TESIS DE LA UNIVERSIDAD DE ZARAGOZA
2013
116

Jorge Enrique Barnett Lawton

Location models applied to the design of a distribution network for hydrogen fuel produced from renewable energy sources

Departamento
Zaragoza Logistics Center

Director/es
Gonsalvez, David
Menezes, Mozart Batista de Castro

http://zaguan.unizar.es/collection/Tesis

ISSN 2254-7606

Prensas de la Universidad
Universidad Zaragoza

Universidad Zaragoza
1542

Tesis Doctoral

LOCATION MODELS APPLIED TO THE DESIGN OF A DISTRIBUTION NETWORK FOR HYDROGEN FUEL PRODUCED FROM RENEWABLE ENERGY SOURCES

Autor

Jorge Enrique Barnett Lawton

Director/es

Gonsalvez, David
Menezes, Mozart Batista de Castro

UNIVERSIDAD DE ZARAGOZA

Zaragoza Logistics Center

2013

UNIVERSIDAD DE ZARAGOZA

TESIS DOCTORAL

Modelos de localización aplicados al diseño de una red de distribución de combustible de hidrógeno producido a partir de fuentes de energía renovables

JORGE ENRIQUE BARNETT LAWTON

M.S., Ingeniería Industrial, Georgia Institute of Technology, EE.UU.

Ingeniero Industrial, Universidad Tecnológica de Panamá, Panamá

8 de julio de 2013

Modelos de localización aplicados al diseño de una red de distribución de combustible de hidrógeno producido a partir de fuentes de energía renovables

por

Jorge Enrique Barnett Lawton

M.S., Georgia Institute of Technology
B.S., Universidad Tecnológica de Panamá

Presentado al Programa Internacional de Logística MIT-Zaragoza
como parte de las condiciones para la obtención del título de

Doctor en Filosofía en Logística y Gestión de la Cadena de
Suministro

del

ZARAGOZA LOGISTICS CENTER

Julio 2013

Autor . . . .
Programa Internacional de Logística MIT-Zaragoza
8 de julio de 2013

Director de tesis . . . .
Mozart Batista de Castro Menezes
Profesor Titular de Gestión de Operaciones
University of Calgary, Calgary, Canada

Co-director de tesis . . . .
David Gonsalvez
Profesor de Gestión de la Cadena de Suministro
Programa Internacional de Logística MIT-Zaragoza

A mis padres, Claudia y Jorge

Agradecimientos

Decir que esta travesía ha sido difícil sería quedarse corto. De igual forma, decir que esta meta fue alcanzada sólo por mi esfuerzo sería subestimar lo importante e influyente que fue el apoyo de mis colegas y seres queridos durante esta aventura. Muchas personas –muchas más que las que estas páginas me permiten– merecen mi aprecio por su estímulo o apoyo durante los últimos años, y lo más que puedo hacer es expresarles a todos mi gratitud.

Aun así me gustaría aprovechar esta oportunidad para reconocer a algunas personas que jugaron un papel importante a lo largo de esta travesía, y sin las cuales esta tesis no sería más que una mera ilusión.

Primero debo agradecer a mi director de tesis, Mozart Menezes, quien aparte de mentor fue una amable voz de razón que se aseguró de mantenerme enfocado cuando lo necesitaba, y me animó a dar mi mejor esfuerzo cuando mi voluntad menguaba. También doy gracias a mi co-director de tesis, Jarrod Goentzel, por su apoyo a lo largo de este proceso, en especial durante el desafiante año en MIT.

Quisiera agradecer a Mustafa C̃agri Gürbüz, David Gonsalvez, María Jesús Sáenz, Santiago Kraiselburd y Elena Fernández por los roles que asumieron durante mi tiempo en el ZLC. Extiendo esa gratitud al resto de la familia del ZLC.

Estoy agradecido con mis colegas y amigos en Zaragoza, Boston y Panamá –en especial Gerardo, Lijie, Eirini, Mario, Marios, Roberto M., Roberto P.-F., Emiliano, y Angel– que con sus opiniones, consejos y humor hicieron que este fuera un camino más interesante y manejable. Extiendo mi más sincero aprecio a Laura que, en un tiempo en que me hubiera resultado fácil rendirme, siempre fue una voz de aliento; no puedo agradecer lo suficiente tu generosidad y amistad. También quiero agradecer a Lineth, quien estuvo a mi lado y fue mi apoyo durante parte sustancial de este desafío.

Finalmente, quiero expresar mi más inmensa gratitud a mis padres –Claudia y Jorge– y mis hermanas –Lyra y Lisa– quienes siempre creyeron en mí. Espero haberlos llenado de orgullo.

Índice

Lista de Ilustraciones

Lista de Tablas

Modelos de localización aplicados al diseño de una red de distribución de combustible de hidrógeno producido a partir de fuentes de energía renovables

por

Jorge Enrique Barnett Lawton

Presentado al Programa Internacional de Logística MIT-Zaragoza
como parte de las condiciones para la obtención del título de
Doctor en Filosofía en Logística y Gestión de la Cadena de Suministro

Resumen

El agotamiento de las reservas de combustibles fósiles, la creciente preocupación sobre el posible impacto a corto y mediano plazo de las emisiones de gases invernadero en la sociedad y el medio ambiente, y las resultantes políticas orientadas hacia energías más limpias y menor dependencia de combustibles fósiles, han todos puesto mayor presión en incrementar la capacidad de generar energía a partir de fuentes renovables, de las cuales la energía eólica ha sido uno de los segmentos de más rápido desarrollo. Sin embargo, la naturaleza intermitente e impredecible de estas fuentes, combinada con las limitaciones de la actual infraestructura de transmisión, se traduce en una cantidad significativa de energía eólica desperdiciada y en que no se llegue a explotar adecuadamente la capacidad instalada de generación.

El hidrógeno ha sido identificado como un vector energético limpio para almacenar la energía producida a partir de dichas fuentes variables. Adicionalmente, el hidrógeno ha demostrado su factibilidad como combustible de cero emisiones para uso en transporte, sea a través de sistemas de combustión o celdas de combustible. El uso del hidrógeno como mecanismo de almacenaje puede ayudar a absorber las fluctuaciones en el suministro de energía, y potencialmente aumentar la utilización de la capacidad disponible de generación renovable.

El desarrollo de una infraestructura de producción y distribución de hidrógeno es un tema de gran relevancia en la evaluación de la viabilidad a largo plazo de este vector energético. Esta tesis doctoral se centra en las decisiones enfrentadas por una empresa con capacidad de generación de energía renovable, cuando también le es posible producir hidrógeno por medio de electrólisis conectada a la red. Primero nos enfocamos en las decisiones de diseño de la cadena de suministro para el productor; específicamente, dónde localizar las instalaciones de producción de hidrógeno, cuánta capacidad construir en cada instalación, y cuánto hidrógeno debe ser distribuido a instancias posteriores en la red.

Presentamos una serie de modelos de optimización para el diseño de la red de

producción y distribución, considerando el valor adicional creado por la producción de hidrógeno para una empresa generadora, incorporando la incertidumbre en el suministro de energía renovable y en los precios de electricidad. La empresa opera en un mercado regulado por una agencia externa con la capacidad de fijar precios para el combustible de hidrógeno y establecer políticas generales de servicio para el mercado siendo considerado por el productor. Los Capítulos 2 y 3 de la tesis están enfocados en la problemática del productor bajo dos políticas regulatorias de distribución: selección de mercados y asignación proporcional.

Para abordar este problema presentamos procedimientos de solución basados en métodos de generación de columnas. A manera de fundamentar nuestro trabajo en un contexto práctico, en el Capítulo 4 implementamos nuestros modelos para evaluar una futura red de combustible de hidrógeno en España, y evaluamos el efecto de los incentivos externos (subsidios) sobre las decisiones del productor, permitiéndonos ganar intuición sobre cómo dichos incentivos pueden ser puestos en marcha para hacer viable una economía de hidrógeno.

Conclusiones y Extensiones

Esta tesis ha abordado el problema de maximización de ganancias enfrentado por una empresa con capacidad de generación de energía renovable, cuando cuenta con la opción de producir hidrógeno por medio de un proceso de electrólisis conectada a la red, para su distribución a consumidores que lo utilizan como combustible alternativo. Presentamos las propiedades de la función de ganancia esperada de la empresa productora, además de un método exacto basado en generación de columnas para resolver este problema utilizando dos políticas de distribución alternativas.

A manera de fundamentar nuestro trabajo en un contexto práctico, también presentamos un caso de estudio numérico que extiende esfuerzos anteriores de investigación sobre una potencial cadena de suministros para combustible de hidrógeno en España. Nuestros resultados indican la prevalencia de producción descentralizada en el caso que el productor puede seleccionar los mercados a servir, inducido por las distancias relativamente largas entre los puntos de demanda, que se acentúa en los escenarios de alta penetración de mercado del combustible de hidrógeno y de eficiencia futura del proceso de electrólisis. Adicionalmente, en base al estado actual de la tecnología de producción, nuestros experimentos computacionales indican que un esquema de incentivos basado en reducir los costos de equipo e infraestructura asumidos por los productores resultaría más viable para la agencia reguladora que la implementación de un subsidio del precio del producto final, en el caso que se desee lograr cobertura total de la demanda a nivel local y espacial (satisfacción completa de la demanda esperada para todos los nodos de la red).

Nuestro trabajo, a nuestro saber, es el primer modelo de diseño de redes que incorpora incertidumbre tanto para el suministro y precio de energía eléctrica para una empresa con capacidad dual de generación eólica y producción de hidrógeno, y que proporciona un método de solución exacto que es notablemente eficiente y adaptable a datos existentes del mercado de energía. Esto nos permite resolver múltiples escenarios para valorar el balance entre dos dimensiones de incentivos monetarios de producción.

La habilidad de combinar las capacidades de generación y almacenaje de energía para maximizar ganancias podría resultar crucial en incrementar la viabilidad del hidrógeno como un vector energético alternativo a gran escala, así como mejorar la utilización de las actuales (y futuras) infraestructuras de generación de energía renovable. Por tal razón, consideramos este trabajo como una herramienta de valor tanto para empresas que desean evaluar sus decisiones de diseño de red, y también para los diseñadores de políticas que deseen estimar sistemas de incentivos equitativos para la producción de energía limpia. Demostramos cómo el modelo puede ser utilizado para estimar el costo monetario (para la agencia reguladora) de las reducciones en emisiones derivadas del uso de combustible limpio en el sector de transporte (vehículos particulares para movilización de personas), así como el costo de asegurar una adopción equitativa del combustible alternativo a lo largo y ancho de un área geográfica por medio de una política de asignación proporcional. Nuestro enfoque no consiste en hacer sugerencias específicas de políticas públicas, sino proveer un marco matemático a los participantes del sector que les permita obtener una representación más precisa de los resultados inducidos por un conjunto de políticas, posibilitando una comparación adecuada con otras tecnologías de almacenaje y transporte de energía.

Existen muchas direcciones interesantes en las cuales este trabajo puede ser extendido. Primero, otras políticas alternativas de distribución pueden ser exploradas para balancear rentabilidad (para el productor), capital y costo social (para la sociedad), y costos de incentivos (para el regulador); por ejemplo, la inclusión de niveles de servicio mínimos y/o diferenciados por ubicación. Dichas políticas podrían afectar la estructura del problema y requerir el uo de metodologías diferentes a las propuestas en esta tesis.

Segundo, nuestro trabajo está basado en un esquema zonal de precios, una suposición que podría ser relajada para considerar precios nodales con posible correlación entre nodos.

En el ámbito de la política energética, la utilización del modelo para un análisis más exhaustivo de los efectos regulatorios y de precios sería de gran valor para entender con mayor detalle el efecto de los esquemas de incentivos cuando los parámetros de mercado (precios, demanda) y tecnología (eficiencia, costos de capacidad) están sujetos a cambio. Dichos cambios podrían deberse a mejoras en la tecnología de electrólisis o de compresión/licuefacción, precios o disponibilidad de agua para electrólisis (un punto relevante en ciertas regiones geográficas, incluyendo España), o cambios a largo plazo en el comportamiento estocástico de los precios de energía causados por la introducción de almacenaje a gran escala en el sistema energético. Esto resultaría en un marco analítico para evaluar la asignación de recursos para el desarrollo de vectores energéticos limpios, incluyendo el desarrollo de tecnologías de producción, infraestructuras, programas de incentivo al consumidor, disponibilidad de insumos, y estrategias alternativas de mitigación de emisiones de carbono.

Una extensión directa de nuestro modelo en el campo de diseño de políticas consiste en incorporar dinámicas de mercados de emisiones como parte de las funciones de utilidad de los participantes del sistema. Si al productor se le permite recibir crédito por las emisiones reducidas debido al desplazamiento de combustibles fósiles de la mezcla de combustibles para transporte, entonces podría reducirse la dependencia en incentivos (subsidios) por parte del regulador para el desarrollo sostenible de una infraestructura de producción-distribución de hidrógeno.

Üçtuğ et al. (2011) llevan a cabo un análisis de viabilidad incorporando el comercio de emisiones de carbono para una planta de producción de hidrógeno utilizando un proceso de reformación de metano. Ellos muestran que el comercio de emisiones puede ser una herramienta financiera efectiva (en cuanto a costos) para plantas de producción de hidrógeno durante sus primeros años de operación, cuando el retorno a la inversión es bajo y el riesgo asumido por los productores es alto. Adicionalmente, los resultados de Aflaki y Netessine (2012) indican que los impuestos sobre el carbono por sí solos podrían desalentar la inversión en capacidad de generación renovable. En ese sentido, sería interesante ver el efecto de integrar hidrógeno (como un producto final, en vez de ser únicamente un mecanismo de almacenaje a corto plazo o para desplazamiento temporal de carga dentro del sistema) en su marco analítico para evaluar si el efecto que detectan se mantiene.

Finalmente, tratamos a las instalaciones como entes aceptadores de precios, que es razonable cuando la congestión en la red no es un factor crítico (por ejemplo, debido a la prevalencia del viento en horarios nocturnos –fuera de las horas pico). Como parte de la meta de una tecnología de almacenaje es incrementar la penetración de energías renovables en los mercados energéticos, es evidente que los efectos de la congestión de la red ameritan consideración en mercados de alta penetración de generación eólica.

Nuestra intuición es que el considerar la ubicación de las instalaciones de generación y la congestión en la red crearía una necesidad de ubicar cierta capacidad de producción cerca de los parques eólicos, pues dichos puntos de almacenaje permitirían desacoplar de manera puntual los sistemas de transmisión eléctrica y distribución de hidrógeno durante los períodos de congestión. Con los parques eólicos usualmente localizados en áreas de baja densidad poblacional, este tipo de producción localizada podría ser rentable para servir nichos de mercado, tales como pequeñas aglomeraciones urbanas cerca de los parques eólicos que generalmente serían dejadas fuera de los planes piloto de combustibles alternativos, creando una válvula de escape para la generación excedente en períodos de mayor demanda, mientras provee un beneficio social a través de la adopción de tecnologías limpias en áreas que de otra forma serían ignoradas debido a su relativamente bajo nivel de demanda.

Otra consecuencia de relajar el supuesto de instalaciones aceptadoras de precios son las dinámicas de retroalimentación (feedback) causadas por un incremento en generación renovable, dada la presencia de tecnologías de almacenaje (sea hidrógeno o algún otro vector), cuando la penetración de renovables es alta. Zhou et al. (2012) resaltan este punto y sugieren analizar el efecto combinado de múltiples parques eólicos e instalaciones de almacenaje sobre los precios de electricidad. El extender nuestro modelo para hacer que los precios de energía (un parámetro que afecta las decisiones de localización y capacidad) sean dependientes en la capacidad total de almacenaje vía hidrógeno involucraría una formulación más compleja para capturar esta dinámica, que no sería factible solucionar utilizando las metodologías sugeridas en esta tesis. Dado que parte del propósito de almacenar energía es incrementar la factibilidad de los proyectos de generación renovable, esta relación sería de interés para expandir el modelo.

La utilidad de nuestro trabajo trasciende el campo de sistemas energéticos. Consideren una empresa tomando una decisión estratégica de producción y distribución asociada a un surtido de bienes que comparten un mismo (potencialmente perecedero) insumo, donde el insumo puede ser vendido directamente a un mercado a un precio que vara de manera incierta, y donde los productos finales pueden ser obtenidos mediante procesos de producción compartidos por subconjuntos de productos. Ese tipo de escenario guarda ciertas similitudes estructurales con nuestro problema. En este contexto, las decisiones de localización y capacidad están relacionadas con la selección de los procesos de transformación; mientras que el problema de asignación de mercados podría relacionarse con cuánta capacidad, si alguna, debe ser asignada a cada producto, con los costos de transporte representando los costos de transformación para cada par producto - proceso, con la polı́tica de selección de mercados actuando como mecanismo de filtración para que le empresa escoja qué productos finales debe producir. El modelo puede ser adaptado para representar los precios de los diferentes productos finales, en contraste con un precio único (como es nuestro caso). En el contexto de diferentes tecnologías de producción para industrias con insumos de alto costo (metales preciosos para semiconductores), volatilidad de precio (procesos con alto consumo de energía) o naturaleza perecedera (insumos agrícolas con productos derivados de mayor durabilidad pero menor valor), un modelo de este tipo ciertamente podría ser aplicable.

Location models applied to the design of a distribution network for hydrogen fuel produced from renewable energy sources

by

Jorge Enrique Barnett Lawton

M.S., Georgia Institute of Technology
B.S., Universidad Tecnológica de Panamá

Submitted to the MIT-Zaragoza International Logistics Program
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Logistics and Supply Chain Management

at the

ZARAGOZA LOGISTICS CENTER

July 2013

Author . . . .
MIT-Zaragoza International Logistics Program
July 8, 2013

Thesis Director . . . .
Mozart Batista de Castro Menezes
Associate Professor of Operations Management
University of Calgary, Calgary, Canada

Thesis Co-director . . . .
David Gonsalvez
Professor of Supply Chain Management
MIT-Zaragoza International Logistics Program

Location models applied to the design of a distribution network for hydrogen fuel produced from renewable energy sources

by

Jorge Enrique Barnett Lawton

Submitted to the MIT-Zaragoza International Logistics Program
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Logistics and Supply Chain Management

Abstract

The depletion of fossil fuel reserves, rising concerns over the potential short- and long-term impact of carbon emissions on society and the environment, and the resulting policies aiming for cleaner energy and reduced fossil-fuel dependance, have all placed greater pressure in increasing generation capacity for new renewables, of which wind power has been one of the fastest developing segments. However, the intermittent and unpredictable nature of these sources, coupled with limitations in current transmission infrastructure, translates into significant amounts of wind energy being curtailed and the full potential of generation capacity not being realized.

Hydrogen has been identified as a clean carrier to store the energy produced from such volatile sources. Additionally, it has been proven to be a feasible zero-emission fuel for transportation, either in combustion systems or through the use of fuel cells. Using hydrogen as a storage mechanism would help absorb fluctuations in the energy supply, and potentially increase the utilization of existing renewable generation capacity.

The development of a hydrogen production and distribution infrastructure is a most relevant issue in evaluating the long-term viability of hydrogen as an energy carrier. This work addresses the decisions faced by a firm with renewable energy generation capacity, when also able to produce hydrogen by means of grid-connected electrolysis. We first focus on the supply chain design decisions for the producer; namely, where to locate the hydrogen production facilities, how much capacity to build at each facility, and how hydrogen should be distributed downstream.

We present a series of optimization models for the design of the production-distribution network, considering the additional profits of hydrogen production for an energy generating firm, incorporating the uncertainty of renewable energy supply and electricity prices. The firm operates in a market regulated by an external agency with the capability of setting hydrogen prices and establishing general service policies for the market under consideration by the producer.

For addressing this problem, we make use of solution procedures based on

column generation. As a way of grounding our work in a practical context, we implement our models in assessing a future hydrogen network in Spain and evaluate the effect of incentives (subsidies) on the firm's decisions, gaining insights on how such incentive schemes can be set in place to make a hydrogen economy viable.

To my parents, Claudia and Jorge

Acknowledgments

To say this journey has been challenging would be a severe understatement. Likewise, to say that this milestone was reached by means of my lone effort would be an unfair assessment of how meaningful and influential the support of my peers and loved ones was throughout this endeavour. Many people –more than these pages can hold– deserve my appreciation for their encouragement or support during the last few years, and I can only express my gratitude to you all.

Still, I would like to take this opportunity to recognize some people who played an important role throughout this journey, and without whom this thesis would be little more than wishful thinking.

First I have to thank my advisor, Mozart Menezes, who was not only an academic mentor, but also a kind voice of reason that made sure to keep me on track whenever I needed it, and encouraged me to give my best whenever my strength was dwindling. I also wish to thank my co-advisor, Jarrod Goentzel, for his support throughout this process, especially during the challenging time I spent at MIT.

I wish to thank Mustafa C̃agri Gürbüz, David Gonsalvez, María Jesús Sáenz, Santiago Kraiselburd and Elena Fernández for the roles they assumed during my time at ZLC. I would also like to extend my gratitude to the rest of the faculty and staff of the Zaragoza Logistics Center.

I'm thankful to my colleagues and friends in Zaragoza, Boston and Panama –especially Gerardo, Lijie, Eirini, Mario, Marios, Roberto M., Roberto P.-F., Emiliano, and Angel– who through their opinions, advice and humor made this journey much more interesting and manageable. My most sincere appreciation goes to Laura who, at a time when it would have been easy to give up, always brought out the best in me. I could not thank you enough for your kindness and friendship. I wish to also thank Lineth, who stood by me during a substantial part of this journey.

