-statistic	t - statistic (difference from 1)
Previous period choice of port	β_1	4.751	5.296	-
Previous period choice of carrier	β_2	3.921	4.377	-
Transportation cost	β_3	-0.0002	-0.019	-
Port temporal parameter	α	0.684	4.761	2.196
Spatial parameter	ρ	0.394	3.615	5.560
Port standard deviation	ω	-0.074	-1.579	-
Carrier temporal parameter	θ	0.582	4.405	3.159
Carrier standard deviation	σ	-0.045	-1.226	-
Port – carrier temporal parameter	ϕ	0.607	4.889	3.156
Port – carrier standard deviation	κ	1.547	3.621	-
Log- likelihood	$L(\theta)$	-277.906	-	-
Average log- likelihood	$L(\theta)/N$	-0.334	-	-

Anexo E: Código ACPEP

```
unit Unit1;
```

```
{ $mode objfpc } { $H+ }
```

```
interface
```

```
uses
```

```
Classes. SysUtils. FileUtil. TAgGraph. TAsSeries. Forms. Controls. Graphics.  
Dialogs. ExtCtrls. StdCtrls;
```

```
type
```

```
{ TForm1 }
```

```
TForm1 = class(TForm)
```

```
    Button1: TButton;
```

```
    Button2: TButton;
```

```
    Button3: TButton;
```

```
    Button4: TButton;
```

```
    Button5: TButton;
```

```
    Button6: TButton;
```

```
    Button7: TButton;
```

```
    Button8: TButton;
```

```
    Button9: TButton;
```

```
    Chart1: TChart;
```

```
    Chart1LineSeries1: TLineSeries;
```

```
    Chart1LineSeries2: TLineSeries;
```

```
    Chart1LineSeries3: TLineSeries;
```

```
    Chart1LineSeries4: TLineSeries;
```

```
    Chart2: TChart;
```

```
    Chart2LineSeries1: TLineSeries;
```

```
    Chart2LineSeries2: TLineSeries;
```

```
    Chart2LineSeries3: TLineSeries;
```

```
    Chart2LineSeries4: TLineSeries;
```

```
    Chart3: TChart;
```

```
    Chart3LineSeries1: TLineSeries;
```

```
    Chart3LineSeries2: TLineSeries;
```

```
    Chart3LineSeries3: TLineSeries;
```

```
    Chart3LineSeries4: TLineSeries;
```

```
    Edit1: TEdit;
```

```
    Edit2: TEdit;
```

```
    Edit3: TEdit;
```

```
    Edit4: TEdit;
```

```
    Edit5: TEdit;
```

```
    Edit6: TEdit;
```

Image1: TImage;
Image2: TImage;
Image3: TImage;
Label1: TLabel;
Label10: TLabel;
Label11: TLabel;
Label12: TLabel;
Label13: TLabel;
Label14: TLabel;
Label15: TLabel;
Label16: TLabel;
Label17: TLabel;
Label18: TLabel;
Label19: TLabel;
Label2: TLabel;
Label20: TLabel;
Label21: TLabel;
Label22: TLabel;
Label23: TLabel;
Label24: TLabel;
Label3: TLabel;
Label4: TLabel;
Label5: TLabel;
Label6: TLabel;
Label7: TLabel;
Label8: TLabel;
Label9: TLabel;
Panel1: TPanel;
Panel2: TPanel;
Panel3: TPanel;
Panel4: TPanel;
ScrollBar1: TScrollBar;
ScrollBar2: TScrollBar;
ScrollBar3: TScrollBar;
ScrollBar4: TScrollBar;
ScrollBar5: TScrollBar;
ScrollBar6: TScrollBar;
ScrollBar7: TScrollBar;
ScrollBar8: TScrollBar;
Shape1: TShape;
Shape2: TShape;
Timer1: TTimer;
Timer2: TTimer;
ToggleBox2: TToggleBox;
ToggleBox3: TToggleBox;
ToggleBox4: TToggleBox;
ToggleBox1: TToggleBox;

```

procedure Button1Click(Sender: TObject);
procedure Button2Click(Sender: TObject);
procedure Button3Click(Sender: TObject);
procedure Button4Click(Sender: TObject);
procedure Button5Click(Sender: TObject);
procedure Button6Click(Sender: TObject);
procedure Button7Click(Sender: TObject);
procedure Button8Click(Sender: TObject);
procedure Button9Click(Sender: TObject);
procedure Image2Click(Sender: TObject);
procedure ScrollBar1Change(Sender: TObject);
procedure ScrollBar2Change(Sender: TObject);
procedure ScrollBar3Change(Sender: TObject);
procedure ScrollBar4Change(Sender: TObject);
procedure ScrollBar5Change(Sender: TObject);
procedure ScrollBar6Change(Sender: TObject);
procedure ScrollBar7Change(Sender: TObject);
procedure ScrollBar8Change(Sender: TObject);
procedure ToggleBox1Change(Sender: TObject);
procedure ToggleBox2Change(Sender: TObject);
procedure ToggleBox3Change(Sender: TObject);
procedure ToggleBox4Change(Sender: TObject);
private
  { private declarations }
public
  { public declarations }
end;

var
  Form1: TForm1;
  type puntos=record
    x:integer;
    y:integer;
    umb:real;
    estado: array [1..4] of 0..1;
    d: array [1..69] of real;//distancia a los vecinos
    dp: array [1..4] of real;//distancia a los puertos
    np: array [1..4] of integer;
    ep: array [1..4] of real;//cuantifica la calidad del puerto
    npf: array [1..4] of real;
    nscore: array [1..4] of integer;
    fi1: array [1..4] of real;
    score: array [1..4] of real;
  end;
  type individuos=record
    radio:real;
    alfa:real;