Last but not least, I wish to express immense gratitude towards my parents –Claudia and Jorge– and my sisters –Lyra and Lisa– who always believed in me. I hope I made you proud.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The emphasis placed in recent years on the use of renewable energy sources has had a significant effect on the share of investments in energy generation capacity around the world. One of the fastest growing segments in the generation portfolio has been that of wind powered electricity production (GWEC 2008). However, as happens with other renewable power sources, supply presents a significant degree of uncertainty –in the case of wind power, the intermittence of wind patterns. Aggregated, the supply of energy from wind can be predicted to a significant degree; but the absence of an economically-viable, standardized, and massively adopted energy storage mechanism and the limitations on transmission grid capacity causes this variability to translate into short-term shutdown of wind energy generation to avoid the risk of overloading the network, resulting in waste of potential energy supply.

Although regulations exist for minimizing the curtailing of electricity generated from renewable sources (as reference, European Union 2009) by giving dispatch priority to installations using such sources for generation, curtailing of electricity from wind generation is still significant due to network capacity issues and the natural prevalence of wind during off-peak periods. For instance, Rogers et al. (2010) describe wind curtailing practices around the world, indicating that for the specific case of Spain, up until 2009 the majority of curtailments were due to network congestion issues, while since late 2009 the majority were due to wind generation exceeding the minimum load.

Hence, the use of a storable clean energy carrier, such as hydrogen, is critical in markets where a significant portion of the energy generation capacity comes from intermittent renewable sources (Jørgensen and Ropenus 2008).

A potential hydrogen market represents an opportunity for wind farm owners and other renewable electricity producers to obtain additional profits, as hydrogen can act as a load balancing mechanism to mitigate the fluctuations on the supply side by allowing for storage of energy output (U.S. Department of Energy 2009b). More importantly, such a balancing mechanism would lead to better integration of renewable generation capacity to the energy supply portfolio, higher utilization of existing (and improved viability of future) generation infrastructures, and lower dependence on fossil fuels –both as energy generation feedstocks and transportation energy carriers— with its consequent effects on the environment and energy security.

Different alternatives have been evaluated for storing surplus energy; for instance, compressed air energy storage (U.S. Department of Energy 2009a, Denholm and Sioshansi 2009) and high-capacity batteries (Dell and Rand 2001, Wald 2007, Saran et al. 2010). A particular factor creating interest in the development of a hydrogen-based economy is its direct applicability as a substitute for fossil-based fuels in transportation (Ogden 1999, Turner 1999), a sector which as of 2005 was responsible for about two thirds of oil consumption in the United States of America, and 55% of oil consumption worldwide (Hirsch et al. 2005). Furthermore, hydrogen produced from a renewable source represents a sustainable carrier for delivering clean energy to consumers (Wang 2002).

Designing the supply chain for the production and distribution of hydrogen is then an important task in evaluating the feasibility of incorporating this energy carrier into the energy system. This thesis focuses on the decision faced by a firm with the capability of generating electricity from a renewable source, when a hydrogen market exists, providing the firm with the option of building facilities for producing hydrogen using a grid-connected electrolysis process. The firm needs to select the location and size of the electrolysis plants, as well as the markets to be served from each plant.

This work is motivated from a research project funded by the Spanish Ministry of Science and Innovation in direct collaboration with Acciona Energy –one of the world's leading wind power developers and operators— to design a future hydrogen supply network for Spain (Goentzel 2010). In that project some modeling choices were made to employ averages for representing uncertain parameters, with the objective of the project being the design of the production-distribution network that minimized the cost of hydrogen delivered at demand points.

This thesis builds on that work by explicitly incorporating the stochasticity of energy supply and electricity prices to the decision model and addressing the decision problem from a profit maximization viewpoint, attempting to recreate the business conditions faced by firms in the renewable energy sector.

There are four main contributions from our work. First, we incorporate two critical attributes of renewable sources and power systems in the design of supply networks for clean energy carriers: the uncertainty of supply generated by intermittent sources, and the fluctuations of energy prices caused by supply-demand imbalance and the cost profile of generation technologies. Second, we treat the problem from the point of view of an integrated electricity-generating/hydrogen-producing firm, a departure from previous works focusing on hydrogen production as the sole economic activity. This setting aligns with the interest of power generation firms of increasing the effective utilization of existing renewable generation capacity, by means of an energy storage technology that would allow them to time-shift generation and reduce the amount of wind curtailment. Third, we frame the problem in the context of profit maximization, and show how focusing only on minimizing transportation costs fails to find an optimal (and in some instances, even a good) solution. To address this shortcoming we propose solution methods based on column generation. We use a static approach for the problem, with our model being a representation of a system operating in steady state over a long time horizon (i.e., when a stable demand for hydrogen exists in the market due to wider adoption of hydrogen-powered internal combustion engine or fuel-cell vehicles, and the process yield of the technology used for production –including compression and/or liquefaction— is also stable). Finally, we aim to provide policy makers with some understanding of the effect of regulatory decisions, including hydrogen prices and subsidies for producers, on the producer's profitability. Providing the right incentives or conditions for making hydrogen production financially viable on a large scale is an important aspect in properly designing the supply system, as this would enable the right amount of investment to be placed on production and distribution infrastructure, as well as increasing the potential of expanding renewable generation capacity.

1.1 Research Objectives

The main objective of this thesis is to create a model that provides a decision maker with answers to the following questions:

- Given the capability of producing hydrogen using renewable energy, how should the hydrogen production-distribution network be designed? Specifically, how many electrolysis plants are required, where should they be located, how much capacity should they have, and which demand locations are to be served from each facility?

- What is the monetary gain from introduction hydrogen production capability to a firm that already obtains revenues from the sale of electricity to an energy market?

- What is the effect on firm profits and network structure for different levels of stable hydrogen prices set by an external regulatory agency? How do profits and network structure change when the regulator attempts to incentivize the firm by assuming a portion of facility-associated costs?

- What is the effect on firm profits when the regulator attempts to enforce spatial (geographic) coverage of demand locations by means of a strict distribution policy? Consequently, what would be the cost of implementing such a distribution policy for the regulator, if the same level of production is wished to be maintained?

- What is the cost for the regulator (and, consequently, for society) of achieving a certain level of environmental benefits derived from reducing fossil-fuel emissions?

Aside from the direct answers that can be obtained from the model, there is a more general purpose we hope to achieve with this work. We believe such a modeling framework can be used, by stakeholders from the public and private sectors alike, for assessing the long-term potential of a hydrogen economy and comparing this potential to that of other energy storage and emission reduction/control technologies. We consider the integration of supply chain design and policy issues to be an important element in understanding the tradeoffs associated to the introduction of future energy technologies, and an element that can aid in creating policies which align the interests of industry with those of society.

We point out that it is not an objective of this thesis to perform a comprehensive numerical study of every possible scenario that could arise in an uncertain and ever-evolving energy landscape. It is our goal to provide a quantitative approach to address this particular problem, and understand (on a methodological level) the implications and challenges of incorporating uncertainty to the analysis of energy storage technologies, while also gaining some general insights into the properties of this complex decision problem. As greater knowledge is gathered on the values of the parameters associated to hydrogen production technology, the tool would represent a foundation which could then be adapted to perform such numerical studies.

1.2 Literature Review

From a methodological point of view, this thesis is set in the broad field of location analysis and supply chain design. From an application perspective, it incorporates technical and economic aspects of energy systems to represent the market and operational conditions faced by decision makers in our problem.

This section will present a summary of the extant literature to which we can relate our problem, based on three main intersecting themes: energy economics, location models, and hydrogen network design. In particular, the last theme will highlight previous research addressing the design of hydrogen distribution systems, but differing in the specific settings and assumptions on which the problem is based.

1.2.1 Energy economics

Some works in this field have relevance in motivating and establishing a proper context for our work. One stream of literature addresses the economic implications of further integrating renewable sources to the energy system. Owen (2004) and Owen (2006) find that renewable technologies can be competitive with generation from fossil feedstocks if the estimates of environmental damage from the combustion of these fossil fuels is internalized into the price of the resulting electricity. Welch and Venkateswaran (2009) address the environmental and financial sustainability of wind energy based on trends of improving generation technology and increasing costs of fossil fuels.

Aflaki and Netessine (2012) study the relationship between electricity market liberalization, carbon taxes and intermittency of renewables, and suggest that efforts in reducing source intermittency (by means of a storage technology) might prove more effective in incentivizing capacity investments in renewable generation than just the use of carbon taxes. The main challenge in addressing the natural intermittency of renewable sources (and, consequently, one of the main roadblocks for a more rapid growth of renewable generation capacity) is the lack of a large-scale, cost-efficient storage technology.

Although a large stream of literature focuses on the technical aspects of energy storage technologies, we will only highlight those works that shed a light on the strategic implications of incorporating these storage technologies to the existing energy system. For a review of the technical characteristics of different storage technologies, the reader may refer to the concise account by Ibrahim et al. (2008) or a more comprehensive reference in Huggins (2010).

Significant research has been done evaluating the management of storage for wind energy under particular settings (see: Black and Strbac 2007, Denholm and Sioshansi 2009, Sioshansi 2010, Zhou et al. 2012). In particular, Zhou et al. (2012) assess the effect of storage in the monetary value of a revenue-maximizing wind farm, caused by reducing wind curtailment and time-shifting generation, pointing out that increased storage capacity might reduce the average amount of wind energy sold when transmission capacity is abundant. Kim and Powell (2011) address the advance energy commitment problem faced by wind farms in the presence of finite storage capacity. Although their work is not in the context of network design, their model does share with ours the consideration of stochastic prices and supply.

Finally, Sundararagavan and Baker (2012) perform a comparative analysis of different types of storage systems, including niquel-cadium and lithium-ion batteries, flywheels, pumped hydro, and compressed air energy storage. The authors evaluate the costs of these different technologies, and proceed to identify the characteristics that affect whether they can become viable. Notably, hydrogen is not one of the alternatives considered.[1] We should note that the paper mentioned focuses on storage for the purpose of load shifting, delaying the delivery of renewable energy from off-peak periods (where most generation naturally occurs) to peak demand periods; meaning that any hydrogen produced solely for that purpose would require reconversion to electricity for delivery, a process that is currently not viable from the point of view of energy efficiency losses. This setting is then different

[1] The author of this thesis had a direct conversation with the second author of that paper at the INFORMS Annual Meeting on October 14, 2012. When asked about this omission, the response was that their initial analysis resulted in a very high cost for hydrogen given the state of production technology.

from ours, as we evaluate hydrogen as an end-product, but we choose to mention it given some materials .

Given our consideration of uncertain energy prices as part of the model, we refer the reader to the extensive literature in energy pricing and its relationship to network congestion, supply and demand (Rivier and Pérez-Arriaga 1993, Stoft 2002, Barquín 2006).

1.2.2 Location analysis and supply chain design

The existing work in location analysis is quite extensive and has evolved significantly over the years. For a comprehensive overview of discrete location models, the reader may refer to Labbé et al. (1995), Daskin (1995) and Drezner and Hamacher (2004).

Over time, many authors have reviewed the extant literature for specific streams of location science research. Hale and Moberg (2003) provide a broad account of research in the field since its origins. Klose and Drexl (2005) review the literature in the context of distribution network design, focusing on continuous location models, network location models, mixed-integer programming models, and applications. Snyder (2006) focuses on facility location under uncertainty, differentiating between stochastic and robust location problems. A more recent review by Melo et al. (2009) explores the literature in the context of supply chain management, comparing models in terms of the supply chain decisions (apart from those pertaining to location and allocation) incorporated to the models.

We highlight the relevance of previous research incorporating uncertainty to location models (e.g., Mirchandani 1980, Weaver and Church 1983, Louveaux 1986, Berman et al. 2007, to name a few). The chapter by Berman and Krass (2002) provides in-depth coverage of stochastic location models. The seminal works of Hakimi (1964, 1965) play an important role in establishing the discrete nature of the solution space, enabling the use of the solution techniques shown in the thesis. Finally, given the combinatorial nature of location problems, relevant references on networks and combinatorial optimization are fundamental for analyzing the structural properties of our problem (Cook et al. 1997, Schrijver 2003, Larson and Odoni 2007), while specific works in decomposition and column generation techniques (Dantzig and Wolfe 1960, Barnhart et al. 1998, Lübbecke and Desrosiers 2005) form a strong base for the methodology used for solving the models presented. Although here we only mention them in passing, in the methodological section of this thesis we will directly highlight specific results or properties that have appeared in the literature whenever we find it relevant.

1.2.3 Hydrogen system design

Ogden (1999) and Ogden et al. (1999) discuss possible infrastructure configurations for production and delivery of hydrogen, without explicit modeling of the location-allocation problem. Yang and Ogden (2007) focus on determining optimal (minimum cost) hydrogen delivery modes from a large central production plant assuming a centralized production scheme, focusing on the hydrogen transportation portion of the supply chain (from the central plant to a central warehouse, or to a network of demand points). They apply an idealized city model to represent the spatial density of population; however, they do not consider the economic implications of (deterministic) hydrogen prices or (uncertain) energy prices in their distribution model.

Some works explicitly incorporate the location-allocation decisions to their problems. Dagdougui (2012) reviews the extant literature in hydrogen supply chains, highlighting the different approaches used in the planning of hydrogen infrastructure.

Almansoori and Shah (2006) present a snapshot model for the design of a hydrogen supply chain, formulating it as a deterministic mixed-integer linear program (MILP) and applying the model for the case of Great Britain. Almansoori and Shah (2009) extend their first paper by allowing decisions in multiple periods and multiple energy sources for hydrogen production (natural gas, oil, coal, biomass and solar power). Notably, the quantity of available sources is deterministic, which may partially be justified given they do not consider wind power as one of the sources in their model. Kim et al. (2008) introduce demand uncertainty and model a hypothetical Korean hydrogen supply chain. Finally, Huang et al. (2010b) introduce a stochastic dynamic programming model for evaluating the timing and location of hydrogen production sites when hydrogen demand is considered uncertain and present a case study for Northern California. These works focus on hydrogen as the standalone economic activity, and do not incorporate the potential tradeoffs faced by a firm using its own saleable generated electricity for hydrogen production as part of a profit-maximizing model.

Brey et al. (2006) present a multi-objective deterministic optimization model for the gradual rollout of hydrogen as a substitute for existing transportation fuels in Spain. They include a cost-minimization objective, a component for minimizing the deviations of production from the regional preferences for particular sources for hydrogen production, and an environmental equity component in their model to induce that environmental benefits of fuel replacement reach all regions to a certain extent. Their reasoning for this environmental equity criterion is that national targets could be possibly achieved by focusing on a greater adoption rate on a few locations with significant renewable energy potential (thereby focusing the environmental benefits of emission reduction on those locations), while pollution reduction has greater utility in regions with higher initial pollution levels; thus, they wish to model a mechanism that encourages production (and realizes benefits) on a more balanced manner.

Han et al. (2012) propose the design of a hydrogen distribution network incorporating differentiated pricing across regions, by means of a MILP. Note that, in the spirit of Almansoori and Shah (2006), all parameters are deterministic and, although framed as a profit-maximizing model, hydrogen is treated as the sole source of revenues for the firm and not a complementary activity to electricity production.

Ball et al. (2007) develop a model for assessing the deployment of infrastructure to support a future hydrogen-based transport system in Germany through the year 2030. Tzimas et al. (2007) estimate the infrastructure requirements for achieving three different degrees of penetration for hydrogen fuel at the European level. They do so by employing general guidelines for calculating infrastructure requirements, without explicitly modeling the configuration of the (optimal) required network. Their work is meant to provide benchmark figures for the magnitude of investment required to deploy a EU-level hydrogen infrastructure up to the year 2050. De Wolf et al. (2009) focus on the optimal design of a pipeline network for transportation of hydrogen fuel, when the supply points for hydrogen are known; i.e., the locations of production plants are fixed, not part of the firm's decision.

Giannakoudis et al. (2010) use a stochastic annealing algorithm to address the optimization of a power generation system with hydrogen storage and renewable energy sources; however, the problem does not consider spatial optimization of the hydrogen production and distribution network as part of the model, with the role of hydrogen storage in their context being an input for a fuel cell that will feed electrical power back into the system when required.

Parker et al. (2010) present a network design model for the production of hydrogen from biomass (agricultural residues). The formulate the profit-maximization location-allocation problem as a mixed-integer, non-linear program, with different hydrogen prices for the demand points. The complexity of the problem forces them to make some assumptions regarding the viability of pipeline links to reduce the number of binary variables. They apply their methodology to the case of California's Sacramento Valley, using rice straw as feedstock for the biomass process. The difference in structure from our problem comes from their use of deterministic costs for feedstocks.

Finally, it's worth mentioning that the design of supply chains for biofuels can be considered close in nature to our work. Chen and Fan (2012) use a mixed-integer stochastic programming model to address a problem with multiple scenarios of supply and demand, to capture some of the uncertainty associated to feedstock availability, aiming to minimize the total costs of capital investments, feedstock procurement and transportation, and distribution of end-product. Huang et al. (2010a) look at a multistage mixed-integer linear program to minimize the total system cost of building and operating a network of biorefineries under deterministic parameters over a finite planning horizon. In both cases, the authors present numerical case studies for the state of California. Papapostolou et al. (2011) formulate a mixed-integer linear programming model and implement it for the case of Greece. A review of research in biofuel and petroleum-based supply chains can be found in An et al. (2006).

Comparing these previous works, we can observe that the main differentiators of this paper from the existing literature are: the explicit modeling of network design problem for a hydrogen producer, coupled with the integration of stochastic energy prices to the profitability analysis, as part of a profit-maximization model where the supply is also uncertain. All those elements are not presented as part of a single model in any of the works reviewed and, to the best of our knowledge, in the literature. Dagdougui (2012) emphasizes, as part of the author's overview of hydrogen supply chain literature, that the literature has been heavily populated by models minimizing the cost or the environmental impacts of hydrogen supply chains, with fewer studies focusing on risk issues in the design of this infrastructure. Additionally, the author states a need to do further research on hydrogen supply chains operating on clean feedstocks, including renewable energy sources. Thus, we can consider our work to fill a present gap in the body of research in hydrogen supply chains.

We do acknowledge that there are many other elements presented (or at times only mentioned) by other authors that could be incorporated as extensions of our model, such as multi-period decisions with demand growth to simulate adoption. Thus, the literature presented serves not only as comparison for placement of our model in the field, but also as motivator for identifying future research directions where our integrated framework can be enhanced to incorporate other relevant effects. We will identify those research directions in §5.2.

1.3 Structure

The remainder of this work will be divided in four main sections, the scope of each is briefly introduced below.

- **Chapter 2**: We address the supply chain design problem faced by the producer, under a distribution policy that enforces full local coverage at the (potential) expense of not achieving full spatial coverage. In virtue of our problem having a profit-maximizing objective, the producer focuses on trading off the addition of new markets to her coverage area.

- **Chapter 3**: We formulate the problem under a different distribution policy, where the regulator prioritizes full spatial coverage at a uniform service level across the network (which may be less than 100%, indicating full local coverage might not be achieved). The producer's tradeoff, in the context of this policy, shifts towards increasing her presence in all local markets at the same pace.

- **Chapter 4**: We shift our attention towards the effect of the actions of a (well-intentioned) regulatory agency, which can control the price of hydrogen to be received by the producer, as well as set policies on how the producer shall allocate production to satisfy the regulator's plans of adoption of hydrogen technology in his region of jurisdiction. This regulator also has the option of providing a monetary incentive to the producer by either assuring the producer a captive profit from marginal hydrogen prices (i.e., a price subsidy) or by partially covering the costs of building capacity in the network (i.e., an equipment subsidy). We incorporate these incentives in our network model, and use the proposed solution procedures to evaluate the impact of these incentives in the producer's profitability and network configuration. The main purpose is evaluating these incentive systems as part of an energy policy framework, using an actual example of a potential hydrogen distribution network as a basis for evaluation. To offer a practical scenario for this analysis, we present a case study for a potential future hydrogen network in Spain, noting that the model is suitable for any geographic region, independent of size, where energy prices can be treated as those of a single zone (see §2.2 for an explanation of our model assumptions).

- **Chapter 5**: This final section presents a summary of our results and lists potential paths for extending our model to capture other elements of energy system design.

Chapter 2

Decision Under a Market Selection Policy

We start by defining the first distribution policy to be covered in this thesis.

Definition 1. *A* **market selection policy** *consists of the following conditions:*

- *The producer can choose to serve a subset from a finite list of locations, each with an associated hydrogen demand.*

- *If a location is chosen for service, then the firm commits to building the necessary capacity to satisfy the projected demand of each location; i.e., deliberate partial exclusion of a market in the planning stage is not allowed.*

The interpretation for this policy is that full local coverage of hydrogen demand at a given node takes precedence over the spatial coverage of hydrogen distribution at the network level. Intuitively, a market selection policy would –all other elements being the same– be preferred by potential producers over any alternative strategy, as it allows prioritizing the markets of greater revenue potential.

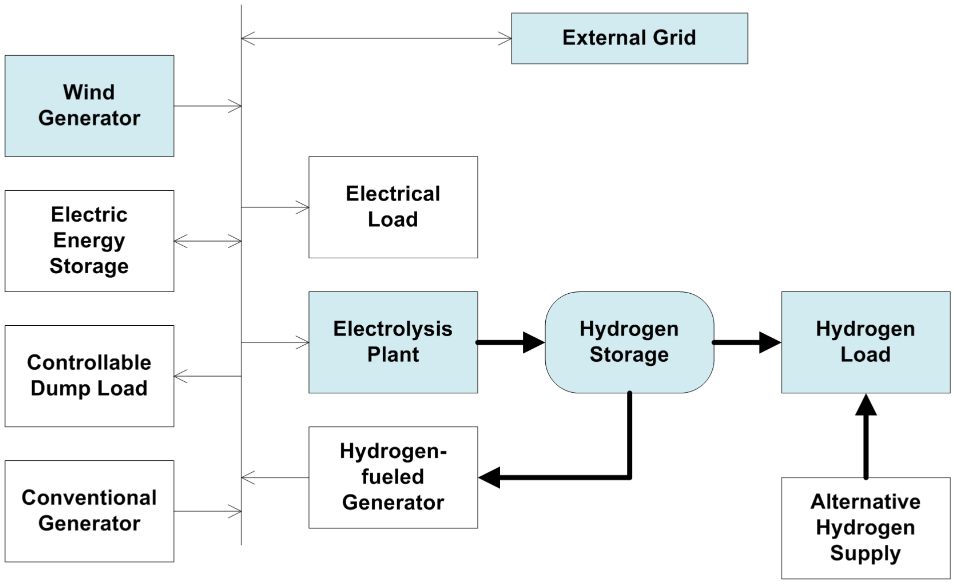


Figure 2-1: Components of a wind-hydrogen energy system (Source: Korpås and Greiner 2008). Shaded boxes represent the components considered in our model.