```

```
beta:real;
p1:real;
p2:real;
p3:real;
p4:real;
z:array [1..69] of real;
end;

type puntos1=record
x:integer;
y:integer;
z:real;
k:integer;
end;
  type puerto=record
    p:real;
    t:real;
  end;
type puntos2=record
d:array[1..69] of puntos1;
end;
  type vectorreal=record
    c:array[1..69] of integer;
    x:array[1..69] of integer;
    y:array[1..69] of integer;
    fun:array[1..69] of real;
  end;
type qalidad=record
q:real;
ij:integer;
end;
type tipo2=record
d:real;
k:integer;
end;
var
taco:array[1..100] of integer;
cony:array[1..100] of tipo2;
bolsa:array[1..76] of integer;
ind.ind2:array[1..100] of individuos;
cal.cal2:array[1..100] of qalidad;
t1.t2:array[1..76] of vectorreal;
port:array[1..4] of puerto;
zona: array[1..76] of puntos;
zona2: array[1..76] of puntos2;
ene.zz.tiempo:integer;
```

p1x.p1y.p2x.p2y.p3x.p3y.p4x.p4y.tiempo2.nhijos.nmutantes.naciertos.ntotal.ValorMaximo.semilla:integer;

ese.radio.alfa.beta.umbral.nt1.nt2.nt3.nt4.nt1s.nt2s.nt3s.nt4s.SumaTotal.SumaTotal2.nac:real;

 x1.y1.x2.y2.f1.f2.f1p.f2p.m.mp.PorcentajeAciertos.p:real;

 error1.error2.error3.error4.DatoError:real;

 p1ac.p2ac.p3ac.p4ac:0..1;

 entrada.entrada2.entrada3.entrada4.salida1.salida2.salida3.salida4.salida5:text;

implementation

{ \$R *.lfm }

{ TForm1 }

function xf(x:real): integer;

begin

 xf:=round(m*(x-x1)+f1);

end;

function yf(y:real): integer;

begin

 yf:=round(mp*(y-y1)+f1p);

end;

function lot(x:integer):integer;

begin

 lot:=-x+random(2*x);

end;

procedure ValoresAzarIndividuos;

var i.j.k.k1.k2:integer;

begin

 k2:=1;

 if semilla=1 then

 begin

 assignfile(entrada3.'ParametrosGlobales.dat');

 reset(entrada3);

 assignfile(entrada4.'UmbralZonas.dat');

 reset(entrada4);

 k2:=2;

readln(entrada3.ind[1].radio.ind[1].alfa.ind[1].beta.ind[1].p1.ind[1].p2.ind[1].p3.ind[1].p4);

 for k:=1 to 69 do

 begin

 readln(entrada4.k1.ind[1].z[k]);

 end;

```
closefile(entrada3);
closefile(entrada4);
end;
for i:=k2 to 100 do
begin
ind[i].radio:=10+random(800);
ind[i].alfa:=lot(ValorMaximo);
ind[i].beta:=lot(ValorMaximo);
ind[i].p1:=lot(ValorMaximo);
ind[i].p2:=lot(ValorMaximo);
ind[i].p3:=0;
ind[i].p4:=lot(ValorMaximo);
for j:=1 to 69 do
begin
ind[i].z[j]:=random;
end;
end;
end;
procedure ValoresFoto;
begin
x1:=58+20/60+80.33/3600;
x2:=38+0/60+2.22/3600;
y1:=(14+53/60+43.8/3600);
y2:=28+53/60+19.81/3600;
f1:=0;
f1p:=0;
f2:=392;
f2p:=288;
m:=(f2-f1)/(x2-x1);
mp:=(f2p-f1p)/(y2-y1);
end;
procedure CoordenadasPuertos;
var i,j:integer;
begin
p1y:=(yf(23.8063));
p1x:=(xf(45.4016));
p2y:=(yf(22.9028));
p2x:=(xf(43.2075));
p3y:=(yf(24.6778));
p3x:=(xf(47.4153));
p4y:=(yf(25.52));
p4x:=(xf(48.5092));
end;
procedure EstadoInicial;
var i:integer;
begin
for i:=1 to ene do
```

```
begin
  zona[i].estado[1]:=0;
  zona[i].estado[2]:=0;
  zona[i].estado[3]:=0;
  zona[i].estado[4]:=0;
end;
end;
```

```
procedure Coordenadas;
var i,j,i1:integer;
begin
  for i:=1 to ene do
    begin
      zona[i].x:=t2[1].x[i];
      zona[i].y:=t2[1].y[i];
      writeln(salida1.i.' '.zona[i].x.' '.zona[i].y);
    end;
  closefile(salida1);
end;
```

```
procedure orden(ip:integer);
var temp1.temp2.temp3.temp4.n.i.loc:integer;
temp0:real;
begin
  for loc:=1 to 3 do
    for i:=loc+1 to 4 do
      if zona[ip].ep[i]<zona[ip].ep[loc] then
        begin
          temp0:=zona[ip].ep[loc];
          temp1:=zona[ip].np[loc];
          zona[ip].ep[loc]:=zona[ip].ep[i];
          zona[ip].np[loc]:=zona[ip].np[i];
          zona[ip].ep[i]:=temp0;
          zona[ip].np[i]:=temp1;
        end
      end;
  end;
end;

procedure orden2(ip:integer);
var temp1.temp2.temp3.temp4.n.i.loc:integer;
temp0:real;
begin
  for loc:=1 to 3 do
    for i:=loc+1 to 4 do
      if zona[ip].score[i]>zona[ip].score[loc] then
        begin
          temp0:=zona[ip].score[loc];
          temp1:=zona[ip].nscore[loc];
          zona[ip].score[loc]:=zona[ip].score[i];
```

```
zona[ip].nscore[loc]:=zona[ip].nscore[i];
zona[ip].score[i]:=temp0;
zona[ip].nscore[i]:=temp1;
end
end; //fin orden
procedure VectorDistancia;
var i,j:integer;
pedo1.pedo2.pedo3.pedo4:real;
begin
pedo1:=p;
pedo2:=p*(1-p);
pedo3:=p*(1-(pedo1+pedo2));
pedo4:=p*(1-(pedo1+pedo2+pedo3));
for i:=1 to ene do
begin
zona[i].npf[zona[i].np[1]]:=pedo1;
zona[i].npf[zona[i].np[2]]:=pedo2;
zona[i].npf[zona[i].np[3]]:=pedo3;
zona[i].npf[zona[i].np[4]]:=pedo4;
end;
end;

procedure DistanciaPuertos;
var i,j:integer;
razon1.d1.d2.d3.d4:real;
begin
razon1:=1;
for i:=1 to ene do
begin
readln(entrada2.j.d1.d2.d3.d4);
zona[i].dp[1]:=d4/razon1;
zona[i].dp[2]:=d3/razon1;
zona[i].dp[3]:=d1/razon1;
zona[i].dp[4]:=d2/razon1;
zona[i].np[1]:=1;
zona[i].np[2]:=2;
zona[i].np[3]:=3;
zona[i].np[4]:=4;

writeln(salida5.i.' 'd1:3:0.' 'd2:3:0.' 'd3:3:0.' 'd4:3:0);
end;
closefile(salida5);
closefile(entrada2);
end;
procedure CalidadPuerto;
var i,j:integer;
begin
```

```
for i:=1 to ene do
begin
  for j:=1 to 4 do
  begin
    zona[i].ep[j]:=zona[i].dp[j]/500;
  end;
end;
end;
procedure DistanciaZonas;
var i,j:integer;
begin
  for i:=1 to ene do
  begin
    for j:=1 to ene do
    begin
      if i<>j then zona[i].d[j]:=sqrt(sqr(zona[i].x-zona[j].x)+sqr(zona[i].y-zona[j].y));
      if i=j then zona[i].d[i]:=1e6;
    end;
  end;
end;
procedure CalcularDistancias;
var i,j:integer;
begin
  for i:=1 to ene do
  begin
    orden(i);
  end;
end;
procedure NumeroScore;
var i:integer;
begin
  for i:=1 to ene do
  begin
    zona[i].nscore[1]:=1;
    zona[i].nscore[2]:=2;
    zona[i].nscore[3]:=3;
    zona[i].nscore[4]:=4;
  end;
end;

procedure TForm1.ScrollBar1Change(Sender: TObject);
begin
  port[1].p:=ScrollBar1.Position/100;
end;

procedure TForm1.ScrollBar2Change(Sender: TObject);
begin
```

```
port[1].t:=ScrollBar2.Position/100;  
end;
```

```
procedure TForm1.ScrollBar3Change(Sender: TObject);  
begin  
    port[2].p:=ScrollBar3.Position/100;  
end;
```

```
procedure TForm1.ScrollBar4Change(Sender: TObject);  
begin  
    port[2].t:=ScrollBar4.Position/100;  
end;
```

```
procedure TForm1.ScrollBar5Change(Sender: TObject);  
begin  
    port[3].p:=ScrollBar5.Position/100;  
end;
```

```
procedure TForm1.ScrollBar6Change(Sender: TObject);  
begin  
    port[3].t:=ScrollBar6.Position/100;  
end;
```

```
procedure TForm1.ScrollBar7Change(Sender: TObject);  
begin  
    port[4].p:=ScrollBar7.Position/100;  
end;  
procedure TForm1.ScrollBar8Change(Sender: TObject);  
begin  
    port[4].t:=ScrollBar8.Position/100;  
end;  
procedure TForm1.ToggleBox1Change(Sender: TObject);  
begin  
    // timer1.Enabled:=true;  
end;
```

```
procedure TForm1.ToggleBox2Change(Sender: TObject);  
begin  
    // timer1.Enabled:=true;  
end;
```

```
procedure TForm1.ToggleBox3Change(Sender: TObject);  
begin  
    // timer1.Enabled:=true;  
end;
```

```
procedure TForm1.ToggleBox4Change(Sender: TObject);
```

```

begin
// timer1.Enabled:=true;
end;

procedure Vecindad;
var i,j,suma1,suma2,suma3,suma4,total:integer;
begin
for i:=1 to ene do
begin
total:=0;
suma1:=0;
suma2:=0;
suma3:=0;
suma4:=0;
for j:=1 to ene do
begin
if zona[i].d[j]<radio then
begin
//total:=total+1;
total:=total+zona[j].estado[1]+zona[j].estado[2]+zona[j].estado[3]+zona[j].estado[4];
suma1:=suma1+zona[j].estado[1];
suma2:=suma2+zona[j].estado[2];
suma3:=suma3+zona[j].estado[3];
suma4:=suma4+zona[j].estado[4];
end;
end;
if total<>0 then zona[i].fi1[1]:=suma1/total else zona[i].fi1[1]:=0;
if total<>0 then zona[i].fi1[2]:=suma2/total else zona[i].fi1[2]:=0;
if total<>0 then zona[i].fi1[3]:=suma3/total else zona[i].fi1[3]:=0;
if total<>0 then zona[i].fi1[4]:=suma4/total else zona[i].fi1[4]:=0;
end;
end;
procedure Activacion;
var i,j:integer;
begin
for i:=1 to ene do
begin
zona[i].score[1]:=alfa*zona[i].fi1[1]+p1ac*beta*zona[i].dp[1]+port[1].p;
zona[i].score[2]:=alfa*zona[i].fi1[2]+p2ac*beta*zona[i].dp[2]+port[2].p;
zona[i].score[3]:=alfa*zona[i].fi1[3]+p3ac*beta*zona[i].dp[3]+port[3].p;
zona[i].score[4]:=alfa*zona[i].fi1[4]+p4ac*beta*zona[i].dp[4]+port[4].p;
end;
end;
procedure OrdenScore;
var i:integer;
begin
for i:=1 to ene do