2.1 Problem Description

In our problem, the production of hydrogen using electricity generated from wind energy creates a market complementary to that of electricity. Hydrogen will be transported to downstream depots and dispatched to final consumers at forecourt stations. A representative wind power generation and hydrogen production system is depicted in Figure 2-1 (Korpås and Greiner 2008), where the components that we include in our decision model are highlighted.

We will focus on the location of the electrolysis plants, which shall have integrated storage capacity, and will assume the firm does not purchase hydrogen from an outside source for serving its customer requirements (represented by the *hydrogen load*). Note there is also the possibility of the producer storing the energy as hydrogen, and then transform it into electricity using a *hydrogen-fueled generator* when electricity prices make it profitable, although this last possibility will not be explicitly considered within the scope of this work. Note the role of a hydrogen-fueled generator in the system would be relevant in the case the market for hydrogen as an end-product is small in comparison to the load-shifting needs of renewable generators, although that would involve even more energy losses (35 to 65%[1]) in the reconversion. Thus, use of hydrogen for storage and reconversion of energy to electricity is far too costly to be a justified alternative (Menanteau et al. 2011). In our problem, we assume a market for hydrogen as an alternative transportation fuel is in place.

Price fluctuations allow the producer to choose the best option at any given time between selling electricity to the grid, or producing hydrogen and selling it to its respective market. We assume energy price fluctuations in the grid to fully characterize the state of the system, which includes the state of the network (capacity and congestion) and market (balance between supply and demand) at any given time.

We set our analysis in a market where a potential hydrogen producer is restricted to using renewable sources –such as wind or solar— and a grid-connected electrolysis process for hydrogen production; i.e., purchasing energy proceeding from alternate sources for hydrogen production is not allowed.

The main reasoning for such a setting evolves from the clean nature of hydrogen as an energy carrier, which would otherwise be undermined by the potential environmental impact of using electricity generated by means of fossil feedstocks; additionally, this matches one of the main pathways suggested for clean, emission-free hydrogen (Wang 2002).

To enforce this restriction, a contractual mechanism would need to be put in place to certify that energy being retrieved from the grid for hydrogen production has been input by the firm using a renewable source. Though such mechanism is not in place on a broad scale at the moment, we consider its implementation to be practically possible in a reasonable timeframe. We will point out that there are voluntary programs for individual consumers willing to pay premiums for "certified" green energy. In these privately owned programs, the utility receiving payments from consumers will use the premium to purchase Renewable Energy Credits (RECs), the purchase not implying that the premium will be utilized for building clean en-

[1]Values obtained from "Energy storage: Could hydrogen be the answer?" by Nadya Anscombe, as appeared on http://www.solarnovus.com, on June 4, 2012.

ergy infrastructure or cover costs from renewable generation. An overview of such programs is available at U.S. Department of Energy (2012). Note such programs, as well as the Compliance Markets established in the United States of America, do not necessarily enforce that the firm holding certificates actually built renewable infrastructure, but a modification of such programs could be a foundation for the accountability requirements of a clean hydrogen market.

2.2 Modeling Framework

2.2.1 Network topology

Let $G = (N, A)$ be a connected graph representing the electric transmission system (from now on, the grid or network), where N represents the finite set of demand nodes, with $N = \{1, \ldots, n\}$. The set of edges A represents transportation paths between pairs of demand nodes. A facility can be located anywhere in G. We later show that, without loss of generality, the set of potential locations for facilities can be reduced to the set of nodes N.

2.2.2 Action choices

The firm makes use of a renewable source that provides an uncertain quantity of energy, characterized by a random variable R, with each unit of energy generated having a cost w. The firm has knowledge of the current market price for electricity φ (a realization of a random variable representing electricity price, p^E). The firm can then decide whether the actual amount of energy supplied by the source (r, a realization of R) shall be sold to the grid at unit price φ or used for producing hydrogen.

This second option requires a transformation process (electrolysis) at a facility located at $j \in G$, which yields e units of hydrogen per unit of energy introduced, and has a processing cost m per unit of hydrogen output. Each resulting unit of hydrogen can then be transported to a demand node i at a cost t_{ij}, and be sold to a downstream buyer at a price p^H. These downstream buyers, for the purpose of our analysis, are depots located at the ”city edge” that will then be responsible for last-mile delivery to filling stations. This convention has been used in prior studies to decouple the main distribution system from local logistics. A given demand point $i \in N$ requires an amount D_i. The rate of hydrogen production is defined by the capacity (C) of the electrolyzers present at a given facility, which is built a priori by the firm at a cost $Q(C)$ representing the amortization of the investment the firm makes on building and operating this production capacity.

We point out that even when hydrogen is a storable good (and, precisely, its storable nature is its main value proposition as an energy carrier), our problem does not explicitly model the storage decision associated to the production facilities and demand points, both of which require, by design, such storage infrastructure). The need then arises to argue when should inventory be shifted from the production site to the depots located at demand points. In other words, this would involve explicitly addressing ordering policies from downstream depots, and evaluating whether the uncertainty in supply and storage capacity hinder in any way the transfer of hydrogen downstream. We do not consider this portion of the problem in this model, as we attempt to abstract a long-term planning decision, focusing on the location and distribution components. We can do so under the assumption of (1) the system operating in steady state; (2) there being *sufficient* storage capacity at the production site and demand points, preventing the case of production stoppages due to storage capacity constraints; (3) there are no significant holding costs that affect the choice of dispatch at either storage point, which is reasonable considering the investment in storage capacity is fixed and considerably large. In the case of constant hydrogen demand, which is the setting we have explored, these assumptions can be supported.

Likewise, in the initial approach with the firm collaborating in the research project that motivated this work, the assumption was that any safety stock necessary to hedge against shortages would be held downstream at demand points, outside of the scope of the producing firm. Even if there were considerations for the effect of supply uncertainty on possible inventory shortfalls or capacity constraints, this stochasticity would not be an issue in a long-run strategic model, as the variability of consumption is relatively small when aggregated, given the (daily) cyclical nature of wind generation makes it unlikely to have extended periods of non-generation. Thus, the thesis assumes the goal to be meeting demand in expectation. Extensions to the model could attempt to capture these dynamics.

2.2.3 Timeframe for analysis

We present a static model, where we assume the system to be operating in steady state over a long time horizon. In other words, the system has reached a point where demand for hydrogen is considerable stable within a time period due to some sustained level of adoption of hydrogen-powered vehicles, and the yield of the production process is sustained. Figure 2-2 presents a general timeline with the different phases as experienced by the producer, and the decision timeframe on which our problem is based. The timing of the investment decision is indicated, while operational decisions (when to produce hydrogen and/or sell electricity to the grid) taken upon observing realization of random system elements, are taken during the phase indicated as *operating period*. The timing is similar in these terms to Aflaki and Netessine (2012).

Expected profits for the firm are calculated for an arbitrary period of analysis (i.e., units are normalized to correspond to the length of that period) within this long-term timeframe, the only requirement being that supply, production capacity, and amortized facility costs are consistently defined for that same time period. Note that market prices for electricity are usually updated, depending on the market, in intervals ranging from five minutes to one hour; thus it would be reasonable that the period of analysis shall be consistent with the period of price updating.

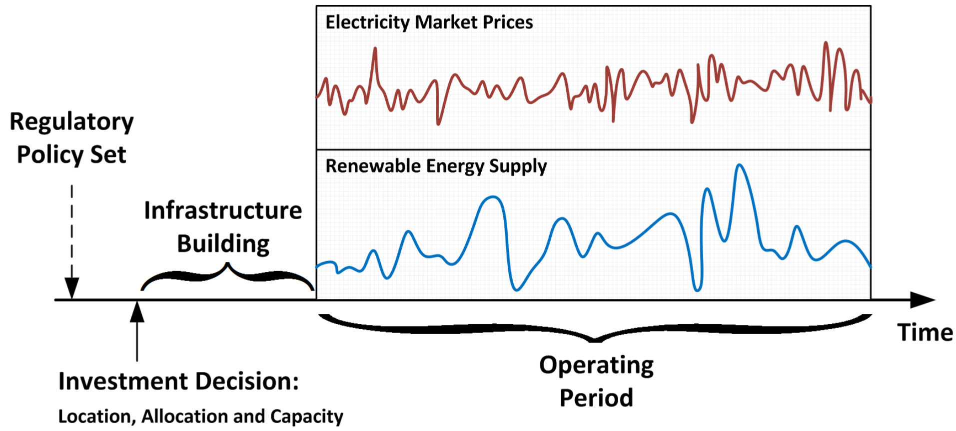


Figure 2-2: Producer's decision timeline

2.2.4 Pricing of goods

Hydrogen price is assumed to be constant, as in a regulated market. We consider this a valid assumption, at least in the initial stages of the hydrogen economy, substantiated in the need to ensure sufficient and stable output and accelerate technology adoption either through subsidies or enforcement of environmental policy (Rogner 1998).

A reasonable question is whether wind energy output is capable of affecting energy prices, and in that case, whether the supply at a given node is correlated with energy prices at that or any other node in the grid. For this thesis we will assume that any single renewable energy generation facility (e.g., wind farm) and hydrogen production plant is a price-taking facility, having no considerable unilateral effect on electricity prices or market structure (Jørgensen and Ropenus 2008, Greiner et al. 2008), which allows us to treat. This also implies that no single generation facility can fully congest all possible paths between two nodes as this would provide it with price-setting capability which would contradict our assumption (Barquín 2006). A consequence of the price-taking assumption is that electricity prices and wind-based supply as independent (Kim and Powell 2011). We will focus on a network with identical electricity price across all nodes; i.e., a single zone in a zonal pricing scheme, which is the case in Spain.

2.2.5 Distribution of supply

For the purpose of this thesis, we do not make any prior assumptions on the probability distribution of energy supply from the renewable source. We make, however, a brief remark on the properties of such distribution for the benefit of the reader.

Many different factors aside from local wind speeds and patterns influence the likelihood of a certain energy output from a wind farm. Figure 2-3 presents some of these main factors.

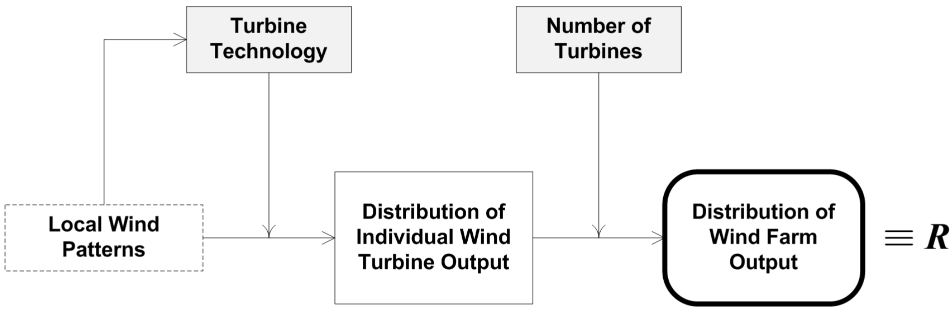


Figure 2-3: Factors directly inciding in the energy output from a wind farm

In general, wind speeds at a certain location can be approximated by the use of a two parameter Weibull distribution (generally using a shape factor of 2), which is a widespread and well accepted practice in the wind energy research literature (Carta et al. 2009). This input is used by firms to choose the proper turbines. An individual turbine's output as a function of wind speed can be described by an output curve like that shown in Figure 2-4.

Output only fluctuates between the *cut in* and *cut out* wind speeds (the speeds below and above which the turbine will not generate any electricity, either because there is not sufficient input or because speeds are too high to sustain generation causing turbine shutdown). These speeds are by design given for any specific type of turbine. Above a certain wind speed, output approaches 100% of the rated output value.

The probability distribution of power output from a single turbine at any given time is obtained by using the output curve to transform the probability distribution

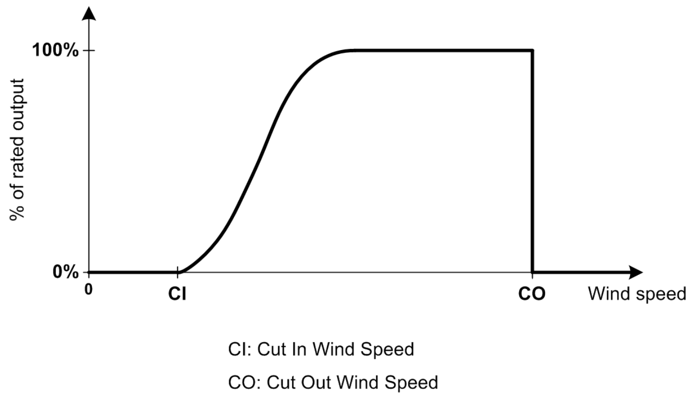


Figure 2-4: Output from single turbine as function of wind speed (Source: Boyle 2004)

of wind speeds into their equivalent expected output values. The aggregated power output for a wind farm is then obtained through aggregation of the individual outputs. Power values can be mapped into energy values by using an arbitrary time interval (e.g., hours). The energy generation curve and the power output curves will be affinely shaped. The probability distribution of power output from any number of wind turbines, when obtained from a generic output-speed curve from two-parameter Weibull distributed wind speeds with shape parameter 2, has a single mode. Hence we can infer R to have a unimodal distribution.

Although this property does not play a role in our current model, it can become relevant in extensions where structural properties are held only for specific families of distributions, an issue we discuss briefly in §5.2.

2.2.6 Hydrogen demand

Aggregate hydrogen demand at a location for use in transportation may reasonably be expected to present a behavior similar to that of current automotive fuels. Such behavior may be assumed as stable and highly predictable, conditioned on the absence of external system disruptions that can induce unforeseeable variability Ganslandt and Norbäck (2004). Extensive studies on gasoline demand indicate it is quite inelastic in the short run (Hughes et al. 2008), although demand does present a significant seasonal/cyclical nature (Menanteau et al. 2011).

On the other hand, demand for fuel substitutes such as ethanol can be highly sensitive to relative price fluctuations with respect to gasoline (i.e., in the variation of the *differential* between ethanol and gas prices), but we should note that substitution is easy between these two fuels, playing a role in the higher price-responsiveness of ethanol (Anderson 2012). Thus, a price-regulated market for a gasoline substitute would mitigate price-elasticity issues, especially given the specificity of vehicles when it comes to fuel use that restricts substitution. The seasonality issue, however, remains unadressed in our model, and would need to be captured in a multi-period formulation which could form the basis for future work.

Finally, we are considering the situation of a single firm generating electricity and producing hydrogen, with no local competition in its markets of operation. Certainly, in the case of multiple firms operating in the same markets, the problem would take a different structure. However, it is likely that in the initial stages of hydrogen fuel adoption, when demand is still at relatively low levels, potential long-term risk of companies entering the market will induce the government to take a more active role in establishing a market structure that enables firms to operate without the added risk of competition (Bento 2008, Wang and Wang 2010).

2.2.7 Capacity cost function

We use hydrogen production capacity as a proxy for facility size, and treat storage capacity as linked with production capacity. We assume there exists a mapping $f : C \to Q$ between production capacity and facility costs. For a single technology configuration (i.e., specific combination of different electrolyzers), facility costs have two distinct components.

The first is a fixed cost that depends solely on the location of the facility, regardless of its size, and may include operating permits, grid connection rights, land acquisition, and site overhead expenses.

The second component is a variable cost, which depends on production capacity and may include operation and maintenance of electrolyzers, liquefiers, storage facilities and flow systems. Variable costs are affected by the system's scale and are assumed to be convex increasing in capacity, an assumption based on the behavior of some manufacturing systems when subjected to increasing utilization rates.

We should point out that at the current state of technology, electrolyzer size is such that, for utility-level deployment of hydrogen infrastructure, increases in capacity would likely be achieved by the addition of (relatively small) electrolyzer units. This would be approximated in a large-scale planning model as a linear function. The inclusion of an alternative with economies of scale is feasible in the current model, as the solution methodology proposed in the case of the market selection policy is unaffected by the shape of the capacity cost function. Since (as will be presented in §3) this procedure can be suitably adapted for the proportional allocation case, then both problems can be solved numerically independently of the shape of $Q(C)$.

The capacity decision is made at time zero by the producer and we assume it cannot be reversed; in fact, it could be done at considerable cost, but we will not consider this possibility. Table 2.1 summarizes the parameters used throughout this thesis, while Figure 2-5 illustrates the association of the model parameters with the components of the wind-hydrogen system.

2.3 The Producer's Problem

Based on the hydrogen price set by the regulator, the producer faces the problem of choosing the locations for the electrolysis plants, how much capacity to build at each site, and deciding which markets to serve from each plant.

Let $\rho_i = \frac{D_i}{\sum_{k \in N} D_k}$ be the proportion of demand corresponding to node $i \in N$, $S_k \subseteq N$ a (fixed) subset of demand nodes in the network served by a single facility, and $\rho_{S_k} = \sum_{i \in S_k} \rho_i$ the proportion of demand corresponding to subset S_k. Then,

$$\bar{t}_{j|S_k} = \frac{\sum_{i \in S_k} \rho_i t_{ij}}{\rho_{S_k}}$$

R	Supply of renewable energy (kWh). Generally distributed random variable, with density $f_R(\cdot)$ and full support over $[0, R_{max}]$, where $R_{max} \in \mathbb{R}$ represents maximum generation capacity at the source.
D_i	Demand rate for hydrogen at node i in kilograms.
p^E	Price of electricity ($/kWh). Generally distributed random variable with density function $f_{p^E}(\cdot)$.
p^H	Wholesale price of hydrogen at demand node i ($/kg). Deterministic.
w	Unit variable cost of electricity generation at source ($/kWh). Deterministic.
e	Hydrogen production efficiency, as units of hydrogen output per unit of energy input (kg/kWh). Deterministic.
m	Unit variable cost of hydrogen production ($/kg). Deterministic, and the same for all potential locations.
t_{ij}	Unit variable cost of hydrogen transport to demand site i from location j ($/kg). Assumed constant, thus treating transportation costs as linear in flow.
$Q(\cdot)$	Amortized facility costs ($). Assumed the same for all locations.

Table 2.1: Summary of model parameters.

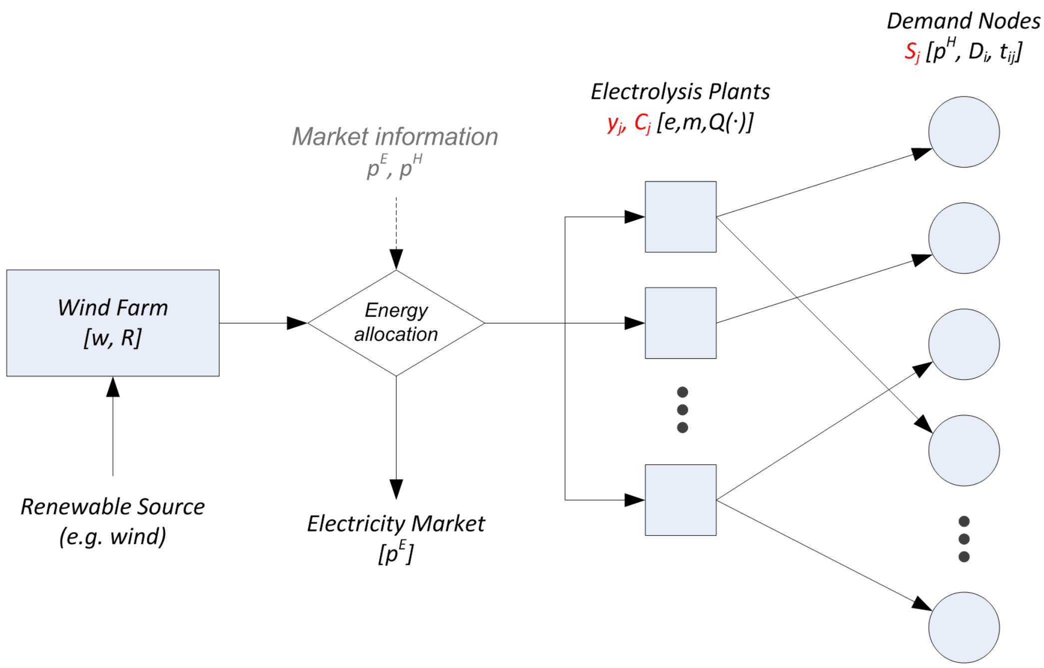


Figure 2-5: Illustration of problem setting, including model parameters.

is the weighted average transportation cost per unit of hydrogen from location $j \in S_k$ to such a subset of demand nodes.

Effectively, every unit of hydrogen produced at and distributed from location j to some portion of the network will translate into a constant transportation cost $\bar{t}_{j|S_k}$; thus, any subset of demand nodes $S_k \subseteq N$ being supplied from a single facility at j can be treated as a single entity, with demand $D_{S_k} = \sum_{i \in S_k} D_i$ and associated transportation cost $\bar{t}_{j|S_k}$. For the remainder of this section, we slightly abuse notation by using $\bar{t}_j$ for expressing such transportation cost for a facility at j, whenever the target subset S_k is implied as fixed. Given the allocation decision is done at an earlier time than operational decisions (i.e., decisions associated to individual realizations of the random variables), then such a notational choice does not have any effect on the understanding of the problem.