```

```
begin
  orden2(i);
end;
end;
procedure Actualiza;
var i,j,k:integer;
perruno:real;
begin

  for i:=1 to ene do
  begin
    perruno:=random;
    if zona[i].umb>perruno then
    begin
      if (zona[i].nscore[1]=1) and (p1ac=1) then zona[i].estado[1]:=1 else
zona[i].estado[1]:=0;
      if (zona[i].nscore[1]=2) and (p2ac=1) then zona[i].estado[2]:=1 else
zona[i].estado[2]:=0;
      if (zona[i].nscore[1]=3) and (p3ac=1) then zona[i].estado[3]:=1 else
zona[i].estado[3]:=0;
      if (zona[i].nscore[1]=4) and (p4ac=1) then zona[i].estado[4]:=1 else
zona[i].estado[4]:=0;
    end
  else
    for j:=1 to 4 do
    begin
      zona[i].estado[j]:=0;
    end;
  end;
end;

procedure Conteo;
var i,j,n1,n2,n3,n4:integer;
begin
  n1:=0;
  n2:=0;
  n3:=0;
  n4:=0;
  for i:=1 to ene do
  begin
    if t1[tiempo].c[i]=1 then n1:=n1+1;
    if t1[tiempo].c[i]=2 then n2:=n2+1;
    if t1[tiempo].c[i]=3 then n3:=n3+1;
    if t1[tiempo].c[i]=4 then n4:=n4+1;
  end;
  nt1:=n1/ene*100;
  nt2:=n2/ene*100;
```

```
    nt3:=n3/ene*100;
    nt4:=n4/ene*100;
end;
procedure Conteo2;
var i,j.n1.n2.n3.n4:integer;
begin
    n1:=0;
    n2:=0;
    n3:=0;
    n4:=0;
    for i:=1 to ene do
        begin
            if t2[tiempo].c[i]=1 then n1:=n1+1;
            if t2[tiempo].c[i]=2 then n2:=n2+1;
            if t2[tiempo].c[i]=3 then n3:=n3+1;
            if t2[tiempo].c[i]=4 then n4:=n4+1;
        end;
        nt1s:=n1/ene*100;
        nt2s:=n2/ene*100;
        nt3s:=n3/ene*100;
        nt4s:=n4/ene*100;
    end;
procedure ErrorCuadrado;
var suma1.suma2.suma3.suma4:real;
begin
    suma1:=sqr(nt1-nt1s);
    suma2:=sqr(nt2-nt2s);
    suma3:=sqr(nt3-nt3s);
    suma4:=sqr(nt4-nt4s);
    error1:=error1+suma1;
    error2:=error2+suma2;
    error3:=error3+suma3;
    error4:=error4+suma4;
end;
procedure TForm1.Button2Click(Sender: TObject);
begin
    close;
end;
procedure LeerDatos;
var i,j.a1.a2.a3.perro.a7:integer;
    a4.a5.a6:real;
begin
    perro:=1;
    j:=0;
    for i:=1 to 5244 do
        begin
            readln(entrada.a1.a2.a3.a4.a5.a6.a7);
```

```
if perro=a1 then
begin
j:=j+1;
zona2[a1].d[j].x:=xf(a5);
zona2[a1].d[j].y:=yf(a4);
zona2[a1].d[j].z:=a6;
zona2[a1].d[j].k:=a7;
if a6<0.2 then
begin
t2[a1].c[a7]:=0;
t2[a1].x[a7]:=xf(a5);
t2[a1].y[a7]:=yf(a4);
t1[a1].x[a7]:=xf(a5);
t1[a1].y[a7]:=yf(a4);
end;
if (a6>=0.2) and (a6<0.4) then
begin
t2[a1].c[a7]:=1;
t2[a1].x[a7]:=xf(a5);
t2[a1].y[a7]:=yf(a4);
t1[a1].x[a7]:=xf(a5);
t1[a1].y[a7]:=yf(a4);
end;
if (a6>=0.4) and (a6<0.6) then
begin
t2[a1].c[a7]:=2;
t2[a1].x[a7]:=xf(a5);
t2[a1].y[a7]:=yf(a4);
t1[a1].x[a7]:=xf(a5);
t1[a1].y[a7]:=yf(a4);
end;
if (a6>=0.6) and (a6<0.8) then
begin
t2[a1].c[a7]:=3;
t2[a1].x[a7]:=xf(a5);
t2[a1].y[a7]:=yf(a4);
t1[a1].x[a7]:=xf(a5);
t1[a1].y[a7]:=yf(a4);
end;
if (a6>=0.8) then
begin
t2[a1].c[a7]:=4;
t2[a1].x[a7]:=xf(a5);
t2[a1].y[a7]:=yf(a4);
t1[a1].x[a7]:=xf(a5);
t1[a1].y[a7]:=yf(a4);
end;
```

```
end
else
begin
perro:=perro+1;
j:=1;
zona2[a1].d[j].x:=xf(a5);
zona2[a1].d[j].y:=yf(a4);
zona2[a1].d[j].z:=a6;
zona2[a1].d[j].k:=a7;
if a6<0.2 then
begin
t2[a1].c[a7]:=0;
t2[a1].x[a7]:=xf(a5);
t2[a1].y[a7]:=yf(a4);
t1[a1].x[a7]:=xf(a5);
t1[a1].y[a7]:=yf(a4);
end;
if (a6>=0.2) and (a6<0.4) then
begin
t2[a1].c[a7]:=1;
t2[a1].x[a7]:=xf(a5);
t2[a1].y[a7]:=yf(a4);
t1[a1].x[a7]:=xf(a5);
t1[a1].y[a7]:=yf(a4);
end;
if (a6>=0.4) and (a6<0.6) then
begin
t2[a1].c[a7]:=2;
t2[a1].x[a7]:=xf(a5);
t2[a1].y[a7]:=yf(a4);
t1[a1].x[a7]:=xf(a5);
t1[a1].y[a7]:=yf(a4);
end;
if (a6>=0.6) and (a6<0.8) then
begin
t2[a1].c[a7]:=3;
t2[a1].x[a7]:=xf(a5);
t2[a1].y[a7]:=yf(a4);
t1[a1].x[a7]:=xf(a5);
t1[a1].y[a7]:=yf(a4);
end;
if (a6>=0.8) then
begin
t2[a1].c[a7]:=4;
t2[a1].x[a7]:=xf(a5);
t2[a1].y[a7]:=yf(a4);
t1[a1].x[a7]:=xf(a5);
```

```

    if (t1[j].c[k]=t2[j].c[k]) then
    begin
    suma3:=1;
    suma2:=1;
    suma4:=1;
    end;
    suma1:=suma1+suma2+suma3;
    suma5:=suma5+suma3+suma4;
    end;
end;
SumaTotal:=(suma1/(76*69));
Sumatotal2:=suma5/(76*69);
end;
procedure DeterminaCalidad2;
var j,k.suma1.suma2.suma3.suma4.suma5:integer;
begin
    suma1:=0;
    suma4:=0;
    for j:=1 to 76 do
    begin
    for k:=1 to 69 do
    begin
    if (t1[j].c[k]<>t2[j].c[k]) then suma2:=0;
    if (t1[j].c[k]<>t2[j].c[k]) then suma3:=0;
    if (t1[j].c[k]=t2[j].c[k]) and (t1[j].c[k]<>0) then suma2:=naciertos;
    if (t1[j].c[k]=t2[j].c[k]) and (t1[j].c[k]<>0) then suma3:=1;
    if (t1[j].c[k]=t2[j].c[k]) and (t1[j].c[k]=0) then suma2:=1;
    if (t1[j].c[k]=t2[j].c[k]) and (t1[j].c[k]=0) then suma3:=1;
    suma1:=suma1+suma2;
    suma4:=suma4+suma3;
    end;
    end;
    SumaTotal:=(suma1/(ntotal));
    SumaTotal2:=(suma4/(76*69));
end;
procedure ConteoAciertos;
var j,k.suma1.suma2.suma3.suma4.suma5:integer;
begin
    suma1:=0;
    for j:=1 to 76 do
    begin
    for k:=1 to 69 do
    begin
    if (t2[j].c[k]<>0) then suma2:=naciertos;
    if (t2[j].c[k]=0) then suma2:=1;
    suma1:=suma1+suma2;
    end;

```

```
end;
ntotal:=suma1;
end;
procedure corrida;
var j,k:integer;
begin
    tiempo:=tiempo+1;
    CalidadPuerto;
    CalcularDistancias;
    VectorDistancia;
    for k:=1 to ene do
begin
    zona[k].np[1]:=1;
    zona[k].np[2]:=2;
    zona[k].np[3]:=3;
    zona[k].np[4]:=4;
end;

    NumeroScore;
    Vecindad;
    Activacion;
    OrdenScore;
    Actualiza;
    Embolsar;
    GuardaDATos;
end;

procedure Simulacion(i:integer);
var j,k:integer;
begin
    error1:=0;
    error2:=0;
    error3:=0;
    error4:=0;
    radio:=ind[i].radio;
    alfa:=ind[i].alfa;
    beta:=ind[i].beta;
    port[1].p:=ind[i].p1;
    port[2].p:=ind[i].p2;
    port[3].p:=ind[i].p3;
    port[4].p:=ind[i].p4;
    port[1].t:=1;
    port[2].t:=1;
    port[3].t:=1;
    port[4].t:=1;
    for j:=1 to 69 do
begin
```

```

    zona[j].umb:=ind[i].z[j];
end;
EstadoInicial;
tiempo:=0;
for k:=1 to 76 do
begin
Corrida;
Conteo;
Conteo2;
ErrorCuadrado;
end;
DatoError:=(1/(sqrt(error1)/76+sqrt(error2)/76+sqrt(error3)/76+sqrt(error4)/76);
DeterminaCalidad2;
cal[i].q:=SumaTotal+DatoError;
cal2[i].q:=SumaTotal2;
// cal[i].q:=DatoError;
// cal2[i].q:=DatoError;
cal[i].ij:=i;
cal2[i].ij:=i;
end;
procedure CalculoParaUnaGeneracion;
var i,j,k:integer;
begin
for i:=1 to 100 do //Esto permite la simulacion para cada individuo
begin
Simulacion(i);
end; //fin for i
end;
procedure OrdeneLosIndividuos;
var temp1.temp2.temp3.temp4.n.i.loc.temp12:integer;
temp0.temp02:real;
begin
for loc:=1 to 99 do
for i:=loc+1 to 100 do
if cal[i].q>cal[loc].q then
begin
temp0:=cal[loc].q;
temp1:=cal[loc].ij;
cal[loc].q:=cal[i].q;
cal[loc].ij:=cal[i].ij;
cal[i].q:=temp0;
cal[i].ij:=temp1;
end
end; //fin orden
procedure OrdeneLosIndividuos2;
var temp1.temp2.temp3.temp4.n.i.loc.temp12:integer;
temp0.temp02:real;