At any given time, the producer compares the profit rate of selling a kilowatt-hour (kWh) of electricity from the source to the grid (π^E), with that of using a kWh of electricity to produce and deliver hydrogen from a location j (π^H_j). We define:

$$\pi^E = \varphi - w$$
$$\pi^H_j(\bar{t}_j) = e(p^H - m) - w - \bar{t}_j$$

For adequate comparison, π^E and $\pi^H_j(\bar{t}_j)$ are in equivalent units (\$/kWh). Clearly, if $\pi^E \geq \pi^H_j(\bar{t}_j)$ then it is more profitable for the producer to sell electricity to the grid than producing hydrogen at location $j \in G$.

The relation of the profit rates with the zero profit benchmark yields six scenarios, of which only those where $\pi^H_j \geq 0$ will be feasible for firms operating in this setting (see Figure 2-6), given the condition that hydrogen production using the firm's own electricity generation must be profitable in order to justify any electrolysis capacity investment.

	Scenario 1	$\pi^H \geq 0 > \pi^E$
Feasible scenarios	***Scenario 2***	$\pi^E \geq \pi^H \geq 0$
	Scenario 3	$\pi^H > \pi^E \geq 0$

Figure 2-6: Feasible profitability scenarios for the producer.

2.3.1 Analysis of profitability scenarios

We proceed to analyze each scenario separately, leading to the aggregated (expected) profit function for the producer. Within this section we look only at a particular action choice given a pair of realizations of energy prices and supply, when production is done at a particular facility j. An implicit assumption here is that the firm acts rationally while facing each market situation, meaning no external factor (e.g., a competing firm, a conflicting strategic objective) is present that will cause the firm to deviate from that particular action.

We introduce the following notation to be used throughout. Let X be a generally distributed random variable with continuous and differentiable probability density function $f_X(x)$. Let $a, b \in D_X$, and the following functions be defined:

$$\begin{aligned} F_X(a) &= \int_{-\infty}^{a} f_X(x)dx \\ F_X^c(a) &= 1 - F_X(a) = \int_{a}^{\infty} f_X(x)dx \\ G_X(a) &\equiv \int_{-\infty}^{a} x f_X(x)dx \\ H_X(a,b) &\equiv F_X(b) - F_X(a). \end{aligned}$$

Note $G_X(a)$ is nondecreasing in a and converges to $E[X]$ as $a \to \infty$. Also, $\bar{R}$ and $\bar{p}^E$ will represent the expected value of energy supply and prices, respectively. We also define $\beta_j(\bar{t}_j) = p^H - m - \frac{w}{e} - \bar{t}_j$ as the marginal profit per kg of hydrogen produced at location j and delivered, and $\bar{\Pi}_k$ as the contribution to expected profits from scenario $k = \{1, 2, 3\}$.

Scenario 1: $\pi^H \geq 0 > \pi^E$

In this scenario, hydrogen production is the only profitable alternative for the firm, hence it will produce as much hydrogen as plant capacity or electricity supply allows. In the case electricity supply exceeds hydrogen plant capacity, the firm has no alternative profitable use for the excess electricity resulting, in the absence of alternative storage mechanisms, in a loss of the potential supply. The profitability conditions lead to the following:

$$\begin{aligned} \pi^E < 0 \quad &\Rightarrow \quad \varphi < w \\ \pi^H > \pi^E \quad &\Rightarrow \quad \varphi < e(p^H - m - \bar{t}_j) \end{aligned}$$

Notice the first expression imposes a stronger condition than the second on electricity price. Hence $p^E < w$ sets the general bound on electricity prices for this scenario. We denote the case where the firm chooses to only produce hydrogen as *Case 1*. The amount of hydrogen to be produced by the firm will be $min\{er, C\}$. Then, the contribution to total expected profits from this case is:

$$\begin{aligned} \bar{\Pi}_1 &= \beta_j(\bar{t}_j)\, \mathbb{E}_R[\min\{eR, C_j\} \int_{-\infty}^{w} f_{p^E}(y)dy] \\ &= \beta_j(\bar{t}_j) F_{p^E}(w) \left[C_j F_R^c(C_j/e) + e\, G_R(C_j/e) \right] \end{aligned} \tag{2.1}$$

Scenario 2: $\pi^E \geq \pi^H \geq 0$

Here, generating electricity and selling it to the grid is more profitable than hydrogen production. Hence, when facing this scenario and in the absence of external constraints enforcing some minimum production requirement, the firm will resort only to the former activity. From the profitability conditions we have:

$$\begin{aligned} \pi^E \geq 0 \quad &\Rightarrow \quad \varphi \geq w \\ \pi^E \geq \pi^H \geq 0 \quad &\Rightarrow \quad \varphi \geq e(p^H - m - \bar{t}_j) \end{aligned}$$

As $\pi^H \geq 0$, then $e(p^H - m - \bar{t}_j) \geq w$ and $\varphi \geq e(p^H - m - \bar{t}_j)$ becomes the tighter price condition . For convenience, define $\delta_j(\bar{t}_j) = e(p^H - m - \bar{t}_j)$. We denote the occurrence of this scenario as *Case 2*, and estimate its contribution to expected profits as:

$$\begin{aligned}\bar{\Pi}_2 &= \int_0^\infty \int_{\delta_j(\bar{t}_j)}^\infty x\ (y-w)\, f_{p^E}(y) f_R(x) dy\, dx \\ &= \bar{R}\left(\bar{p}^E - w\right) - \bar{R}\left(G_{p^E}(\delta_j(\bar{t}_j)) - wF_{p^E}(\delta_j(\bar{t}_j))\right) \end{aligned} \tag{2.2}$$

Scenario 3: $\pi^H > \pi^E \geq 0$

With hydrogen production the most profitable alternative, the firm will focus on that activity first, as long as capacity allows and there is sufficient energy supply available. However, if supply exceeds the equivalent production capacity, excess electricity can be sold to the grid, as it is also profitable to do so. From the profitability conditions we have $\varphi \geq w$ and $\varphi < e(p^H - m - \bar{t}_j) = \delta_j(\bar{t}_j)$. Thus, this scenario is defined for the nodal prices in the interval $w \leq \varphi < \delta_j(\bar{t}_j)$.

The sale of excess electricity to the grid is subject to the additional condition $r \geq C_j/e$. We denote this instance as *Case 3a*, while the instance where $r < C_j/e$ is denoted as *Case 3b*. We can then estimate the contribution to total expected profits derived from the occurrence of this scenario as follows:

$$\begin{aligned}\bar{\Pi}_3 &= \beta_j(\bar{t}_j)\, \mathbb{E}_R[\min\{eR, C_j\} \int_w^{\delta_j(\bar{t}_j)} f_{p^E}(y)dy] \\ &\quad + \int_{C_j/e}^\infty \int_w^{\delta_j(\bar{t}_j)} \left(x - C_j/e\right)(y-w)\, f_{p^E}(y) f_R(x) dy\, dx \\ &= \beta_j(\bar{t}_j) H_{p^E}(w, \delta_j(\bar{t}_j)) \left(C_j F_R^c(C_j/e) + e\, G_R(C_j/e)\right) \\ &\quad + \left(\frac{w}{e} H_{p^E}(w, \delta_j(\bar{t}_j)) - \frac{1}{e}\left(G_{p^E}(\delta_j(\bar{t}_j)) - G_{p^E}(w)\right)\right) \\ &\qquad \cdot \left(C_j F_R^c(C_j/e) + eG_R(C_j/e) - e\bar{R}\right)\end{aligned} \tag{2.3}$$

We can see these scenarios create a set of mutually exclusive intervals for p^E and R (Figure 2-7) such that for any realization of these random parameters there is a best course of action by the producer, given its selected production capacity C_j

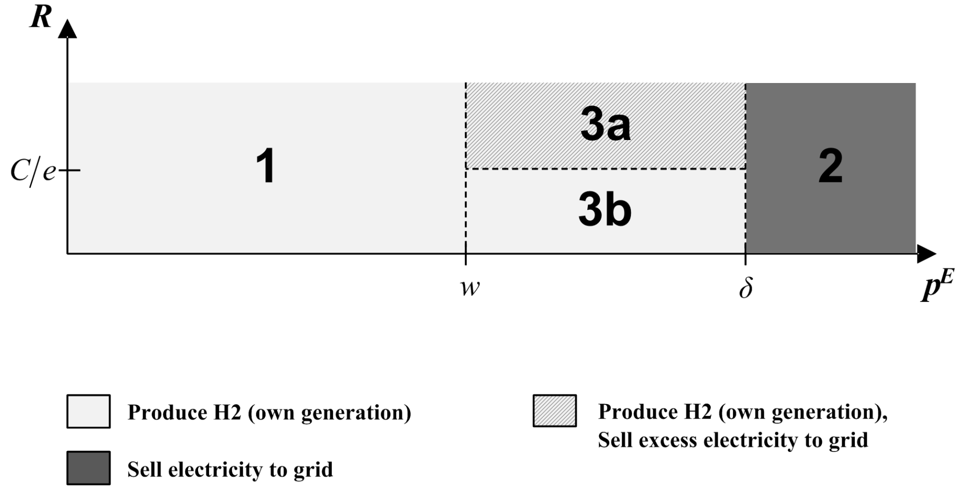


Figure 2-7: Energy price and supply conditions.

(Table 2.2). Here, $\delta_j(\bar{t}_j)$ represents then the minimum electricity price at which the producer prefers selling energy to the grid rather than produce hydrogen at location j and deliver it to its assigned markets.

2.3.2 The producer's expected profit function

By combining expressions (2.1), (2.2) and (2.3), we can obtain the firm's total expected profits. In the absence of hydrogen production capacity, the firm's expected profits from electricity generation can be estimated as

$$K^E = \bar{R}\left(\bar{p}^E - w\right) + \bar{R}wF_{p^E}(w) - \bar{R}G_{p^E}(w).$$

We define $\hat{\Pi}_j = \bar{\Pi}_1 + \bar{\Pi}_2 + \bar{\Pi}_3 - K^E$ as the expected additional profits from hydrogen production (derived from the cases in Figure 2-7) for a firm operating a facility at location $j \in G$ beyond the base profits from electricity generation. In other words, $\hat{\Pi}_j$ represents the contribution of hydrogen to firm profits. We then have:

$$\hat{\Pi}_j = K_j^H(\bar{t}_j)C_jF_R^c(C_j/e) + eK_j^H(\bar{t}_j)G_R(C_j/e) - Q(C_j) \tag{2.4}$$

Profit Scenario	Case	Primary Activity	Secondary Activity	Price conditions	Supply conditions
$\pi^H \geq 0 > \pi^E$	1	Produce H2	None	$\varphi < w$	None
$\pi^E \geq \pi^H \geq 0$	2	Sell electricity	None	$\varphi \geq \delta_j(\bar{t}_j)$	None
$\pi^H > \pi^E \geq 0$	3a	Produce H2	Sell electricity	$w \leq \varphi < \delta_j(\bar{t}_j)$	$r \geq C_j/e$
	3b	Produce H2	None	$w \leq \varphi < \delta_j(\bar{t}_j)$	$r < C_j/e$

Table 2.2: Feasible scenarios, corresponding optimal actions and random variable conditions.

where:

$$K^H(\bar{t}_j) = \beta_j(\bar{t}_j)F_{p^E}(\delta_j(\bar{t}_j)) - \frac{1}{e}\left(G_{p^E}(\delta_j(\bar{t}_j)) - G_{p^E}(w) - wH_{p^E}(w, \delta_j(\bar{t}_j))\right).$$

Here, $K_j^H(\bar{t}_j)$ represents the marginal gains from hydrogen production and delivery; its first term capturing the profits derived from this activity when it is the most profitable choice ($p^E < \delta_j(\bar{t}_j)$), and the second term representing the profits from selling electricity surrendered by the firm in order to produce hydrogen, when both options are profitable but hydrogen production is preferred ($w \leq p^E < \delta_j(\bar{t}_j)$). In general, for the market selection case, $K_j^H(\bar{t}_j)$ will be nonnegative, as the producer will not select a subset with negative marginal hydrogen profits; the same, however, can't be said of the proportional allocation case where the producer might be induced to serve subsets with negative marginal hydrogen contribution in its solution.

By considering only the additional profits attainable from hydrogen production, we can focus our capacity optimization problem on the interval of potential positive capacity values $(0, eR_{max}]$, and disregard the discontinuity caused by fixed costs at $C_j = 0$. Intuitively, the decision for building a given capacity $C_j^* > 0$ will rely on determining whether $\hat{\Pi}_j(C_j^*) \geq 0$.

First and second order differentiation for (2.4) with respect to production capacity yield:

$$\frac{d\hat{\Pi}_j}{dC_j} = K_j^H(\bar{t}_j)F_R^c(C_j/e) - Q'(C_j) \tag{2.5}$$

$$\frac{d^2\hat{\Pi}_j}{dC_j^2} = -\frac{K_j^H(\bar{t}_j)}{e}f_R(C_j/e) - Q''(C_j) \tag{2.6}$$

Thus, we can establish the following result regarding the producer's profit function.

Proposition 1. *The total expected profit function for the producer ($\hat{\Pi}_j$) associated to a facility located at j serving a fixed subset of nodes is concave when capacity* *costs are convex.*

Proof. Follows from expression (2.6), as $K_j^H(\bar{t}_j) \geq 0$, and $Q''(C_j) \geq 0$ for convex capacity costs. $\square$

The following sections will focus on the producer's problem when subject to the alternative distribution policies.

We now address the problem of locating one and multiple facilities in the network when the producer can choose the markets which it will serve. For clarity, we can use (2.4) to define the output associated to a production capacity C_j, given production always occurs when supply is available, as $\bar{H}_j(C_j) = C_j F_R^c(C_j/e) + eG_R(C_j/e)$.

However, no production would occur when the price of electricity is higher than the threshold $\delta_y(\bar{t}_{y|S_y})$. To account for those instances where electricity prices may induce no production even when supply is available (Scenario 2), we adjust the output as follows:

$$\hat{H}_y(C_y, \bar{t}_{y|S_y}) = [C_y F_R^c(C_y/e) + eG_R(C_y/e)]F_{p^E}(\delta_y(\bar{t}_{y|S_y}))$$

2.4 Locating a Single Facility on the Network

The problem of locating a single facility in a network under a market selection policy is defined as follows. Let $S_y \subseteq N$ represent the set of demand nodes to be served from a production facility located at $y \in G$, and $D_{S_y} = \sum_{i \in S_y} D_i$ the aggregated hydrogen demand for such node subset, and $\bar{t}_{y|S_y}$ the associated unit transportation cost of from the facility to its chosen demand nodes.

$$\begin{aligned} Z^{MS}(y, S_y, C_y) = &\underset{y \in G, S_y \subseteq N, C_y \geq 0}{\text{Max}} \; K_y^H(\bar{t}_{y|S_y})\bar{H}_y(C_y) - Q(C_y) \\ &s.t. \quad \bar{H}_y(C_y)F_{p^E}(\delta_y(\bar{t}_{y|S_y})) = D_{S_y} \end{aligned}$$

Note the following properties of total expected (hydrogen) output $\bar{H}_y(C_y)$:

$$\frac{d\bar{H}_y(C_y)}{dC_y} = F_R^c(C_y/e) \geq 0$$
$$\frac{d^2\bar{H}_y(C_y)}{dC_y^2} = -\frac{1}{e} f_R(C_y/e) < 0$$
$$\bar{H}_y(C_y) \to e\bar{R} \quad \text{as} \quad C_y \to eR_{max}.$$

As the distribution of energy supply has full support over $[0, eR_{max}]$, $\bar{H}_y(C_y)$ will be concave increasing on capacity for that same range of capacity values, provided that production always occurs when supply is available. The same holds for $\hat{H}_y(C_y)$ when the threshold $\delta_\cdot$ is fixed; i.e., when a pair (y, S_y) is chosen. Thus, for $D_{S_y} \leq e\bar{R}$ and a fixed pair (y, S_y), there exists a one-to-one mapping between capacity and output (both general output $\bar{H}_y(\cdot)$ and adjusted output $\hat{H}_y(\cdot)$, and C_y can be obtained from numerically solving $\hat{H}_y(C_y) = D_{S_y}$. We will denote this solution as $C_y^*(D_{S_y}, \bar{t}_{y|S_y})$, which makes capacity costs directly dependent on the selected subset S_y, where $C_y^*(D_{S_y}, \bar{t}_{y|S_y})$ is convex increasing in subset demand. This also sets a natural upper bound on market size D_{S_y}, as no production can be achieved beyond $e\bar{R}$. We can rewrite the producer's profit function as follows.

$$\text{(SFMS)} \quad Z^{MS}(y, S_y) = \underset{y \in G, S_y \subseteq N}{\text{Max}} D_{S_y} \frac{K_y^H(\bar{t}_{y|S_y})}{F_{p^E}(\delta_y(\bar{t}_{y|S_y}))} - Q\left(C_y^*(D_{S_y}, \bar{t}_{y|S_y})\right) \tag{2.7}$$

The following property holds for the producer's profit function.

Proposition 2. *Given a fixed location j and demand subset S_j, $Z^{MS}(j, S_j)$ is decreasing in $\bar{t}_{j|S_j}$ if $K_j^H(\bar{t}_{j|S_j}) < \frac{F_{p^E}\left(\delta_j(\bar{t}_{j|S_j})\right)^2}{f_{p^E}\left(\delta_j(\bar{t}_{j|S_j})\right)}$.*

Proof. For a given location j and allocation subset S_j we have expected profits:

$$D_{S_j} \frac{K_j^H(\bar{t}_{j|S_j})}{F_{p^E}(\delta_j(\bar{t}_{j|S_j}))} - Q\left(C_j^*(D_{S_j}, \bar{t}_{j|S_j})\right)$$

The threshold value $\delta_j(\cdot)$ is decreasing in $\bar{t}_{j|S_j}$, so the proportion of time that supply is available and the producer finds itself in the hydrogen production region (scenarios 1, 3a and 3b in Figure 2-7) gets smaller. This means the producer requires greater capacity to achieve the same level of output to meet demand. Thus, $C_j^*(D_{S_j}, \bar{t}_{j|S_j})$ is increasing in $\bar{t}_{j|S_j}$. Since $Q(C_j^*(\cdot,\cdot))$ is increasing in $C_j^*(\cdot,\cdot)$, then $\frac{\partial Q(C_j^*(\cdot,\bar{t}_{j|S_j}))}{\partial \bar{t}_{j|S_j}} > 0$ From the revenue term we obtain:

$$\frac{\partial}{\partial \bar{t}_{j|S_j}} \frac{K_j^H(\bar{t}_{j|S_j})}{F_{p^E}(\delta_j(\bar{t}_{j|S_j}))} = \frac{f_{p^E}(\delta_j(\bar{t}_{j|S_j}))K_j^H(\bar{t}_{j|S_j})}{F_{p^E}(\delta_j(\bar{t}_{j|S_j}))^2} - 1$$

It is then assured $Z^{MS}(\cdot, \bar{t}_{j|S_j})$ is decreasing in $\bar{t}_{j|S_j}$ if $K_j^H(\bar{t}_{j|S_j}) < \frac{F_{p^E}(\delta_j(\bar{t}_{j|S_j}))^2}{f_{p^E}(\delta_j(\bar{t}_{j|S_j}))}$. □

Note the condition described in Proposition 2 is sufficient, but not necessary. Likewise, we can evaluate the second order conditions:

$$\begin{aligned}\frac{\partial^2 Z^{MS}(j, S_j)}{\partial \bar{t}_{j|S_j}^2} = \quad & \frac{f'_{p^E}(\delta_j(\bar{t}_{j|S_j}))K_j^H(\bar{t}_{j|S_j})}{F_{p^E}(\delta_j(\bar{t}_{j|S_j}))^2} - \frac{f_{p^E}(\delta_j(\bar{t}_{j|S_j}))}{F_{p^E}(\delta_j(\bar{t}_{j|S_j}))} \\ & - \frac{2 f_{p^E}(\delta_j(\bar{t}_{j|S_j}))^2 K_j^H(\bar{t}_{j|S_j})}{F_{p^E}(\delta_j(\bar{t}_{j|S_j}))^3} + \frac{\partial^2 Q(C_j^*(\cdot, \bar{t}_{j|S_j}))}{\partial \bar{t}_{j|S_j}^2}\end{aligned}$$

As can be observed, convexity can't be assured for general distributions and sets of parameters. However, it can be verified for particular instances of the problem.

Proposition 2, assuming the conditions for preserving monotonicity are met, implies that for two subsets S_1, S_2 of equal size, the subset with lowest weighted transportation cost will be preferred for yielding higher profits. Because the total size of the market to be served affects expected profits, this preference relationship does not immediately extend to comparing differently sized subsets. However, we can derive a structural property of the (SFMS) that will enable the derivation of solution procedures.

For simplicity, we will incur in some slight abuse of terminology and refer to a node with lower (higher) transportation cost as being closer (farther) to a location j. Thus, $j_{(i)}$ represents the ith unserved demand node closest to j. For a fixed facility at j, we can define $\{S_{j_{(1)}}, \ldots, S_{j_{(n)}}\}$ as the series of subsets formed from the consecutive addition of nodes $j_{(1)}, \ldots, j_{(n)}$ (i.e., $S_{j_{(i)}} = S_{j_{(i-1)}} \bigcup \{j_{(i)}\}$). Recall that the market selection policy requires the nodes selected to be fully served; however, we establish some properties that are relevant for the case where partial fulfillment of a node's demand is allowed, and then adjust this result to incorporate the integrality constraint. In order to capture partial fulfillment we can use variables $x_{ij} \in [0, 1]$ to define the portion of demand of node i that is served from a facility j (with $\mathbf{x}_j^*$ the optimal such vector for a facility j). Note that for the remainder of this section we will assume the monotonicity condition described in Proposition 2 is met.