```

```
begin
  for loc:=1 to 99 do
    for i:=loc+1 to 100 do
      if cal2[i].q>cal2[loc].q then
        begin
          temp0:=cal2[loc].q;
          temp1:=cal2[loc].ij;
          cal2[loc].q:=cal2[i].q;
          cal2[loc].ij:=cal2[i].ij;
          cal2[i].q:=temp0;
          cal2[i].ij:=temp1;
        end
      end;
    end;
  end;
end;

procedure TrasladeLaInformacion;
var i,j:integer;
begin
  for i:=1 to 100 do
    begin
      ind2[i].radio:=ind[i].radio;
      ind2[i].alfa:=ind[i].alfa;
      ind2[i].p1:=ind[i].p1;
      ind2[i].p2:=ind[i].p2;
      ind2[i].p3:=ind[i].p3;
      ind2[i].p4:=ind[i].p4;
      ind2[i].beta:=ind[i].beta;
      for j:=1 to 69 do
        begin
          ind2[i].z[j]:=ind[i].z[j];
        end;
      end;
    end;
  end;
end;

procedure MejorIndividuo;
var i,j,k:integer;
begin
  i:=cal[1].ij;
  ind[1].radio:=ind2[i].radio;
  ind[1].alfa:=ind2[i].alfa;
  ind[1].p1:=ind2[i].p1;
  ind[1].p2:=ind2[i].p2;
  ind[1].p3:=ind2[i].p3;
  ind[1].p4:=ind2[i].p4;
  ind[1].beta:=ind2[i].beta;
  for j:=1 to 69 do
    begin
      ind[1].z[j]:=ind[i].z[j];
    end;
  end;
end;
```

```

procedure SetentaHijos;
var i,j,k,j1,j2,l,j0,k0:integer;
s1:real;
begin
  for i:=2 to nhijos do
    begin
      { j0:=1+random(20);
        k0:=1+random(100);
        j:=taco[j0];
        k:=taco[k0];}
      j:=1;
      k:=1+random(100);
      s1:=random;
      if s1<0.5 then ind[i].radio:=ind2[j].radio else ind[i].radio:=ind2[k].radio;
      s1:=random;
      if s1<0.5 then ind[i].alfa:=ind2[j].alfa else ind[i].alfa:=ind2[k].alfa;
      s1:=random;
      if s1<0.5 then ind[i].p1:=ind2[j].p1 else ind[i].p1:=ind2[k].p1;
      s1:=random;
      if s1<0.5 then ind[i].p2:=ind2[j].p2 else ind[i].p2:=ind2[k].p2;
      s1:=random;
      if s1<0.5 then ind[i].p3:=ind2[j].p3 else ind[i].p3:=ind2[k].p3;
      s1:=random;
      if s1<0.5 then ind[i].p4:=ind2[j].p4 else ind[i].p4:=ind2[k].p4;
      s1:=random;
      if s1<0.5 then ind[i].beta:=ind2[j].beta else ind[i].beta:=ind2[k].beta;
      l:=5+random(60);
      for j1:=1 to l do
        begin
          ind[i].z[j1]:=ind2[i].z[j1];
        end;
      for j2:=l+1 to 69 do
        begin
          ind[i].z[j2]:=ind2[k].z[j2];
        end;
      end;
    end;
end;
procedure Mutantes;
var i,j,k,j1,k2,j0,k3,k4,k5,k6,k7:integer;
begin
  for i:=nhijos+1 to nhijos+nmutantes do
    begin
      {j0:=1+random(100);
        j:=1+taco[j0];}
      j:=1;
      ind[i].radio:=ind2[j].radio;
      ind[i].alfa:=ind2[j].alfa;
    end;
end;