Lemma 1. *Given two nodes $a, b \in N$ not fully served by j ($x_{aj}, x_{bj} < 1$), where $t_{aj} < t_{bj}$, then the producer obtains greater total expected profits from serving a unit of demand of node a rather than serving a unit of demand of node b.*

Proof. Let j be the potential location for a facility, and a and b represent the two currently unserved nodes with lowest transportation cost from j, with $t_{aj} < t_{bj}$. From the definition of $\bar{t}_{j|S_j}$, we can rewrite the weighted transportation cost of a unit of hydrogen from j to its subset S_j as

$$\bar{t}_{j|S_j} = \bar{t}_{j|S_0} + \frac{D_i}{D_i + D_{S_0}}(t_{ij} - \bar{t}_{j|S_0})$$

where $S_0 = S_j \setminus \{i\}$ is the subset served prior to the inclusion of a node i. Assume $D_i \in \mathbb{Q}$, thus allowing to split any node i into a finite number of nodes of size $\epsilon > 0$ located at the original distance t_{ij}. Let a_ϵ and b_ϵ denote the ϵ-sized nodes derived from a and b, respectively. From (2.7) we can derive the producer's expected profits for a fixed facility located at j:

$$Z_j^{MS}(j, S_j) = \operatorname*{Max}_{S_j \subseteq N} \; D_{S_j} \frac{K_j^H(\bar{t}_{j|S_j})}{F_{p^E}(\delta_j(\bar{t}_{j|S_j}))} - Q\left(C_j^*(D_{S_j}, \bar{t}_{j|S_j})\right)$$

Then, if a node of size ϵ with transportation cost t_{ij} is included in S_j, the resulting expected profits are given by the above expression, with parameters $\bar{t}_{j|S_j} \leftarrow \bar{t}_{j|S_0} + \frac{\epsilon}{D_{S_0}+\epsilon}(t_{ij} - \bar{t}_{j|S_0})$ and $D_{S_j} \leftarrow D_{S_0} + \epsilon$. The producer will be indifferent in choosing between any two nodes a_ϵ, as their equal size and transportation cost yields the same expected profits when included in S_j. Likewise, it will be indifferent between any two nodes b_ϵ. Thus, the only remaining marginal comparison required for establishing the optimal addition of a node to S_j will involve comparing the expected profits of adding either a node a_ϵ or a node b_ϵ. Given both are of equal size, the only differentiating components of the profit function are those factors involving transportation costs, namely $K_j^H(\bar{t}_{j|S_j})/F_{p^E}(\delta_j(\bar{t}_{j|S_j}))$. For clarifying notation, we'll define $\psi_j\left(\bar{t}_{j|S_j}\right) \equiv \frac{K_j^H\left(\bar{t}_{j|S_j}\right)}{F_{p^E}\left(\delta_y(\bar{t}_{j|S_j})\right)}$.

Let $S_a = S_0 \bigcup\{a_\epsilon\}$ and $S_b = S_0 \bigcup\{b_\epsilon\}$. As $t_{a_\epsilon} < t_{b_\epsilon}$ then $\bar{t}_{j|S_a} < t_{j|S_b}$. By Proposition 2 we have $\psi_j\left(\bar{t}_{j|S_a}\right) = \frac{K_j^H\left(\bar{t}_{j|S_a}\right)}{F_{p^E}\left(\delta_j(\bar{t}_{j|S_a})\right)} > \frac{K_j^H\left(\bar{t}_{j|S_b}\right)}{F_{p^E}\left(\delta_j(\bar{t}_{j|S_b})\right)} = \psi_j\left(\bar{t}_{j|S_b}\right)$; thus, the addition of a node a_ϵ to any current subset S_0 provides greater profits than the addition of a node b_ϵ. This property holds for any fixed initial subset S_0, including $S_0 = \{\varnothing\}$. Thus, the first node to be assigned to a facility located at j will be a node a_ϵ.

For the second inclusion, the subset S_0 is updated to $S_0 \bigcup\{a_\epsilon\} = \{a_\epsilon\}$, with associated transportation cost $\bar{t}_{j|S_0} = t_{aj} \geq 0$ and demand $D_{S_0} = \epsilon \geq 0$. Thus, by extension of the previous condition, as long as there is a node a_ϵ available for inclusion, no node b_ϵ will be included in S_j, so the entirety of node a will be served before any portion of node b is served. Note that each subsequent inclusion of the node with smallest transportation cost available will result in equal or greater value for $\bar{t}_{j|S_j}$ (i.e., the weighted transportation cost will be non-decreasing), and $t_{ij} \geq t_{j|S_0}$ as $\bar{t}_{j|S_0}$ results from a weighted average of a set of transportation costs less than or equal to t_{ij}, thus $\psi_j(\bar{t}_{j|S_j})$ will be non-increasing.

Finally, once addition of all nodes a_ϵ has been completed, the tradeoff moves to comparing nodes b_ϵ with the corresponding nodes of size ϵ formed from splitting the next closest unassigned node (e.g., c_ϵ). As $t_{cj} > t_{bj}$, then node b will be included in its entirety before including any portion of node c, yielding the generalized result that the addition of any portion of a node with lower transportation cost still not in a subset is locally optimal given a facility located at j, completing the proof. □

This enables us to establish the following results for the producer's profit function for a fixed location $j \in G$.

Theorem 1. *There exists a node $j_{(k)}$ for which the following conditions are satisfied.*

(1) $Z^{MS}(j, S_{j_{(i)}}) < Z^{MS}(j, S_{j_{(i+1)}})$, for all $i = 1, \ldots, k-1$, and $Z^{MS}(j, S_{j_{(i)}}) > Z^{MS}(j, S_{j_{(i+1)}})$, for all $i = k, \ldots, n-1$.

(2) $\mathbf{x}_j^*$ contains at most one non-integer element. Thus, an optimal subset contains at most one partially served node $j_{(i^*)}$, which will either be $j_{(k)}$ or $j_{(k+1)}$.

(3) $x^*_{j_{(k)}j} = 1, \forall k < i^*$, and $x^*_{j_{(k)}j} = 0, \forall k > i^*$.

Proof. (1) Because of Lemma 1, for any current solution $\mathbf{x}_j \in [0, 1]^n$ only one node i with $x_{ij} < 1$ (the nearest) is considered for inclusion. All other components of $\mathbf{x}_j$ being fixed, weighted transportation costs are increasing in x_{ij}. Consequently, each marginal increase of x_{ij} results in a gradually smaller decrease of $K_j^H(\cdot)$; in other words, $K_j^H(\cdot)$ will be decreasing with respect to x_{ij}.

The expression $Z^{MS}(j, S_{j_{(k)}})$ considers the full addition of node $j_{(k)}$ to the subset served by location j; i.e., $x_{j_{(k)}j} = 1$. Assume a node $j_{(k)}$ is found for which $Z^{MS}(j, S_{j_{(k)}}) < Z^{MS}(j, S_{j_{(k-1)}})$ and $Z^{MS}(j, S_{j_{(k)}}) > Z^{MS}(j, S_{j_{(k+1)}})$. For the first inequality we require marginal profits to be strictly positive. As the subsequent addition of full nodes necessarily increases weighted transportation costs and decreases marginal contribution $K^H(\cdot)$, while it increases capacity costs $Q(\cdot)$, then any full node $j_{(i)}$ included prior to $j_{(k)}$ must have resulted in greater profits than those derived from the previous inclusion $j_{(i)}$, explaining the increasing sequence of total profits for $j_{(i)} : i = 1, \ldots, k-1$. Likewise, if total profits are reduced with the complete addition of $j_{(k+1)}$, then the addition of the node with next highest transportation cost $j_{(k+2)}$ will result in lower total profits. As weighted transportation costs are increasing, any further addition of a node will decrease profits below those of the previous inclusion, explaining the increasing sequence of total profits for $j_{(i)} : i = k, \ldots, n-1$.

(2) For the first part of the proof, assume $0 < x_{ij} < 1$ for some demand node i and facility located at j. A direct result from Lemma 1 is that for any portion of

node i to be the optimal choice for inclusion in the demand served by j, we require that no closer node has any unserved demand, thus $\{l : x_{lj} < 1, t_{lj} < t_{ij}\} = \emptyset$ and $x_{lj} = 1, \forall\ l : t_{lj} < t_{ij}$. By extension of the same property, for any node l farther than i to have $x_{lj} > 0$, this would require $x_{ij} = 1$, which would contradict our initial assumption. Thus, $x_{lj} = 0, \forall l : t_{lj} > t_{ij}$, and only one element of $\mathbf{x}_j^*$ can be non-integer.

For the second part of the proof, note that $j_{(k)}$ is the last node that increases total expected profits when included in its entirety. As each new node included in the service subset has greater or equal transportation costs than the previous included node (and, consequently, than the current weighted average cost) we have that $\bar{t}_{j|\cdot}$ will increase at a steeper rate when the first portion of a new node is included. This means that when the inclusion of node $j_{(k)}$ is completed, if the marginal expected profits $\hat{\Pi}'_j(\cdot)$ are positive, then total expected profits either reach a maximum at $x_{j_{(k)}j} = 1$ or at some partial allocation of node $j_{(k+1)}$ (i.e., $0 < x_{j_{(k+1)}j} < 1$). Likewise, if marginal expected profits at that point are negative, it must be because total expected profits reach a maximum for some partial allocation of node $j_{(k)}$ (i.e., $0 < x_{j_{(k)}} < 1$).

(3) Follows from the proof of part (2). $\square$

Given this result, we can derive a procedure for the formation of service subsets for each candidate location when nodes are not required to be served in their entirety. Since there are infinite potential sites, we first establish the node optimality property (Hakimi 1964) to reduce the set of viable locations. Let y^* be the optimal location of a facility serving a node subset $S \subseteq N$, while $G[S]$ is the graph induced by node subset S. Then,

Proposition 3. *For a service subset $S \subseteq N$ assigned to some facility under a market selection policy,* $y_j = \operatorname{argmin}_{i \in G[S]}\{\bar{t}_{i|S_j}\}$.

Proof. The proof is a result of Proposition 2, resulting in the lowest weighted transportation cost location being optimal for a fixed subset. $\square$

Note that the expression in Proposition 3 is equivalent to determining the 1-median solution to the problem given a subset of nodes to be served. When partial inclusion of nodes is allowed, this property will still hold as any portion of a node can be treated as a node in itself. As Hakimi (1964) established the existence of a nodal solution to the p-median problem (of which the 1-median is a special case), we are guaranteed there exists a location $y^* \in N$ for which $Z^{MS}(y^*, S) \geq Z^{MS}(j, S),\ \forall j \in G[S]$.

Theorem 1 suggests that the linear relaxation of the (SFMS) can be solved by fixing a facility at a given node, sorting the remaining nodes in increasing order of transportation cost, and continuously adding the nearest node until the profit function reaches a maximum which, given the function's unimodality, will be a global optimum (this procedure would be repeated for all n nodes). Thus, this linear relaxation is analogous to a continuous knapsack problem.

The nonlinear integer problem shown in expression (2.7) is, however, NP-hard. To support this claim, we can refer to the definition of the 0-1 knapsack problem (Martello and Toth 1990), with each node i included having a weight D_i and a value v_i, and the knapsack having a capacity $e\bar{R}$. Since v_i is a function of the other nodes in a particular feasible solution, the number of evaluations to guarantee optimality is in the order of 2^n, with a complexity akin to that of a 0-1 knapsack problem, which is NP-complete.

This complexity does not pose an obstacle for small size instances where enumeration is a reasonable alternative; however, larger problems could not reasonably be solved through enumeration. We propose some efficient alternative procedures to address this issue, which will be used for the case when multiple facilities are to be located (i.e., $p > 1$).

2.4.1 Numerical methods

We present two base procedures –nearest-node and greedy inclusion— for obtaining a reasonably good approximation to the optimal service subset for a fixed node.

Given the result presented in Theorem 1 we can expect the nearest-node result to be close to optimality, especially in low-density networks where the difference in added profits from inclusion will be more significant across alternatives (the same effect will also be perceived in the use of the greedy heuristic).

After introducing the two based heuristics, we will also present an extension for exploring improved solutions, which is especially relevant in networks of greater density where the base heuristics may converge to a point significantly far from the true optimal solution. Such an enhancement is based on neighborhood search, and it can be applied to either of the base heuristics.

Nearest-node inclusion heuristic:

Step 0 (Initialization): $S_j^0 = \{\varnothing\}$, $z_j^0 = 0, \forall j = 1, \ldots, n$. Set $j = 1$, $k = 1$.

Step 1 (Node ordering): Sort nodes in ascending order based on transportation cost from node j (t_{ij}); denote the kth lowest-cost node as $j_{(k)}$.

Step 2 (Node inclusion): $S_j^k = S_j^k \bigcup \{j_{(k)}\}$, $z_j^k = Z^{MS}(j, S_j^k)$. If $z_j^k < z_j^{k-1}$ or $k = n$ then proceed to Step 3; else $k \leftarrow k + 1$, repeat Step 2.

Step 3 (Local termination): $z_j^* = \max_k\{z_j^k\}$, $S_j^* = \left\{S_j^k : k = \operatorname{argmax}_k\{z_j^k\}\right\}$. If $j = n$ then proceed to Step 4; else $j \leftarrow j + 1$, go to Step 1.

Step 4 (SFMS solution): Optimal location $y^* = \operatorname{argmax}_{j \in N}\{z_j^*\}$, $S^* = S_y^*$. ■

Greedy inclusion heuristic:

Step 0 (Initialization): $S_j^0 = \{\varnothing\}$, $z_j^0 = 0, \forall j = 1, \ldots, n$. Set $j = 1$.

Step 1 (Facility fixing): $S_j^1 = \{j\}$, $z_j^1 = Z^{MS}(j, \{j\})$. If $z_j^1 < 0$ then proceed to Step 3, otherwise $I \leftarrow N \setminus \{j\}$, $k = 2$.

Step 2 (Node inclusion): While $k \leq n$, find $i^+ = \operatorname{argmax}_{i \in I}\{Z^{MS}(j, S_j^{k-1} \bigcup \{i\})\}$, $S_j^k = S_j^{k-1} \bigcup \{i^+\}$ and $z_j^k = Z^{MS}(j, S_j^k)$; update $k \leftarrow k + 1$ and $I \leftarrow I \setminus \{i^+\}$.

Step 3 (Local termination): $z_j^* = \max_k\{z_j^k\}$, $S_j^* = \left\{S_j^k : k = \operatorname{argmax}_k\{z_j^k\}\right\}$. If $j \leq n$ then proceed to Step 4; else $j \leftarrow j + 1$, go to Step 1.

Step 4 (SFMS solution): Optimal location $y^* = \operatorname{argmax}_{j \in N}\{z_j^*\}$, $S^* = S_{y^*}^*$. ■

Due to the effect of transportation costs on marginal profits from hydrogen production, both heuristics will bear strong resemblance in the construction of solutions for a fixed facility. We can present an enhancement to the nearest-node (alternatively, greedy) heuristic to exploit potential improvements in the neighborhood of the resulting subset.

The algorithm creates a base subset using a feasible solution, such as that obtained by either base heuristic, and explores its neighborhood for an improving direction represented by a node removal, addition, or exchange. At each cycle, the base subset is updated from selecting the neighborhood subset with the highest profits (i.e., in a greedy manner).

Note that this procedure does not guarantee an optimal solution as: (1) only one of (possibly) many increasing paths is chosen at each iteration, and (2) it does not consider the possibility that there is no single-node operation that increases profits but that there are two-node operations that achieve this (we refer the reader to §2.6, where this particular situation arises). However, it does provide a stronger approximation to the optimal subset for a fixed facility without significantly increasing computational cost.

Hybrid heuristic:

Step 0 (Initialization): Set $j = 1, t = 1$.

Step 1 (Base subset): Use nearest-node (or greedy) inclusion heuristic for facility node j; assign solution as base subset B_j, $z_j^0 = Z^{MS}(j, B_j)$.

Step 2 (Neighborhood search): Define the neighborhood $\mathcal{N}_j$ of B_j as the following sets:

- $B_j \setminus \{i\}, \forall i \in B_j$,
- $B_j \bigcup \{k\}, \forall k \notin B_j$, and
- $B_j \setminus \{i\} \bigcup \{k\}, \forall \{i, k\}$ s.t. $i \in B_j, k \notin B_j$

For each subset in $\mathcal{N}_j$, calculate $Z^{MS}(j, \cdot)$, choose set with highest profits S^t_j as the t-degree neighborhood-optimal solution, with profits z^t_j.

Step 3 (Validation): If $z^t_j > z^{t-1}_j$, then $B_j \leftarrow S^t_j$, $t \leftarrow t + 1$, repeat Step 2; otherwise, B_j is nodal solution with profits $z^*_j = z^{t-1}_j$. If $z^*_j < 0$, then $S^*_j = \{\varnothing\}$ (i.e., j is not a feasible location). If $j = n$ then proceed to Step 4; else update $j \leftarrow j + 1$ and return to Step 1.

Step 4 (SFMS solution): Optimal location $y^* = \text{argmax}_{j \in N}\{z^*_j\}$, $S^* = S^*_{y^*}$. ■

The nearest-node and greedy inclusion heuristics require a number of operations in the order of $O(n^2 logn)$ and $O(n^3)$, respectively, for selecting a single location from n nodes.

The hybrid algorithm presented here builds on the result of the nearest-node inclusion procedure, with the main driver of complexity being the formation of the neighborhood sets ($O(n^2)$) and the possible number of cycles before an improving direction is not found. At each cycle, the cardinality of the set increases or decreases by one unit, or stays unchanged. The number of node additions and substractions within a cycle is then bounded by $O(n)$, while the maximum number of swaps is bounded by $O(n^2)$. Since only a single increasing path (if any) is followed at each iteration and no backtracking is allowed, we are assured the procedure will terminate.

We will point out that the hybrid algorithm can use any feasible solution as a starting point; thus, it can be adapted to build on the result of the greedy algorithm if necessary (as will be explained later, this is the case for the branch-and-price procedure used in the multiple facility problem).

Although theoretical worst case performance of the hybrid procedure is significantly inferior to that of the two base heuristics, actual average performance will be considerably better than this upper bound given the relative closeness of the nearest-node solution to the optimal integer allocation, which limits the number of exchanges before converging to a (local) optimum.

In §2.6 we present some numerical results for the three heuristics in order to compare their computational performance. We point out that in all but one of the instances which could be verified by enumeration the hybrid heuristic found the optimal service subset for a fixed facility node.

2.5 Locating Multiple Facilities on the Network

We now address the problem of locating an arbitrary number of facilities under a market selection distribution policy. The general formulation is as follows.

$$
\begin{aligned}
\text{(MFMS)} \quad & Z^{MS}(\mathbf{y}, \mathcal{S}) = \underset{\mathcal{S}, \mathbf{y} \in G}{\text{Max}} \sum_{j=1}^{|\mathcal{S}|} \hat{\Pi}_{y_j}(y_j, S_j) \\
\text{s.t.} \quad & \bigcup_j S_j \subseteq N; \quad S_i \bigcap S_j = \{\emptyset\}, \forall i \neq j \\
& y_j \in S_j, \ \forall j
\end{aligned}
$$

For the special case of an exogenously-defined number of facilities (p), the cardinality condition $|\mathcal{S}| = p$ would need to be incorporated to the formulation. However, our suggested solution method does not require this condition. Note that the constraint set of the original formulation defines a feasible node packing, thus the (MFMS) optimization problem is NP-complete. We can use this structure to derive an exact procedure for the (MFMS) problem based on column generation (Dantzig and Wolfe 1960).

2.5.1 Exact method: Branch-and-price

The column generation form of this problem is equivalent to that of a set packing problem, with the master problem defined as follows:

$$\begin{aligned}
\text{(CGMS)} \quad \underset{\lambda}{\text{Max}} \quad & \sum_k z_k \lambda_k \\
\text{s.t.} \quad & \sum_k \theta_{ik} \lambda_k \leq 1, \quad \forall i = 1, \ldots, n \\
& \lambda_k \in \{0, 1\}, \quad \forall k \in \omega
\end{aligned}$$

where ω is a group of subsets of N, and z_k represents the net profits derived from grouping a subset of demand nodes (S_k) to be served from a single facility, which will be located at y_k (as per Proposition 3), with $z_k = Z^{MS}(y_k, S_k)$. Also, $\theta_{ik} = 1$ if $i \in S_k$ and zero otherwise.

Recall $\psi_j\left(\bar{t}_{j|S_j}\right) \equiv \frac{K_j^H\left(\bar{t}_{j|S_j}\right)}{F_{p^E}\left(\delta_y(\bar{t}_{j|S_j})\right)}$. For a given potential location $j \in N$ we have the following pricing subproblem:

$$\underset{S_j}{\text{Max}} \quad D_{S_j} \frac{K_j^H(\bar{t}_{j|S_j})}{F_{p^E}(\delta_j(\bar{t}_{j|S_j}))} - Q(C_j^*(D_{S_j})) + \sum_{i \in S_j} \mu_i \tag{2.8}$$

or, equivalently,

$$\begin{aligned}
\underset{\mathbf{x}}{\text{Max}} \quad & \sum_{i \in N} \left(D_i \psi_j(\mathbf{x}) - \mu_i\right) x_{ij} - Q(\mathbf{x}) \\
\text{s.t.} \quad & x_{ij} \in \{0, 1\}, \quad \forall i = 1, \ldots, n
\end{aligned}$$

where $\mu_i \leq 0$ is the dual variable associated to the ith convexity constraint from the master problem. Also, $x_{ij} = 1$ if $i \in S_j$, thus weighted average transportation costs can be defined as $\bar{t}_{j|S_j} = \frac{\sum_{i \in N} D_i t_{ij} x_{ij}}{\sum_{i \in N} D_i x_{ij}}$. If all subsets resulting from the solution of the n pricing subproblems yield negative increased profits then we have obtained an optimal solution to (MFMS).

The pricing problem has the same complexity of a fixed-node iteration (i.e., Steps 0–3) of the (SFMS) problem. It is known that the existence of a polynomial time exact algorithm for the pricing subproblem would make the restricted master problem solvable in polynomial time (Lübbecke and Desrosiers 2005). However, we established in §2.4 that this was not the case. We can overcome this issue by using an approximation algorithm for the pricing problem. We refer the reader to Barnhart et al. (1998) for a discussion on that matter. As the hybrid heuristic introduced in §2.4.1 achieves a good practical bound with respect to optimal solutions (see §2.6) it will closely match the column generation pattern that would be obtained from an optimal solution to each pricing problem, while saving significant computation time. If no new entering columns are found through the set of approximated pricing problem solutions, an exact procedure will be required to verify optimality. A branch-and-bound algorithm is suitable for this purpose.