```

```
ind[i].p1:=ind2[j].p1;
ind[i].p2:=ind2[j].p2;
ind[i].p3:=ind2[j].p3;
ind[i].p4:=ind2[j].p4;
ind[i].beta:=ind2[j].beta;
for k:=1 to 69 do
begin
ind[i].z[k]:=ind2[j].z[k];
end;
k:=1+random(9);
if k=1 then ind[i].radio:=10+random(800);
if k=2 then ind[i].alfa:=lot(ValorMaximo);
if k=3 then ind[i].beta:=lot(ValorMaximo);
if k=5 then ind[i].p1:=lot(ValorMaximo);
if k=6 then ind[i].p2:=lot(ValorMaximo);
if k=7 then ind[i].p3:=0;
if k=8 then ind[i].p4:=lot(ValorMaximo);
k2:=1+random(69);
ind[i].z[k2]:=random;
k3:=1+random(69);
ind[i].z[k3]:=random;
k4:=1+random(69);
ind[i].z[k4]:=random;
k5:=1+random(69);
ind[i].z[k5]:=random;
k6:=1+random(69);
ind[i].z[k6]:=random;
k7:=1+random(69);
ind[i].z[k7]:=random;
end;
end;
procedure MutantesAzar;
var i,j:integer;
begin
for i:=nhijos+nmigrantes+1 to 100 do
begin
ind[i].radio:=10+random(800);
ind[i].alfa:=lot(ValorMaximo);
ind[i].beta:=lot(ValorMaximo);
ind[i].p1:=lot(ValorMaximo);
ind[i].p2:=lot(ValorMaximo);
ind[i].p3:=0;
ind[i].p4:=lot(ValorMaximo);
for j:=1 to 69 do
begin
ind[i].z[j]:=random;
end;
end;
```

```

    end;
  end;
procedure Diversidad1;
var
  i,j,k:integer;
  d1.d2.d3.d4.d5.d6.d7.d8.d9.d8b:real;
begin
  i:=cal[1].ij;
  for k:=1 to 100 do
    begin
      if i<>k then
        begin
          d1:=sqr(ind2[i].radio-ind2[k].radio);
          d2:=sqr(ind2[i].alfa-ind2[k].alfa);
          d5:=sqr(ind2[i].p1-ind2[k].p1);
          d6:=sqr(ind2[i].p2-ind2[k].p2);
          d7:=sqr(ind2[i].p3-ind2[k].p3);
          d8:=sqr(ind2[i].p4-ind2[k].p4);
          d8b:=sqr(ind2[i].beta-ind2[k].beta);
          d9:=d1+d2+d3+d4+d5+d6+d7+d8+d8b;
          for j:=1 to 69 do
            begin
              d9:=d9+sqr(ind2[i].z[j]-ind2[k].z[j]);
            end;
          end;
          cony[k].d:=1/(d9+0.0001);
          if i=k then cony[k].d:=1e8;
          cony[k].k:=k;
        end;
      end;
    end;
  end;
procedure Diversidad2;
var
  i,j,k,j1:integer;
  d1.d2.d3.d4.d5.d6.d7.d8.d9.d8b:real;
begin
  i:=cony[1].k;
  for j1:=1 to 100 do
    begin
      k:=cony[j1].k;
      if i<>k then
        begin
          d1:=sqr(ind2[i].radio-ind2[k].radio);
          d2:=sqr(ind2[i].alfa-ind2[k].alfa);
          d5:=sqr(ind2[i].p1-ind2[k].p1);
          d6:=sqr(ind2[i].p2-ind2[k].p2);
          d7:=sqr(ind2[i].p3-ind2[k].p3);

```

```

    d8:=sqr(ind2[i].p4-ind2[k].p4);
    d8b:=sqr(ind2[i].beta-ind2[k].beta);
    d9:=d1+d2+d3+d4+d5+d6+d7+d8+d8b;
    for j:=1 to 69 do
    begin
        d9:=d9+sqr(ind2[i].z[j]-ind2[k].z[j]);
    end;
end;
cony[k].d:=cony[k].d+1/(d9+0.0001);
if i=k then cony[k].d:=1e8;
cony[k].k:=k;
end;
end;
procedure OrdeneCony;
var
    temp1.temp2.temp3.temp4.n.i.loc.temp12.j.k:integer;
    temp0.temp02:real;
begin
    for loc:=1 to 99 do
        for i:=loc+1 to 100 do
            if cony[i].d<cony[loc].d then
                begin
                    temp0:=cony[loc].d;
                    temp1:=cony[loc].k;
                    cony[loc].d:=cony[i].d;
                    cony[loc].k:=cony[i].k;
                    cony[i].d:=temp0;
                    cony[i].k:=temp1;
                end
            end;
end;
procedure DiversidadGenes;
var p1:integer;
begin
    p1:=1;
    Diversidad1;
    OrdeneCony;
    taco[1]:=cony[1].k;
    repeat
        Diversidad2;
        OrdeneCony;
        p1:=p1+1;
        taco[p1]:=cony[1].k
    until p1>10;
end;

procedure RepiteConElMejor;
var j.k:integer;

```

```
begin
  radio:=ind[1].radio;
  alfa:=ind[1].alfa;
  beta:=ind[1].beta;
  port[1].p:=ind[1].p1;
  port[2].p:=ind[1].p2;
  port[3].p:=ind[1].p3;
  port[4].p:=ind[1].p4;
  for j:=1 to 69 do
    begin
      zona[j].umb:=ind[1].z[j];
    end;
  tiempo:=0;
  EstadoInicial;
  for k:=1 to 76 do
    begin
      Corrida;
    end;
  end;
procedure AlgoritmoGenetico;
var i,j,k:integer;
begin
  CalculoParaUnaGeneracion;
  OrdeneLosIndividuos;
  OrdeneLosIndividuos2;
  TrasladeLaInformacion;
  MejorIndividuo;
  DiversidadGenes;
  SetentaHijos;
  Mutantes;
  MutantesAzar;
  RepiteConElMejor;
end;

procedure TForm1.Button3Click(Sender: TObject);
var i,j:integer;
begin
  timer2.Enabled:=false;
  tiempo2:=tiempo2+1;
  //label15.Visible:=false;
  AlgoritmoGenetico;
  //label14.Visible:=true;
  //Chart1LineSeries1.AddXY(tiempo2.cal[1].q);
  //Chart1LineSeries1.AddXY(tiempo2.cal2[cal[1].ij].q);
  //Chart1LineSeries2.AddXY(tiempo2.ntotal);
  //Chart1LineSeries3.AddXY(tiempo2.cal[80].q);
  //Chart1LineSeries4.AddXY(tiempo2.cal[100].q);
```

```
//Form1.Button1.Click;
if tiempo2>1 then Label20.Visible:=true;
if tiempo2>100 then
begin
Label21.Visible:=true;
Label20.Visible:=false;
end;
if tiempo2>500 then
begin
Label22.Visible:=true;
Label21.Visible:=false;
end;
if tiempo2>800 then
begin
Label22.Visible:=true;
Label23.Visible:=false;
end;
if (cal2[cal[1].ij].q)>Nac/100 then
//if tiempo2=2 then
begin
tiempo:=0;
image3.Visible:=true;
Label18.Visible:=true;
end

else timer2.Enabled:=true;
end;
procedure TForm1.Button1Click(Sender: TObject);
var i,j,k.sum:integer;
begin

Image1.Repaint;
Image2.Repaint;
tiempo:=tiempo+1;
if ToggleBox1.Checked=true then p1ac:=1 else p1ac:=0;
if ToggleBox2.Checked=true then p2ac:=1 else p2ac:=0;
if ToggleBox3.Checked=true then p3ac:=1 else p3ac:=0;
if ToggleBox4.Checked=true then p4ac:=1 else p4ac:=0;

// Conteo;
if p1ac=1 then Image1.Canvas.Brush.color:=clblue else
Image1.Canvas.Brush.color:=clblack;
Image1.Canvas.Rectangle(p1x.p1y.p1x+15.p1y+15);
if p2ac=1 then Image1.Canvas.Brush.color:=clred else
Image1.Canvas.Brush.color:=clblack;
Image1.Canvas.Rectangle(p2x.p2y.p2x+15.p2y+15);
```

```

if p3ac=1 then Image1.Canvas.Brush.color:=clyellow else
Image1.Canvas.Brush.color:=clblack;
Image1.Canvas.Rectangle(p3x.p3y.p3x+15.p3y+15);
if p4ac=1 then Image1.Canvas.Brush.color:=clLime else
Image1.Canvas.Brush.color:=clblack;
Image1.Canvas.Rectangle(p4x.p4y.p4x+15.p4y+15);
Image1.Canvas.Brush.color:=clwhite;
//dibujo de los puertos en imagen2
if ToggleBox1.Checked=true then p1ac:=1 else p1ac:=0;
if ToggleBox2.Checked=true then p2ac:=1 else p2ac:=0;
if ToggleBox3.Checked=true then p3ac:=1 else p3ac:=0;
if ToggleBox4.Checked=true then p4ac:=1 else p4ac:=0;
if p1ac=1 then Image2.Canvas.Brush.color:=clblue else
Image2.Canvas.Brush.color:=clblack;
Image2.Canvas.Rectangle(p1x.p1y.p1x+15.p1y+15);
if p2ac=1 then Image2.Canvas.Brush.color:=clred else
Image2.Canvas.Brush.color:=clblack;
Image2.Canvas.Rectangle(p2x.p2y.p2x+15.p2y+15);
if p3ac=1 then Image2.Canvas.Brush.color:=clyellow else
Image2.Canvas.Brush.color:=clblack;
Image2.Canvas.Rectangle(p3x.p3y.p3x+15.p3y+15);
if p4ac=1 then Image2.Canvas.Brush.color:=clLime else
Image2.Canvas.Brush.color:=clblack;
Image2.Canvas.Rectangle(p4x.p4y.p4x+15.p4y+15);
//Fin Dibujo en Imagen2
{ for i:=1 to ene do
begin
sum:=zona[i].estado[1]+zona[i].estado[2]+zona[i].estado[3]+zona[i].estado[4];
if zona[i].estado[1]=1 then Image1.Canvas.Brush.color:=clblue;
if zona[i].estado[2]=1 then Image1.Canvas.Brush.color:=clred;
if zona[i].estado[3]=1 then Image1.Canvas.Brush.color:=clyellow;
if zona[i].estado[4]=1 then Image1.Canvas.Brush.color:=clLime;
if sum=0 then Image1.Canvas.Brush.color:=clwhite;
Image1.Canvas.Ellipse(zona[i].x.zona[i].y.zona[i].x+10.zona[i].y+10);
end; }
for i:=1 to ene do
begin
//Dibujo en Imagen1
if t1[tiempo].c[i]=1 then Image1.Canvas.Brush.color:=clblue;
if t1[tiempo].c[i]=2 then Image1.Canvas.Brush.color:=clred;
if t1[tiempo].c[i]=3 then Image1.Canvas.Brush.color:=clyellow;
if t1[tiempo].c[i]=4 then Image1.Canvas.Brush.color:=clLime;
if t1[tiempo].c[i]=0 then Image1.Canvas.Brush.color:=clwhite;

Image1.Canvas.Ellipse(t1[tiempo].x[i].t1[tiempo].y[i].t1[tiempo].x[i]+9.t1[tiempo].y[i]+
9);
//Dibujo en Imagen2