We need to adapt the hybrid heuristic to address the pricing problem through the following changes: use the greedy heuristic solution for defining the base subset B_j, and use (2.8) as the value function. The rationale for basing the algorithm on the greedy solution (rather than the more efficient nearest-node heuristic) is that, when incorporating the dual variables to the problem, the unimodal behavior of the nearest-node policy described in Theorem 1 will not hold in general (save for very particular instances). The adapted greedy heuristic is structurally unaffected by the inclusion of the dual variables, and will approach more rapidly the optimal subset with little added computational cost.

We can now define an exact procedure based on the branch-and-price method described in Barnhart et al. (1998) to solve the multiple facility problem, as follows.

Branch-and-price algorithm for (MFMS):

Step 0 (Initialization): Define initial set of $n+1$ columns ω as all singleton sets and N.

Step 1 (Relocation): For each column $k \in \omega$, locate facility at the 1-median solution, i.e., $y_k = \mathrm{argmin}_{i \in S_k}\{\bar{t}_{i|S_k}\}$; calculate $z_k = D_{S_j}\psi_j(\bar{t}_{t|S_j}) - Q(C_j^*(D_{S_j}))$.

Step 2 (Relaxed problem): Solve linear relaxation of problem (CGMS) for column set ω.

Step 3 (Insertion): Based on dual variables $\mu_i \leq 0$, solve the pricing subproblem for each designated base node j. Resulting nonempty solutions from pricing subproblems are added to set ω, return to Step 1. If no new columns are formed with positive increased profits, then proceed to Step 4.

Step 4 (Optimality validation): Execute exact algorithm for pricing subproblem for each designated base node j. If no new columns exist with positive increased profits, then proceed to Step 5, otherwise add new columns to ω, return to Step 1.

Step 5 (Branching): If $\lambda_k \in \{0, 1\}, \forall k \in \omega$ then λ is optimal solution to (MFMS) and Stop; else choose one element $\lambda_k \in (0, 1)$ for branching, repeat Step 1. ■

A clarification shall be made regarding Step 0 in this algorithm. Given that the market selection policy allows for demand nodes to remain observed, any initial subset of columns will produce a feasible solution by satisfying the packing constraints. Thus, the choice of the initial set of columns is arbitrary, and left to the judgment of the user. In Chapter 4 we present detailed numerical results for the implementation of this algorithm.

2.5.2 Heuristic method for a fixed number of facilities: P-median location with greedy allocation.

Alternatively, a procedure decoupling the location and allocation components of the problem can be used as a heuristic for the multiple facility problem. We exploit the relationship between transportation costs and marginal hydrogen production profits, and use a cost-minimization linear problem to identify a set of reasonably located facilities. Then, we address the allocation component by using the methods described in §2.4.1. Because the allocation for each facility will be determined separately, we need to eliminate infeasible global allocations (those not satisfying the disjoint set condition from the original formulation or, equivalently, the convexity constraints in the column generation form) by removing nodes that are assigned to more than one facility.

The p-median problem (Mirchandani 1990, Daskin 1995) for $G = (N, A)$ can be expressed as

$$\begin{aligned} &\underset{\mathcal{S},\mathbf{y}\subseteq N}{\text{Min}} \sum_j D_{S_j} \bar{t}_{y_j|S_j} && (2.9) \\ &\text{s.t.} \quad \bigcup_j S_j \subseteq N; \quad S_i \bigcap S_j = \{\emptyset\}, \forall i \neq j; \quad |\mathcal{S}| = p \\ &\qquad y_j \in S_j, \ \forall j \end{aligned}$$

Although the general p-median problem is NP-hard, for a fixed value of p the problem is polynomial-time solvable (Garey and Johnson 1979), and there are numerous algorithms available commercially to solve the problem with relative efficiency for reasonably-sized instances. The heuristic is defined as follows.

P-Median based heuristic:

Step 0 (Initialization): Set p.

Step 1 (Location): Solve p-median problem for $G = (N, A)$ and p. Define locations chosen in p-median solution as facility set $\mathcal{F}^p$.

Step 2 (Initial allocation): For each node $j \in \mathcal{F}_p$, use hybrid procedure to solve the allocation portion of the (SFMS), denote these subsets as S_j^p. Define allocation variables $\theta_{ij} = 1$ if $i \in S_j^P$, and 0 otherwise.

Step 3 (Feasibility test): If resulting subsets are disjoint, then allocation is feasible, go to Step 5. Else, determine the set $\mathcal{I}^p$ of infeasible nodes (i.e., find all i such that $\sum_j \theta_{ij} > 1$).

Step 4 (Duplicate reduction): Select node $i \in \mathcal{I}^p$ with largest demand, let $\mathcal{J}_i^p = \{j : \theta_{ij} = 1\}$ be the set of facilities currently serving i. Calculate $loss_{ij} = Z^{MS}(j, S_j^p) - Z^{MS}(j, S_j^p \setminus \{i\})$. Assign node i to subset S_j^p with the maximum value of $loss_{ij}$. For all other nodes $j \in \mathcal{J}_i^p$, update $S_j^p \leftarrow S_j^p \setminus \{i\}$. Remove i from $\mathcal{I}^p$. If $\mathcal{I}^p = \{\emptyset\}$, then proceed to Step 5, otherwise repeat Step 4 for next largest node in

$\mathcal{I}^p$.

Step 5 (Profit estimation): Current allocation $[\theta_{ij}]_{n\times p}$ is feasible. Calculate total expected profits as $z^p = \sum_j Z^{MS}(j, S_j^p)$. ∎

There are two important observations regarding this heuristic. First, the procedure itself has two phases of approximation: (a) the initial allocations are not guaranteed to be optimal solutions to their respective single facility problem even if the other co-existing subsets were not considered, and (b) the duplicate reduction process (Step 4) involves greedy selection of the node to treat and the facility to which the treated node will be assigned. Thus, the procedure will, at best, provide some lower bound for the producer's expected profits. This bound, however, might not be tight, but could serve as a starting point for speeding up large instances of the branch-and-price procedure by providing an improved set of starting columns.

Second, note that fixing the value of p does not necessarily mean all p facilities will be opened, as some chosen sites from the p-median solution might have an empty set of allocation subsets with positive profits; i.e., if the base subset $S_j = \{j\}$ has negative profits, then any other subset will do so as well. Further work may focus on improving approximation algorithms for the (MFMS) for both fixed and arbitrary number of facilities.

2.6 Numerical Tests: Subset Formation Heuristics for (SFMS)

We present computational results for the solution procedures suggested for the (SFMS), as a base for comparison of performance, relevant due to its structural equivalence to the pricing problem of the (MFMS) and (MFPA).

As reference, all instances of the problems were implemented using a server with a Xeon processor, with 2.93 GHz and 3 GB RAM. Algorithms were programmed using Mathematica, by Wolfram Research, version 7.0.1. However, we place little emphasis on optimizing the computational performance of the algorithms, and the running times presented are for showing relative performance between the proposed procedures, and (whenever reasonable) with respect to enumeration.

The tests were split into two groups. The first, with graphs of sizes $n = \{6, 10, 15\}$ set over a square surface of 100×100 km (roughly the area of Connecticut), serves as a control group to verify the performance of the heuristics against verifiable optimal results obtained through enumeration. Demand node locations and weights were generated randomly. Relative weights were then adjusted to reflect physical demand for hydrogen. The rest of the parameters are as follows: $p^E \sim N(0.05, 0.1)$, $R \sim U(0, 100000)$, $Q(C_j) = 5 + 0.01C_j^{1.5}$, $e = 0.01871$, $m = 0.079$, $w = 0.038$, and $p^H = 3.5$. The results for this first group are summarized in Table 2.3. Note that the running times shown on the table correspond to a single node's allocation problem. Thus, the total time required for selecting the optimal location would then be (on average) n times larger than those shown. The optimality gaps shown are estimated based only on suboptimal instances, thus the average gap including optimal instances would have been substantially lower.

The hybrid heuristic achieved optimality in all but one of the instances (in the interest of full disclosure, the greedy heuristic did find the corresponding optimal allocation in this instance). The performance of the greedy heuristic worsened (both in percentage of optimal solutions reached, as well as the average gap of suboptimal instances) when the number of nodes increased.

The second group, with $n = 50$, allows for comparisons between the heuristics for cases where exhaustive enumeration is not reasonable. These graphs were set on square surfaces of two areas: 100×100 km, and 800×800 km (roughly the size of France). In the case of the larger surface, hydrogen price was set at $p^H = 4.0$ €/kg to partially compensate for the increased distances. The results for this second group are presented in Table 2.4.

As we don't have a validated optimal solution, the gap values for each proce-

Instance		Nearest-node			Greedy			Hybrid			Exhaustive
n	Cases	% Opt.	% Gap	Avg. Time	% Opt.	% Gap	Avg. Time	% Opt.	% Gap	Avg. Time	Avg. Time
6	60	78.3	0.73	0.251	91.7	2.59	0.787	98.3	2.59	1.251	1.86
10	100	81.0	0.90	0.365	89.0	2.51	2.162	100.0	0.00	2.643	32.23
15	75	76.0	0.215	0.527	76.0	3.06	4.950	100.0	0.00	5.628	1063.48

Table 2.3: Comparison of heuristics for the (SFMS) for small size instances (6–15 nodes). The *Cases* column indicates the number of nodes for which the heuristics were executed; *% Opt.* the proportion of nodes for which the optimal allocation (verified by enumeration) was achieved; *% Gap* the average deviation from optimality for each suboptimal result with respect to the enumeration solution; and *Avg. Time* the running time in seconds for each node's allocation problem.

Instance			Nearest-node			Greedy			Hybrid		
Surface	n	Cases	% Best	% Gap	Avg. Time	% Best	% Gap	Avg. Time	% Best	% Gap	Avg. Time
100×100	50	250	74.0	0.028	1.26	43.6	0.62	45.40	99.6	0.013	47.58
800×800	50	250	99.6	2.00	0.46	100.0	0.00	18.47	100.0	0.00	20.22

Table 2.4: Comparison of heuristics for the (SFMS) for $n = 50$. The *Cases* column indicates the number of nodes for which the heuristics were executed; *% Best* the proportion of nodes for which the resulting allocation produced the highest expected profits across the three heuristics; *% Gap* the average deviation from optimality for each suboptimal result with respect to the enumeration solution; and *Avg. Time* the running time in seconds for each node's allocation problem.

dure are measured with respect to the best solution achieved across the three heuristics, and only computed for those instances where that specific procedure does not match the best solution. In general, all heuristics fared much better in the case of the larger surface, converging to the same solution in all but one case. The main reason is the effect of longer distances in transportation costs (and, consequently, on marginal hydrogen profitability), causing the tradeoffs between the different nodes available for inclusion to be more evident. The reduction in computation time is due to the larger surface inducing smaller-cardinality optimal subsets.

The hybrid heuristic was outperformed by the greedy heuristic in only one instance for the smaller surface. Recall that the hybrid heuristic implemented for these experiments used the base subset from the nearest-node solution; thus, for this unique instance, the greedy solution was not reachable through an improving path of single-node exchanges from the nearest-node solution, and may not be itself a global optimum. Still, the hybrid heuristic's performance makes it suitable for a large-scale implementation.

An expanded analysis of computational results for a larger instance of the design problem will be presented as part of a case study in §4.

Chapter 3

Decision Under a Proportional Allocation Policy

Here we will focus on the second distribution policy that the regulator can set in place. In contrast with the local coverage focus of the market selection policy, this second policy will focus on spatial coverage.

Definition 2. *A* **proportional allocation policy** *consists of the following conditions:*

- *The producer has to serve the entire set of demand locations N.*
- *Demand can be served partially within a node, but the proportion of demand served (with respect to the node's total demand) has to be the same across the entire network.*

We consider the proportional allocation policy to be preferred by a regulator wishing to have technology adoption occur at the same rate throughout the network.

The contrast between the proportional allocation (presented in this chapter) and market selection policy (addressed in Chapter 2) is significant. The results from Ball et al. (2007) indicate the introduction of hydrogen fuel in densely populated areas to have a significant impact in reducing infrastructure costs. This approach is consistent with the roadmap for hydrogen infrastructure build-up planned by the

European Union (European Commission 2012b), which focuses on a small number of early user centers. Thus, a cost-centric analysis shifts towards local coverage playing a more relevant role than spatial coverage.

Conversely, spatial coverage could potentially play a significant role in accelerating adoption of the new technology at a more sustainable rate across a greater geographical area, through spatial spillover effects.[1] Some studies present empirical evidence of such spillover effects in the adoption of clean or improved technologies in agricultural settings (Conley and Udry 2010, Lewis et al. 2011) and can influence the overall effect of environmental policy (Banzhaf and Chupp 2010). Additionally, equal access to renewable energy benefits is a key element in the sustainable development of rural and less densely populated areas in Europe (OECD 2012). Finally, spatial spillovers can be related to the reinforcing effect of learning-by-doing, which has been addressed in the context of fuel-cell vehicles by Schwoon (2006a,b). A wider adoption on a geographical level might create enough mass in a larger number of markets, with local learning dynamics playing a greater role in increasing the share of vehicles adopting the new technology.[2] The contrary point of view has been offered by Farrell et al. (2003), suggesting that succesful large-scale deployment of hydrogen as an alternative fuel shall be achieved with greater likelihood if efforts were put in place towards achieving significant market penetration in a single node or a geographically restricted area (i.e., a protected niche), which would maximize societal learning effects while minimizing infrastructure costs and risks. We note such demand dynamics are out of the scope of this thesis, but wished to highlight how the perceived benefit of these dynamics might influence the regulator's choice of distribution policy, and justifies considering both policies to understand their effect on system behavior.

Inducing spatial coverage must come at a cost for the producer, which accentu-

[1] Spatial spillovers are externalities caused by neighboring agents. In the context of our problem, the rate of adoption of a new technology within a region can be shaped by the presence of the technology in neighboring areas.

[2] For a more expanded view on these issues, the reader may refer to the literature in social learning and technology diffusion (Rogers 1995).

ates the role of the regulator in creating appropriate incentive schemes that permit a proper assessment of the trade-off between both types of coverage and the social benefits associated to accelerated adoption or equal access to the new technology.

When the firm is subject to an external (regulatory or contractual) condition requiring coverage of all demand nodes at a constant service level, production capacity implicitly becomes a decision variable for our problem. We first address the producer's decision under this new setting for single and multiple production sites.

3.1 Optimal Capacity Decision

We formulate the unconstrained capacity optimization problem (COP) for a firm with a single renewable generation source, with a fixed production site at j serving a set of nodes $S_j \subseteq N$ under a proportional allocation policy. Using the expected profit formulation (2.4) we have:

$$\text{(COP)} \quad \underset{C_j \geq 0}{\text{Max}} \ \hat{\Pi}_j(C_j) = K_j^H(\bar{t}_{j|S_j})C_jF_R^c(C_j/e) + eK_j^H(\bar{t}_{j|S_j})G_R(C_j/e) - Q(C_j)$$

By Proposition 1, we know that for a single technology configuration, and fixed facility location and service subset, the producer's total expected profit function is concave. Thus, a unique maximum for $\hat{\Pi}_j(C_j)$ will exist and can be determined using the first order conditions defined in (2.5). Let C_j^I be said maximum, then:

$$C_j^I = eF_R^{-1}\left(\frac{K_j^H(\bar{t}_{j|S_j}) - Q'(C_j^I)}{K_j^H(\bar{t}_{j|S_j})}\right)$$

The profit function is illustrated in Figure 3-1. The value C_j^I can be obtained with relative ease by a search procedure such as Newton's method (Bertsekas 1999). This result can clearly be related to the newsvendor problem (Arrow et al. 1951), where the fraction between the brackets is analogous to a critical fractile balancing the cost of overbuilding and underbuilding production capacity, with the marginal costs of overbuilding capacity given by $Q'(C_j^i)$, while the marginal

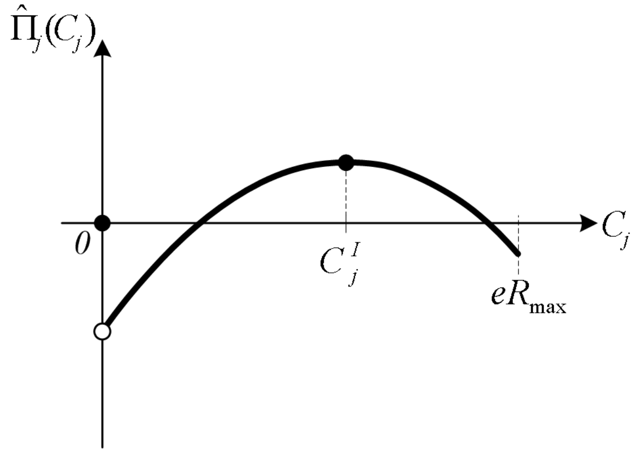


Figure 3-1: Expected profit function for a fixed location and service subset.

costs of underbuilding capacity are given by $K_j^H(\bar{t}_{j|S_j}) - Q'(C_j^I)$. The reader may notice that in this expression there is no need for adjusting $K_j^H(\cdot)$ by the proportion of time that hydrogen production is done, as is done in equation (2.7), as that expression is stated in terms of subset demand, while (2.4) captures the relationship between output and capacity. It is only for the tranformation of the function from capacity-dependent to demand-dependent where this adjustment is necessary.

Note the upper bound on hydrogen plant capacity is implicitly constrained by the support of the energy supply distribution, which we do not assume here to be restricted to any particular family of distributions, as long as it's continuous and twice differentiable. Since fixed capacity costs are only incurred if a positive capacity is built, the proper solution to (COP) is $C_j^* = C_j^I$ if $\hat{\Pi}_j(C_j^I) > 0$, and $C_j^* = 0$ otherwise.

A further issue arises if the optimal solution to (COP) results in an output that exceed what would be the desired demand. In the typical newsvendor model, the optimal inventory/purchase level is determined by the demand distribution, and it is implied the supplier has ample capacity. If this assumptions was not enforced, the newsboy solution would only be optimal if it is smaller than the supplier's capacity, and if the supplier's capacity is insufficient to satisfy the optimal quantity then the highest profits would be achieved at the upper bound of the supplier's capacity. Conversely, in our model, the optimal capacity to be built is dependent on the supply distribution. The solution is not completely independent on demand, as demand acts as a constraint, and can be formalized as $\min\{D_{S_j}, \bar{H}_j(C_j^I)\}$. In the analysis shown above we have dealt with the unconstrained problem, which is equivalent to assuming demand is large enough to assure that the natural constraint placed by the support of the energy supply distribution is tighter than that set by an exogenous upper bound on demand. This assumption does not affect the validity of the newsvendor-type solution shown above, but rather allows us to understand properly the dynamics of producer's profits with respect to system capacity in an unconstrained setting, while noting that the adjustment to a constrained setting requires only a comparison of two values

3.2 Locating a Single Facility on the Network

We extend upon the previous model by considering that the firm wishes to choose the location $y \in G$ and the capacity level C_y^* of a single facility to serve the entire set of demand nodes N. This problem is formulated as follows:

$$\text{(SFPA)} \quad Z^{PA}(y, C_y) = \underset{C_y \geq 0, y \in G}{\text{Max}} \; K_y^H(\bar{t}_y)\left(C_y F_R^c(C_y/e) + eG_R(C_y/e)\right) - Q(C_y)$$

Let $j \in G$ be a fixed plant location with fixed capacity C_j serving a set of nodes S_j, with $\bar{t}_{j|N}$ the demand-weighted unit transportation cost to the demand nodes in N from j. The following property holds for the producer's profit function.

Proposition 4. *For a fixed subset and capacity value, $Z^{PA}(j, S_j, C_j)$ is decreasing in $\bar{t}_{j|S_j}$.*

Proof. For a given location j and allocation subset S_j we have:

$$Z^{PA}(j, S_j) = D_{S_j} K_j^H(\bar{t}_{j|S_j}) - Q\left(C_j^*(D_{S_j})\right)$$

Given capacity C_j is fixed, only K^H is a function of transportation costs $\bar{t}_{j|S_j}$, hence proving $\frac{dZ^{PA}(\bar{t}_{j|S_j})}{d\bar{t}_j} \leq 0$ is equivalent to showing $\frac{dK^H(\bar{t}_{j|S_j})}{d\bar{t}_{j|S_j}} \leq 0$. The same applies for the proof of convexity. As $\beta_j(\bar{t}_{j|S_j}) = p^H - m - \frac{w}{e} - \bar{t}_{j|S_j}$ and $\delta_j(\bar{t}_{j|S_j}) = e(p^H - m - \bar{t}_{j|S_j})$, we have:

$$
\begin{aligned}
\frac{dK^H(\bar{t}_{j|S_j})}{d\bar{t}_{j|S_j}} &= -F_{p^E}(\delta_j(\bar{t}_{j|S_j})) - \beta_j(\bar{t}_{j|S_j}) f_{p^E}(\delta_j(\bar{t}_{j|S_j})) \\
&\quad + \frac{\delta_j(\bar{t}_{j|S_j})}{e} f_{p^E}(\delta_j(\bar{t}_{j|S_j})) - \frac{w}{e} f_{p^E}(\delta_j(\bar{t}_{j|S_j})) \\
&= -F_{p^E}(\delta_j(\bar{t}_{j|S_j})) < 0 \\
\frac{d^2K^H(\bar{t}_{j|S_j})}{d\bar{t}^2_{j|S_j}} &= f_{p^E}(\delta_j(\bar{t}_{j|S_j})) > 0
\end{aligned}
$$

Thus, $Z^{PA}(j, C_j)$ is decreasing and convex in $\bar{t}_{j|S_j}$ for a fixed location j and subset S_j. ■ □

We point out that, unlike Proposition 2, here the conditions for monotonicity and convexity are unrestricted in the fixed parameters, as the properties described in Proposition 4 are for a fixed capacity value that is not necessarily that required to satisfy completely the demand of the subset. In other words, in the market selection case the condition of fully satisfying nodal demand forces the required capacity (which is embedded in the value of $Q(\cdot)$) to be indirectly a function of transportation costs, as the proportion of time dedicated to production for a fixed supply distribution will be different. In the proportional allocation case, the capacity to be built is chosen in an unconstrained manner, hence only the revenue term is affected by transportation costs. When capacity is fixed, any change in transportation costs will affect revenue solely within the estimations of parameter $K^H_j(\cdot)$.