```

```

    if t2[tiempo].c[i]=1 then Image2.Canvas.Brush.color:=clblue;
    if t2[tiempo].c[i]=2 then Image2.Canvas.Brush.color:=clred;
    if t2[tiempo].c[i]=3 then Image2.Canvas.Brush.color:=clyellow;
    if t2[tiempo].c[i]=4 then Image2.Canvas.Brush.color:=clLime;
    if t2[tiempo].c[i]=0 then Image2.Canvas.Brush.color:=clwhite;

Image2.Canvas.Ellipse(t2[tiempo].x[i].t2[tiempo].y[i].t2[tiempo].x[i]+9.t2[tiempo].y[i]+
9);
end;
// Chart1LineSeries1.AddXY(tiempo2.cal[1].q);
// Chart1LineSeries1.AddXY(tiempo.nt1);
// Chart1LineSeries2.AddXY(tiempo.nt2);
// Chart1LineSeries3.AddXY(tiempo.nt3);
// Chart1LineSeries4.AddXY(tiempo.nt4);
conteo;
conteo2;
Chart2LineSeries1.AddXY(tiempo.nt1);
Chart2LineSeries2.AddXY(tiempo.nt2);
Chart2LineSeries3.AddXY(tiempo.nt3);
Chart2LineSeries4.AddXY(tiempo.nt4);
Chart3LineSeries2.AddXY(tiempo.nt1s);
Chart3LineSeries1.AddXY(tiempo.nt2s);
Chart3LineSeries3.AddXY(tiempo.nt3s);
Chart3LineSeries4.AddXY(tiempo.nt4s);
writeln(salida4.tiempo.' '.nt1s:3:2.' '.nt2s:3:2.' '.nt3s:3:2.' '.nt4s:3:2);
if tiempo>=74 then
begin
label14.Visible:=false;
timer1.Enabled:=false;
label15.Visible:=true;
closefile(salida4);
end;

end;
procedure TForm1.Button4Click(Sender: TObject);
var i.sum.j:integer;
begin
Randomize;
p:=0.75;
assignfile(entrada.'mapatri2.dat');
reset (entrada);
assignfile(entrada2.'matrizdistancia.dat');
reset(entrada2);

assignfile(salida1.'coordenadas.dat');
rewrite (salida1);
assignfile(salida4.'DinamicaReal.dat');
```

```
rewrite (salida4);
assignfile(salida5.'DistanciaPuertos.dat');
rewrite (salida5);
ValoresFoto;
CoordenadasPuertos;
LeerDatos;
Chart1LineSeries1.clear;
Chart1LineSeries2.clear;
Chart1LineSeries3.clear;
Chart1LineSeries4.clear;
tiempo:=0;
tiempo2:=0;
ene:=69;
val(Edit1.Text.Nac.zz);
val(Edit2.Text.nhijos.zz);
val(Edit3.Text.nmutantes.zz);
val(Edit4.Text.naciertos.zz);
val(Edit6.Text.ValorMaximo.zz);
// val(Edit5.Text.ese.zz);
port[1].p:=ScrollBar1.Position/100;
port[1].t:=ScrollBar2.Position/100;
port[2].p:=ScrollBar3.Position/100;
port[2].t:=ScrollBar4.Position/100;
port[3].p:=ScrollBar5.Position/100;
port[3].t:=ScrollBar6.Position/100;
port[4].p:=ScrollBar7.Position/100;
port[4].t:=ScrollBar8.Position/100;
Coordenadas;
DistanciaPuertos;
DistanciaZonas;
CalidadPuerto;
CalcularDistancias;
VectorDistancia;
EstadoInicial;
ValoresAzarIndividuos; //Debo sacar este procedimiento
assignfile(salida2.'ParametrosGlobales.dat');
rewrite (salida2);
assignfile(salida3.'UmbralZonas.dat');
rewrite (salida3);
ConteoAciertos;
//Dibujar puertos en Imagen1
if ToggleBox1.Checked=true then p1ac:=1 else p1ac:=0;
if ToggleBox2.Checked=true then p2ac:=1 else p2ac:=0;
if ToggleBox3.Checked=true then p3ac:=1 else p3ac:=0;
if ToggleBox4.Checked=true then p4ac:=1 else p4ac:=0;
if p1ac=1 then Image1.Canvas.Brush.color:=clblue else
Image1.Canvas.Brush.color:=clblack;
```

```

Image1.Canvas.Rectangle(p1x.p1y.p1x+10.p1y+10);
if p2ac=1 then Image1.Canvas.Brush.color:=clred else
Image1.Canvas.Brush.color:=clblack;
Image1.Canvas.Rectangle(p2x.p2y.p2x+10.p2y+10);
if p3ac=1 then Image1.Canvas.Brush.color:=clyellow else
Image1.Canvas.Brush.color:=clblack;
Image1.Canvas.Rectangle(p3x.p3y.p3x+10.p3y+10);
if p4ac=1 then Image1.Canvas.Brush.color:=clLime else
Image1.Canvas.Brush.color:=clblack;
Image1.Canvas.Rectangle(p4x.p4y.p4x+10.p4y+10);
//Dibujar puertos en Imagen2
if ToggleBox1.Checked=true then p1ac:=1 else p1ac:=0;
if ToggleBox2.Checked=true then p2ac:=1 else p2ac:=0;
if ToggleBox3.Checked=true then p3ac:=1 else p3ac:=0;
if ToggleBox4.Checked=true then p4ac:=1 else p4ac:=0;
if p1ac=1 then Image2.Canvas.Brush.color:=clblue else
Image2.Canvas.Brush.color:=clblack;
Image2.Canvas.Rectangle(p1x.p1y.p1x+10.p1y+10);
if p2ac=1 then Image2.Canvas.Brush.color:=clred else
Image2.Canvas.Brush.color:=clblack;
Image2.Canvas.Rectangle(p2x.p2y.p2x+10.p2y+10);
if p3ac=1 then Image2.Canvas.Brush.color:=clyellow else
Image2.Canvas.Brush.color:=clblack;
Image2.Canvas.Rectangle(p3x.p3y.p3x+10.p3y+10);
if p4ac=1 then Image2.Canvas.Brush.color:=clLime else
Image2.Canvas.Brush.color:=clblack;
Image2.Canvas.Rectangle(p4x.p4y.p4x+10.p4y+10);
//Fin Dibujo en Imagen2
Image1.Canvas.Brush.color:=clwhite;
{ for i:=1 to ene do
begin
sum:=zona[i].estado[1]+zona[i].estado[2]+zona[i].estado[3]+zona[i].estado[4];
if zona[i].estado[1]=1 then Image1.Canvas.Brush.color:=clblue;
if zona[i].estado[2]=1 then Image1.Canvas.Brush.color:=clred;
if zona[i].estado[3]=1 then Image1.Canvas.Brush.color:=clyellow;
if zona[i].estado[4]=1 then Image1.Canvas.Brush.color:=clLime;
if sum=0 then Image1.Canvas.Brush.color:=clwhite;
Image1.Canvas.Ellipse(zona[i].x.zona[i].y.zona[i].x+10.zona[i].y+10);
end; }
timer2.Enabled:=true;
end;

procedure TForm1.Button5Click(Sender: TObject);
begin
timer1.Enabled:=false;
end;

```

```
procedure TForm1.Button6Click(Sender: TObject);
begin
    timer1.Enabled:=true;
end;

procedure TForm1.Button7Click(Sender: TObject);
begin
    tiempo:=0;
    timer2.Enabled:=false;
end;

procedure TForm1.Button8Click(Sender: TObject);
var i:integer;
begin
    image3.Visible:=false;
    Label18.Visible:=false;
    { writeln(salida2.'Radio='.radio:3:5);
    writeln(salida2.'Alfa='.alfa:3:5);
    writeln(salida2.'Umbral='.umbral:3:5);
    writeln(salida2.'Sigma='.ese:3:5);
    writeln(salida2.'Puerto1='.port[1].p:3:5);
    writeln(salida2.'Puerto2='.port[2].p:3:5);
    writeln(salida2.'Puerto3='.port[3].p:3:5);
    writeln(salida2.'Puerto4='.port[4].p:3:5); }
    // writeln(salida2.358/67*radio:3:0.' '.alfa:3:4.' '.beta:3:4.' '.port[1].p:3:4.'
    '.port[2].p:3:4.' '.port[3].p:3:4.' '.port[4].p:3:4);
    writeln(salida2.358/67*radio:3:0.' '.alfa:3:0.' '.beta:3:0.' '.port[1].p:3:0.'
    '.port[2].p:3:0.' '.port[3].p:3:0.' '.port[4].p:3:0);
    closefile(salida2);
    for i:=1 to 69 do
    begin
        writeln(salida3.i.' '.ind[1].z[i]:3:5);
    end;
    closefile(salida3);
    timer1.Enabled:=true;
end;

procedure TForm1.Button9Click(Sender: TObject);
begin
    semilla:=1;
end;

procedure TForm1.Image2Click(Sender: TObject);
begin

end;
```

end.