We show that the (SFPA) problem can be separated into two distinct sequential decisions. We define the 1-median problem as choosing the point $j \in G$ which minimizes the weighted transportation cost to the respective set of nodes N. It is formally defined as

$$\operatorname*{Min}_{j \in G} \sum_{i \in N} t_{ij} D_i.$$

Theorem 2. *The optimal location y^* for the (SFPA) problem is equivalent to the 1-median solution (y^{1M}) on the same set of candidate locations.*

Proof. Let $j' \in G$ be a fixed location with associated weighted transportation cost to the set of demand nodes N given by $\bar{t}_{j'|N}$, and $C^*_{j'}$ that solves its corresponding (COP) problem. Likewise, for a given $j'' \in G$ with transportation costs $\bar{t}_{j''|N} \geq \bar{t}_{j'|N}$, $C^*_{j''}$ is the solution to its (COP) problem.

Optimality of $C^*_{j'}$ with respect to j' implies $Z^{PA}(j', N, C^*_{j'}) \geq Z^{PA}(j', N, C^*_{j''})$. From Proposition 4 we have that for a fixed location $j \in G$ and capacity value C_j, $Z^{PA}(j, N, C_j)$ is decreasing in $\bar{t}_{j|N}$. Then, for a fixed capacity value $C^*_{j''}$ we have $Z^{PA}(j', N, C^*_{j''}) \geq Z^{PA}(j'', N, C^*_{j''})$, as $\bar{t}_{j''|N} \geq \bar{t}_{j'|N}$, meaning that a lower-cost location will provide higher profits than a higher-cost location for any given capacity; thus, all other parameters being equal, Proposition 4 holds for any pair of distinct locations in the network. Let $y^* = \text{argmin}_{j \in G}\{\bar{t}_{j|N}\}$. Then, $Z^{PA}(y^*, N, C^*_{y^*}) \geq Z^{PA}(j, N, C^*_j), \forall j \in G$.

Recall $\bar{t}_{j|S_j} = \frac{\sum_{i \in S_j} D_i t_{ij}}{\sum_{k \in S_j} D_k}$. From the 1-median formulation, we have:

$$\underset{j \in G}{\text{Min}} \sum_{i \in N} t_{ij} D_i = \underset{j \in G}{\text{Min}} \sum_{k \in N} D_k \bar{t}_{j|N} = \sum_{k \in N} D_k \underset{j \in G}{\text{Min}}\, \bar{t}_{j|N}$$

Thus, $y^{1M} \equiv \text{argmin}_{j \in G}\{\bar{t}_{j|N}\} \equiv y^*$, meaning the solution to the 1-median problem also selects the optimal location for the (SFPA) problem. □

Corollary 1. *Under a proportional allocation policy, the optimal solution for the (SFPA) can be obtained from* $y^* = \text{argmin}_{j \in G}\{\bar{t}_{j|N}\}$.

Proof. Follows from proof of Theorem 2. □

Corollary 2. *An optimal location y^* for (SFPA) can always be found in the set of nodes N.*

Proof. Follows from the node optimality theorem (Hakimi 1964). □

While Corollary 1 simplifies the problem to searching for the location with the lowest weighted transportation cost value $\bar{t}_j$, Corollary 2 reduces the search space to the node set, thus the single facility location problem can be solved in $O(n)$.

The second step in the decision for the single facility problem involves selecting the optimal capacity level, which can then be obtained by solving the (COP) problem for the chosen location y. By definition, the allocation of product to each market $i \in N$ will be a fraction ρ_i of expected production output $\bar{H}_y(C_y^*)$.

3.3 Locating Multiple Facilities on the Network

We extend the previous formulation to locating multiple facilities on the network with each demand node restricted to being supplied from a single facility.

The placement of multiple facilities on the network is equivalent to finding a feasible partition $\mathcal{S}$ of the set N maximizing total expected profits, where $S_j \in \mathcal{S}$ is the set of nodes served by a single facility located in the graph $G[S_j]$ induced by that node subset.

The allocation proportions $\rho_i = \frac{D_i}{\sum_{k \in N} D_k}$ are, by definition, determined with respect to the entire network independent of the facility serving that demand node; thus, when capacity is under control of the firm, an exogenous condition is required to assure the ratio of produced and required hydrogen to be maintained across subsets. Let $\rho_{S_j} = \sum_{i \in S_j} \rho_i$ be the proportion of demand corresponding to subset S_j, and C_{y_j} the production capacity to be made available at the single facility serving that same subset, with $\mathbf{C}$ the vector of such capacities. Additionally, define $C_N = \sum_j C_{y_j}$ as the total production capacity in the network. We can then formulate the multiple facility location-capacity decision as the following combinatorial problem:

$$
\begin{aligned}
\text{(MFPA)} \quad Z^{PA}(\mathbf{y}, \mathcal{S}, \mathbf{C}) = & \operatorname*{Max}_{\mathcal{S}, \mathbf{C} \in \mathbb{R}_+^{|\mathcal{S}|}, \mathbf{y} \in G} \sum_{j=1}^{|\mathcal{S}|} \hat{\Pi}_{y_j}(y_j, S_j, C_{y_j}) \qquad (3.1) \\
\text{s.t.} \quad & \bigcup_j S_j = N; \quad S_i \bigcap S_j = \{\varnothing\}, \forall i \neq j \\
& y_j \in S_j, \ \forall j \\
& C_{y_j} = \rho_{S_j} C_N, \ \forall j
\end{aligned}
$$

The first set of conditions define a feasible partition; while the second group requires each subset to be served from within its member nodes. The last set of constraints enforces the proportional allocation policy. As all plants share a single stochastic energy source, and facility costs are increasing in capacity, it is assured that no location will have more capacity than that required to comply with the proportional allocation policy, resulting in the last set of constraints. Note this formulation is for an arbitrary number of production sites. Just as in the market selection case, we can represent an exogenously-defined number of facilities through a partition cardinality condition. Let $\bar{H}_N(C_N) = C_N F_R^c(C_N/e) + eG_R(C_N/e)$ represent the total expected output of the hydrogen production network. Then the following results can be derived.

Proposition 5. *For the (MFPA) we have:*

(i) The producer's expected profit function for the (MFPA) can be stated as a function of a single capacity value C_N, as

$$Z^{PA}(\mathbf{y}, \mathcal{S}, C_N) = \underset{\mathcal{S}, C_N \in \mathbb{R}_+, \mathbf{y} \in G}{\text{Max}} \; \bar{H}_N(C_N) \sum_j \rho_{S_j} K^H(\bar{t}_{y_j|S_j}) - \sum_j Q(\rho_{S_j} C_N) \qquad (3.2)$$

(ii) Given a partition $\mathcal{S}$, the optimal production capacity vector $\mathbf{C}$ for the (MFPA) can be obtained by solving the (COP) for a single capacity C_N using the expression

$$C_N = e\, F_R^{-1}\left(\frac{\sum_j \rho_{S_j} K_j^H(\bar{t}_{j|S_j}) - \sum_j \rho_{S_j} Q'(\rho_{S_j} C_N)}{\sum_j \rho_{S_j} K_j^H(\bar{t}_{j|S_j})} \right)$$

and allocating a capacity $C_{y_j} = \rho_{S_j} C_N$ to the facility serving each subset S_j.

Proof. (i) From the proportional supply constraint in (3.1) we have that the global production capacity can be stated as $C_N = \sum_{S_j \in \mathcal{S}} C_{S_j}$. Let $\bar{H}_N(C_N) = C_N F_R^c(C_N/e) + eG_R(C_N/e)$ represent the total expected output of the hydrogen production network. Then, by rearranging terms and replacing the individual facility capacities by $\rho_{S_j} C_N$, we can restate (3.1) as (3.2).

(ii) Total expected hydrogen output $\bar{H}_N(C_N)$ is nondecreasing and concave in global capacity (C_N). For a given partition $\mathcal{S}$, the vector $(\rho_{S_1}, \ldots, \rho_{S_p})$ is fixed. Thus, $\sum_{j=1,\ldots,p} Q(\rho_{S_j} C_N)$ is a convex increasing function of C_N, which satisfies our definition of (COP), thus the solution procedure presented for the (COP) will solve the capacity problem for a fixed partition of the (MFPA). $\square$

Capacity costs $Q(\rho_{S_j} C_N)$ are driven both by total network capacity, and by the proportion of demand assigned to each subset. For that reason, a partition providing highest contribution to profits (i.e. maximizing the first term) might not provide maximum operating profits due to a more costly allocation of capacity costs (in the special case where capacity costs are linear and the number of facilities p is fixed, maximizing operating and net profits are equivalent problems). We establish nodal optimality compliance for the multiple facility problem.

Theorem 3. *The set of optimal locations for the (MFPA) can be found in the set of demand nodes ($\mathbf{y}^* \subseteq N$).*

Proof. Assume $(\mathcal{S}', \mathbf{y}', C'_N)$ is the optimal solution to (MFPA) and $\mathbf{y}'$ (the set of optimal locations) includes at least one component y_k outside of the set of nodes (i.e., $y_k \in A$). Leaving all other locations in $\mathbf{y}'$ and their served subsets fixed, by Corollary 2 there exists a nodal location $y_k \in S_k$ which produces higher expected profits for subset S_k and, in turn, higher total expected profits for the (MFPA) given the profits from all other subsets remain unchanged. This procedure can be iterated for sequentially eliminating other non-nodal components. The existing optimal capacity C'_N is feasible for the new set of locations, hence any optimal capacity $C''_N \neq C'_N$ obtained from the nodal solution will result in profits at least as high as the non-nodal solution. Thus, for every instance of the (MFPA), there exists a profit maximizing set of locations $\mathbf{y}^* \subseteq N$. $\square$

Corollary 3. *For a given partition $\mathcal{S}$, the set of optimal locations $\mathbf{y}^*$ will be defined by $\{y_1, \ldots, y_p\}$, with $y_j = \operatorname{argmin}_{i \in S_j}\{\bar{t}_i\}$.*

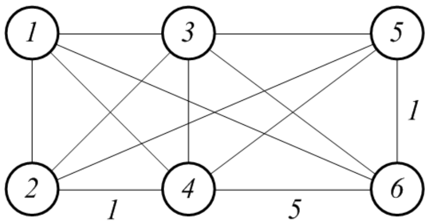
Proof. Follows from Corollary 1 and Theorem 3. $\square$

By Proposition 5 and Corollary 3, the last two sets of constraints in (3.1) will necessarily be satisfied due to the structure of the problem and the behavior of the producer's profit function, even if they are not explicitly included in the formulation. The optimality of the 1-median solution for the single facility problem raises an equivalent question for the special case of locating $p > 1$ facilities; namely, whether the solution to a p-median problem based on transportation costs t_{ij} is optimal for the location of p facilities in the (MFPA) problem. Unlike the (MFMS), the (MFPA) shares with the p-median problem the condition that all nodes in the network shall be served. The p-median problem (Mirchandani 1990, Daskin 1995) for $G = (N, A)$ can be expressed as

$$\begin{aligned}
&\underset{\mathcal{S},\mathbf{y}\subseteq N}{\text{Min}} \sum_j D_{S_j} \bar{t}_{y_j|S_j} \\
&\text{s.t.} \quad \bigcup_j S_j \subseteq N; \quad S_i \bigcap S_j = \{\emptyset\}, \forall i \neq j; \quad |\mathcal{S}| = p \\
&\qquad\quad y_j \in S_j, \ \forall j
\end{aligned}$$

Although the general p-median problem is NP-hard, for a fixed value of p the problem is polynomial-time solvable (Garey and Johnson 1979), and there are numerous algorithms available to solve the problem with relative efficiency for reasonably-sized instances. For that matter, its equivalence to the (MFPA) for a fixed number of facilities would be relevant for simplifying the problem's solution complexity. We show that the two problems are not equivalent by means of a counterexample.

Let $n = 6$ and $\rho_i = \frac{1}{6}$, with Figure 3-2 depicting such a network and the associated symmetric transportation cost matrix $[t_{ij}]$. For illustrative purposes, K^H has been estimated as $K^H(\bar{t}_{j|S_j}) = \frac{10}{\bar{t}_{j|S_j}}$, capacity costs as $Q(C_N) = 5 + 0.1 \sum_k (\rho_{S_k} C_N)^3$, and expected production output as $\bar{H}_N(C_N) = \sqrt{C_N}$. The reader can easily verify these expressions satisfy $K^H(\cdot)$ being convex decreasing in transportation costs (Proposition 4), our assumption of $Q(\cdot)$ being convex increasing on C_N, and $\bar{H}_N(\cdot)$ being concave nondecreasing on C_N. The p-median and (MFPA) solutions for



	1	2	3	4	5	6
1	0.1	1	1	$\sqrt{2}$	6	$\sqrt{37}$
2		0.1	$\sqrt{2}$	1	$\sqrt{37}$	6
3			0.1	1	5	$\sqrt{26}$
4				0.1	$\sqrt{26}$	5
5					0.1	1
6						0.1

Figure 3-2: Counterexample of p-median optimality: Network depiction and transportation cost matrix.

Figure 3-3: Counterexample of p-median optimality: (PMED) and (MFPA) solutions.

$p = 2$ are shown in Table 3.1. The first column shows the resulting partition, with the facility nodes in bold. Due to the symmetry in transportation costs, there are multiple location choices yielding the same objective value; the result presented was chosen to highlight that even when the chosen locations were the same, the allocation (partitioning) had a significant impact in firm profits. The second and third columns show the values of the p-median and (MFPA) objectives for the resulting optimal partitions. The last column indicates the optimal capacity level for each resulting partition.

Not only is the p-median solution not optimal for the (MFPA), but the p-median solution results in 40% lower expected profits than the corresponding (MFPA) solution.

There are some key differences between both problems. First, the convexity of marginal hydrogen profits $K^H(\cdot)$ causes marginal changes in average transportation

Objective	Partition	Z^{PMED}	Z^{PA}	$C_N^*(\mathcal{S}, \mathbf{y})$
(PMED)	$\{1, 2, 3, \mathbf{4}\}, \{\mathbf{5}, 6\}$	**0.769**	21.468	5.415
(MFPA)	$\{1, 2, 3, \mathbf{4}, 6\}, \{\mathbf{5}\}$	1.436	**35.963**	5.198

Table 3.1: Counterexample of p-median optimality: Result summary.

costs to have a different effect on contribution to profits across subsets; specifically, the profitability of subsets with relatively small average delivery cost $\bar{t}_{y_j|S_j}$ is more sensitive to changes in transportation costs resulting from including or excluding a node from the subset. Second, the difference in profitability associated to the assignment of a node to a subset (both by its effect on contribution and capacity costs) means that, as opposed to the p-median problem where the number of alternative solutions is limited to the number of combinations ${}_nC_p$, the set of possible solutions for the (MFPA) extends to all possible partitions of N. An exact procedure involving enumeration of all possible partitions would not be feasible even for problems of relatively small size –the set partitioning problem is NP-complete.

To address the (MFPA) problem we have evaluated different approaches. One such approach was based on Lagrangian relaxation. This method, however, proved impractical as the complexity of the resulting relaxed problem was not reduced with respect to the original problem, as it happens with the UFLP and p-median problems. The next section approaches this problem making use of column generation (Dantzig and Wolfe 1960). The strong connection between Lagrangean relaxation and column generation has been well studied (Nemhauser and Wolsey 1999); still, in this case, the implementation of both methods differs greatly, given the structure of the profit function.

3.3.1 Exact Method for Fixed Service Level: Branch-and-Price

When the value for network production capacity (C_N) is fixed, we can reformulate the (MFPA) as a partitioning problem (Balas and Padberg 1976). Since there is a direct correspondence between capacity and production output, we can relate network capacity directly to a desired network-wide service level; i.e., the fraction of each market's demand that can be satisfied. Thus, an exact method can be derived to optimally design the production-distribution network for hydrogen production given a target (local) demand service level $\phi \in [0, 1]$.

The problem's column generation form is then a variant of the set partitioning formulation established in Barnhart et al. (1998). Since every subset in an optimal solution to the (MFPA) necessarily satisfies the 1-median property (Corollary 3), then each subset is characterized by a unique column, and the set of facility locations is fully defined by the underlying partition subsets. Thus, a more refined version of the set partitioning problem can be used for defining our master problem (Minoux 1987), which will be as follows.

$$\begin{aligned}
\text{(CGPA)} \quad \underset{\lambda}{\text{Max}} \quad & \sum_k z_k \lambda_k && (3.3)\\
\text{s.t.} \quad & \sum_k \theta_{ik} \lambda_k = 1, \quad \forall i = 1, \ldots, n \\
& \lambda_k \in \{0, 1\}, \quad \forall k \in \omega
\end{aligned}$$

where

$$z_k = \hat{D}_{S_j} K_j^H(\bar{t}_j) - Q(C_j^*(\hat{D}_{S_j})),$$

with $\hat{D}_{S_j} = \phi D_{S_j}$ representing the hydrogen demand adjusted for the required local service level, while θ_{ik} maintains the definition established in the previous section. For a given potential location $j \in N$ we have the corresponding pricing subproblem:

$$\underset{S_j}{\text{Max}} \quad \hat{D}_{S_j} K_j^H(\bar{t}_j) - Q(C_j^*(\hat{D}_{S_j})) + \sum_{i \in S_j} \mu_i$$

with $\mu_i \leq 0$ the dual variable of the ith convexity constraint from the master problem. Likewise, an optimal solution is assured if all subsets resulting from solving the n pricing subproblems yield corresponding negative increased profits. The column generation algorithm presented in §2.5.1 is suitable for implementation for the (MFPA) given a network capacity / service level value has been fixed, noting that for Step 1 the new definition of z_k shall be used, and that the relaxed problem to be solved in Step 2 corresponds to the formulation (CGPA). Solutions to this problem for a fixed capacity value enable us to compare the effect of the proportional allocation policy on the producer's profits; in other words, obtain a monetary value for the impact of the regulator enforcing equity in the access to the new technology. We use this algorithm to contrast the results from both policies in §4.

3.3.2 Heuristic Procedures for Endogenous Capacity Decision

The branch-and-price procedure described in §3.3.1 is appropriate when network capacity is a prior to the solution of the problem. A stark contrast of the two distribution policies –market selection and proportional allocation— is the producer's ability to choose its optimal capacity (ergo, expected output) level. However, the solution to the capacity optimization problem presented in §3.1 is applicable only for a fixed partition. For simultaneously optimizing network design and capacity, a more involved procedure would need to be put in place that iteratively optimizes capacity based on improved network solutions. Although not the main focus of this thesis, we explain two procedures to address this variant of the problem.

The first (naive) approach involves repeatedly solving the fixed service level problem described in §3.3.1, for a large set of values of ϕ. The precision of the resulting solution will largely depend on the number of service level scenarios solved (or, conversely, the distance between successive choices of ϕ). Even if a large set of problems are solved, in this naive implementation the maximum profit solution is not even guaranteed to be a local optimum. To obtain a true local optimum, the fixed service level solution achieving the highest expected profits should be subjected to the corresponding capacity optimization problem.

The main advantage of this approach is its ease of implementation; and progressive estimation might allow the detection of patterns in profits, which would allow narrowing the space for locating the (nearly) optimal solution (one such pattern, for instance, involves that if a partition is optimal for two separate service levels, it will also be optimal for every ϕ value in between; thus, when a region of focus has been identified where the optimal solution is considered to be located, then such a repeated optimal partition would guarantee the global optimum is found.

A second approach is more involved in nature, but its inner structure is based on the same column generation algorithm used in the fixed service level (MFPA) problem. We proceed to define this approach.

Let λ be an optimal solution to (CGPA), and μ_i the dual variable associated to each convexity constraint. The following procedure solves the location-allocation problem for a fixed capacity value and then optimizes capacity for the resulting partition. Since for the new capacity a different partition might provide higher expected profits to the producer, the algorithm iteratively repeats this process until convergence to a local maximum. To test for better available solutions, a new capacity value is set from the unexplored space. A large number of location-allocation problems would need to be solved to cover a significant portion the capacity space and increase the guarantee of a global optimum being achieved.

Define ϵ as a small positive number, and $\alpha \in (0, 1)$. Parameter ϑ_{max} defines the maximum number of iterations that will be allowed without finding an improvement in expected profits. Step coefficient α defines how much the capacity value will be perturbed to continue the procedure from a found local optimum. Smaller values for α will result in more exhaustive exploration of the capacity domain, same as a greater value of ϑ_{max}. Because of these arbitrary rules which guarantee termination, the algorithm is not guaranteed to find a (globally) optimal solution.

Modified branch-and-price algorithm for (MFPA):

Step 0 (Initialization): Set iteration limit ϑ_{max}. Define initial set of $n+1$ columns ω as all singleton sets and N. Set initial capacity $C_N^1 = \epsilon(e\bar{R})$ (where ϵ is a small positive value). Set $\mathcal{S}^0 = \{\emptyset\}$, $\mathbf{z}^* = \mathbf{z}^0 = 0$, and $\tau = 1$.

Step 1 (Relocation): For each column $k \in \omega$, locate facility at the 1-median solution, i.e., $y_k = \operatorname{argmin}_{i \in S_k}\{\bar{t}_{i|S_k}\}$; calculate $z_k^\tau = Z^{PA}(y_k, S_k, C_N^\tau)$.

Step 2 (Relaxed problem): Solve (CGPA) for z_k^τ, ω, C_N^τ.

Step 3 (Insertion): Based on dual variables μ_i, solve the following pricing subproblem for each designated base node j:

$$\underset{S_j}{\text{Max}} \quad Z^{PA}(j, S_j, C_N^{\tau}) + \sum_{i \in S_j} \mu_i$$

Resulting nonempty solutions from pricing subproblems are added to set ω, return to Step 1. If no new columns are formed with positive increased profits, then proceed to Step 4.

Step 4 (Branching): If $\lambda_k \in \{0, 1\}, \forall k \in \omega$ then $\mathcal{S}^{\tau}$ is optimal partition for (MFPA) for capacity C_N^{τ}, go to Step 5; else, choose one element $\lambda_k \in (0, 1)$ for branching, return to Step 2.

Step 5 (Validation of local optimality): If $\mathcal{S}^{\tau} = \mathcal{S}^{\tau-1}$, a local optimum has been found, set $C_N^{\tau+1} = C_N^{\tau} + \alpha(e\bar{R} - C_N^{\tau})$, and go to step 7; otherwise proceed to Step 6.

Step 6 (Capacity optimization): Obtain C_N^* for the partition $\mathcal{S}^{\tau}$ by solving:

$$C_N^*(\mathcal{S}^{\tau}) = eF_R^{-1}\left(\frac{\sum_{S_j \in \mathcal{S}^{\tau}} \rho_{S_j}\left(K^H(\bar{t}_{y_j|S_j}) - Q'(\rho_{S_j} C_N)\right)}{\sum_{S_j \in \mathcal{S}^{\tau}} \rho_{S_j} K^H(\bar{t}_{y_j|S_j})}\right)$$

Set $C_N^{\tau+1} = C_N^*(\mathcal{S})$. Calculate $\mathbf{z}^{\tau} = Z^{PA}(\mathbf{y}^{\tau}, \mathcal{S}^{\tau}, C_N^{\tau+1})$. If $\mathbf{z}^{\tau} > \mathbf{z}^*$, then $\mathbf{z}^* \leftarrow \mathbf{z}^{\tau}$, $\vartheta = 1$; otherwise $\vartheta = \vartheta + 1$.

Step 7 (Parameter updating): If $\vartheta = \vartheta_{max}$, stop; otherwise $\tau \leftarrow \tau + 1$, calculate $z_k^{\tau} = Z^{PA}(y_k, S_k, C_N^{\tau}), \forall k \in \omega$; go to Step 2. ■

The parameter ϑ_{max} defines the maximum number of iterations that will be allowed without finding an improvement in total expected profits. Step coefficient α defines how much the capacity value will be perturbed to avoid having the algorithm get caught indefinitely in a local optimum. Smaller values for α will result in more exhaustive exploration of the capacity domain, same as a greater value of ϑ_{max}.

Chapter 4

Incentive and Policy Implications: Case Study on Spain

We apply our methodology for the multiple facility problem under a market selection setting on a realistic network representing a potential future market for hydrogen in Spain. We use this practical setting to understand and discuss the impact of incentive and policy choices on the design of a hydrogen distribution network.

We use the 50 largest cities in Spain as our demand nodes, a list of which is presented in Appendix A. The total target population is 16.4 million people, which amounts to a yearly hydrogen demand of 204,000 kg of hydrogen per hour at 100% local market share (i.e., full replacement of all motor vehicles by fuel cell vehicles). As the renewal rate of the passenger vehicle fleet is approximately 5.4% (European Commission 2012c), we consider it reasonable to use 5% and 20% as short- and medium-term projected market shares for hydrogen-powered vehicle adoption. We point out that Brey et al. (2006) use 15% as an overall target in a medium-term (6 year) horizon, using intermediate targets of 5% and 10% for the earlier stages (2 and 4 years) of technology rollout. Hydrogen demand was derived from demographic and transportation data available from the European Commission (2012c), considering an average hydrogen fuel efficiency of 78.8 km/kgH2 (U.S. Department of Energy 2007), resulting in respective target hydrogen demands of 10,209 and 40,836 kg per hour for 5% and 20% share of the vehicle inventory.

Energy supply and mean price data was obtained from European Commission (2012c) and OMEL (2011), and we assumed the standard deviation of electricity prices to be 40% of the mean price. We use the Euclidean distances between demand nodes, and an average cost of 0.00743 €/kg to obtain our transportation cost values. The rest of the parameters (here defined for a one-hour period) are gathered from the H2A and HyWays frameworks for hydrogen infrastructure development (U.S. Department of Energy 2009c, European Commission 2012b), and from estimations done as part of the SPHERA project and the existing literature on hydrogen pathways (e.g., Levene et al. 2007). The following parameters remain fixed throughout the case study: $R \sim U[0, 1008334]$ [kWh], $p^E \sim N(0.039, 0.0156)$ [€/kWh], $w = 0.01$ €/kWh, $m = 0.079$ €/kg, and $Q(C_j) = 80 + 0.50C_j^{1.1}$.

4.1 Base Scenarios

We present base scenarios based on the following parameter changes to evaluate their effect on the resulting production-distribution network: three values for hydrogen prices $(3.25, 3.5, 4.0)$ to represent the willingness of the regulator to incentivize production, two levels of market share $(5\%, 20\%)$ to represent the rate of adoption of the new technology, and two production efficiencies (0.01871 and 0.02252 kg/kWh) to model current and potential future yield from the electrolysis process. We implement the branch-and-price algorithm presented in §2.5.1 under a market selection policy, with the results summarized in Table 4.1.

Intuitively, expected profits and coverage reflect a nondecreasing behavior with respect to hydrogen prices and production efficiency. Results indicate largely decentralized production, more evidently in the higher market share scenarios, where previously modest markets become attractive for local electrolysis rather than hydrogen transport from a nearby location. We can use a mapping tool to illustrate the network configuration, an example of which (specifically, the result of base scenario 7) is shown in Figure 4-1. As can be observed, production is significantly

Scenario	Efficiency	Market Share	p^H	p^*	z^*	% Coverage	Cycles/Columns	Run Time
1	Current	5%	3.25	4	51.28	23.0	1 / 51	333
2	Current	5%	3.50	17	1602.6	81.6	6 / 78	3397
3	Current	5%	4.00	29	6232.0	100.0	8 / 101	5396
4	Current	20%	3.25	34	508.9	54.8	2 / 54	761
5	Current	20%	3.50	48	8344.0	100.0	2 / 54	727
6	Current	20%	4.00	48	28140.7	100.0	2 / 54	725
7	Future	5%	3.25	19	2377.0	86.5	8 / 100	4077
8	Future	5%	3.50	28	4726.6	99.2	8 / 100	5534
9	Future	5%	4.00	29	9807.9	100.0	7 / 104	4889
10	Future	20%	3.25	48	12088.0	100.0	2 / 54	739
11	Future	20%	3.50	48	22171.0	100.0	2 / 54	727
12	Future	20%	4.00	48	42526.8	100.0	2 / 54	732

Table 4.1: Performance of branch-and-price algorithm for the (MFMS). p^* is the optimal number of facilities, z^* the optimal profits per hour, coverage is measured with respect to target demand, and run time is in seconds.

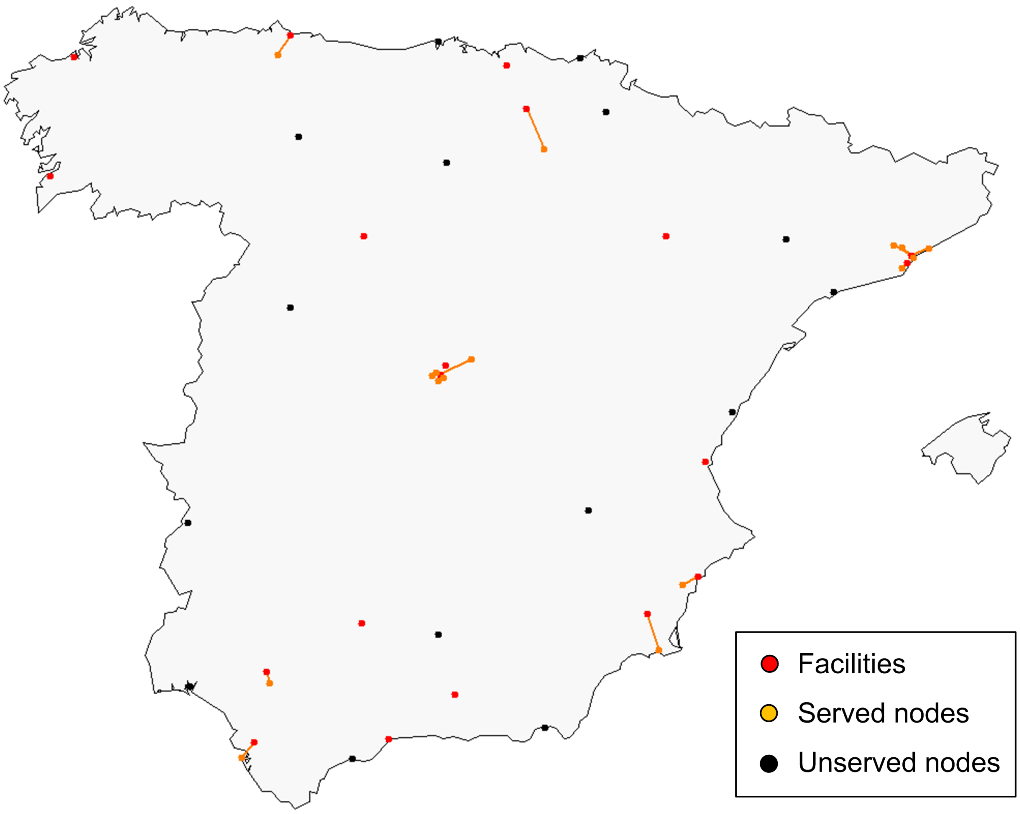


Figure 4-1: Spatial layout of resulting network for base scenario 7. In this result, 19 facilities are opened to serve 35 markets covering 86.5% of estimated demand.

decentralized due to the distance between the population centers. Only some small clusters are formed in regions where the combination of proximity and demand size allows for profitable delivery. We can also see some unserved nodes that are closer to their nearest facility than some served nodes are to theirs. The main reasoning is that the compounded effect of transportation costs and relative size of the nodes makes the threshold for serviceable distance a function of subset size. Finally, we see an instance of a node served by a facility other than its nearest one (in the central portion of the map, specifically the node representing Alcala de Henares being served from the Mostoles facility rather than from Madrid), which is a situation we proved can exist in our model, in opposition to pure cost-minimization models with linear costs such as the p-median problem.

However, to achieve succesful large-scale deployment, the regulator might need to establish an incentive scheme for the producer to assume the risk of investing in capacity, while maintaining hydrogen prices competitive with traditional fuels. It is out of the scope of this thesis to discuss which mechanisms are better for this purpose, but we will briefly present how these incentives could be incorporated to our model, and how this model can be used in comparing such incentive schemes.

4.2 Effect of Incentive Schemes under a Market Selection Policy

We evaluate two different dimensions of monetary incentives for hydrogen production. First, having access to a higher wholesale hydrogen price p^H increases both coverage and producer's profits. Consumers, on the other hand, would more quickly adopt the new technology if the retail price p^T is competitive enough with traditional fuels to offset the higher cost of a new vehicle (i.e., establishing a price equal to that of current fuels might not be sufficient). Thus, the gap between these two prices –at the plant and at the pump— minus any retail costs might need to be covered by the regulator. We refer to this gap $p^H - p^T - c_{ret}$ as the price subsidy.

In contrast, the regulator may cover a fraction $\xi \in [0, 1]$ of the capacity costs. We refer to this incentive as the equipment subsidy. Thus, the producer's profit function will become:

$$Z^{MS}(\mathbf{y}, \mathcal{S}) = \operatorname*{Max}_{\mathcal{S}, \mathbf{y} \in G} \sum_j D_{S_j} K_j^H(\bar{t}_{y_j|S_j}) - (1 - \xi) \sum_j Q\left(C_j^*(D_{S_j})\right)$$

It is evident the regulator's cost function from the combination of these incentiives is $(p^H - p^T - c_{ret})^+ \sum_j D_{S_j} + \xi Q\left(C_j^*(D_{S_j})\right)$, thus costs are nondecreasing in both ξ and p^H.

We set a target retail price for hydrogen of $p^T = 3.50$ €/kg, which would be competitive with existing fuels to offset the additional costs associated to vehicle replacement by consumers. For comparison, note that average retail price for gasoline fuel in Spain for 2011 was 1.318 €/l, or about 5 €/gallon (European Commission 2012a). We also reserve 0.25 €/kg for the cost of retail distribution, which should be sufficient to cover for amortized capital, O&M costs and operating margin of filling stations assuming that existing gas stations can be adapted for hydrogen delivery, thus land costs can be ignored (as a reference, Yang and Ogden 2007 estimate retail point costs excluding land at about $0.20/kg for non-pipeline stations). Thus, for $p^H > 3.25$ €/kg, the regulator will cover the difference between the wholesale and consumer prices.

We can then compare the cost for the regulator (indirectly, for society) of achieving a certain level of coverage for an initial market share of 5%, given different combinations of p^H and ξ. Figure 4-2 graphically represents the coverage level (as a percentage of the target demand) for these incentive schemes, while Table 4.2 presents a comparison of costs incurred by the regulator to achieve full coverage of demand (with an annual production of 89.4M kg), where p^H_ξ is the lowest wholesale price for a given value of ξ that achieves full coverage (due to the monotonicity of the regulator's cost function with respect to p^H, this is also the lowest cost alternative for the regulator given an equipment subsidy level).

Each gallon of gasoline that is replaced by an equivalent amount of hydrogen fuel reduces carbon dioxide emissions by approximately 8.75 kg. Annual production with full geographic coverage at 5% market share would result in over 780,000 tonnes of reduced CO_2 emissions. Thus, we can also present an estimated cost per ton of reduced carbon dioxide emissions, as a way of measuring the potential impact of clean hydrogen, which is approximately 52 €/TCO_2 using an incentive scheme based only on equipment (and not end-product price) subsidy.

α	p_α^H	Profits (M€/yr)	Incentive (€/yr)			Cost (€/TCO_2)
			Price	Equipment	Total	
0%	3.90	45,785,366	58,130,877	0	58,130,877	74.24
10%	3.75	43,276,410	44,716,059	10,509,863	55,225,922	70.53
20%	3.55	36,680,671	26,829,635	21,107,304	47,936,939	61.22
30%	3.40	34,729,861	13,414,818	31,835,206	45,250,024	57.79
40%	3.25	33,179,867	0	42,996,777	42,996,777	54.92
39%	3.25	32,104,948	0	41,921,858	41,921,858	53.54
38%	3.25	31,030,028	0	40,846,939	40,846,939	52.17

Table 4.2: Comparison of regulator incentive costs.

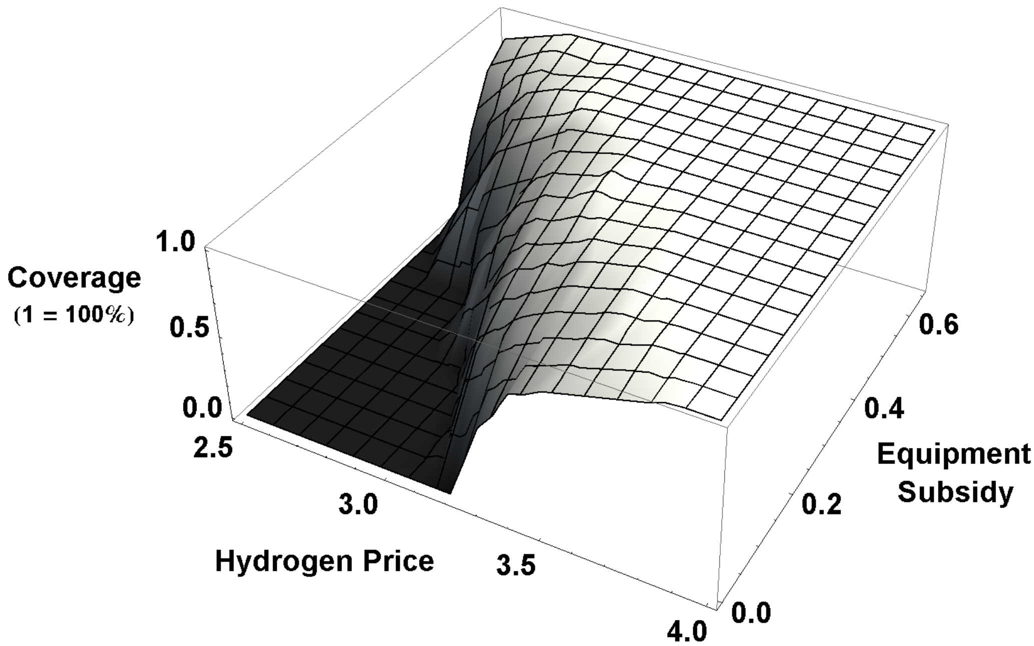
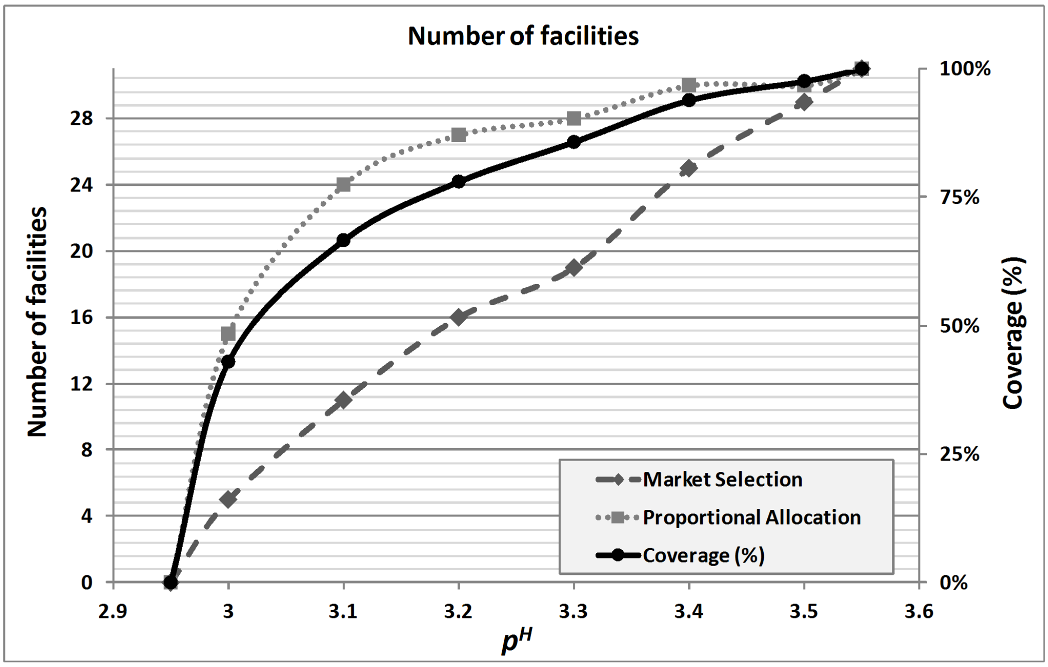


Figure 4-2: Coverage as a function of hydrogen price (p^H) and equipment subsidy (ξ).

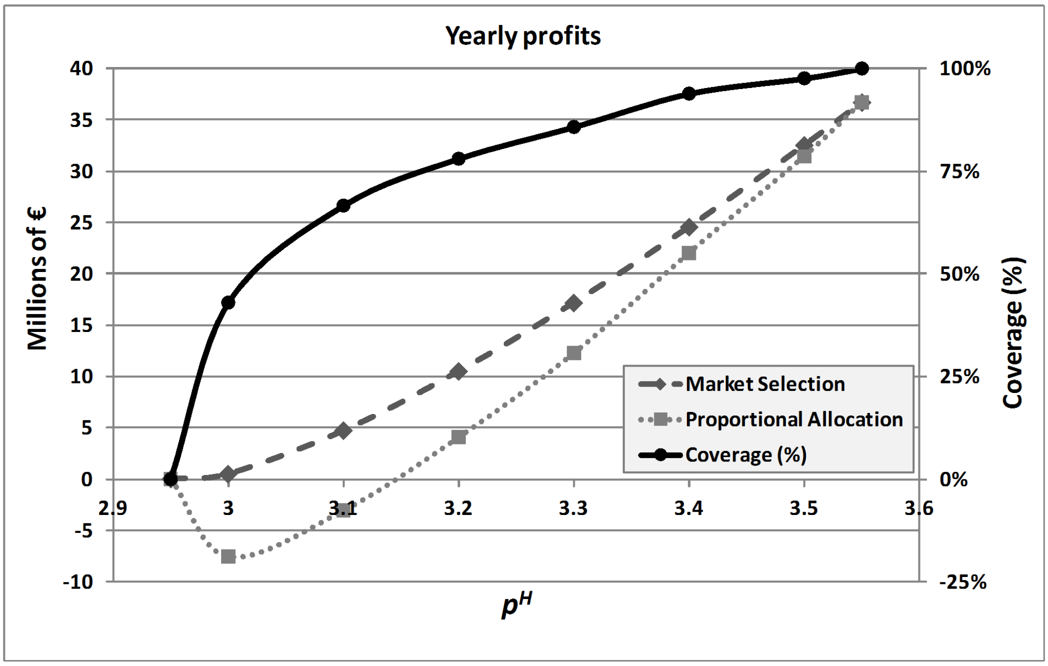
4.3 Effect of Distribution Policy Choice on Network Design and Producer Profitability

We can contrast the firm's performance under both distribution policies. The market selection policy enforces full local coverage of demand, but does not enforce full spatial coverage across the network. The proportional allocation, on the other hand, enforces full spatial coverage and equal (but not full) local coverage across the network.

We can compare firm profitability in the presence of each of these two policies, and measure the impact of inducing spatial equity in the deployment of the new technology for both the firm and the regulator. Given a solution $(\mathcal{S}, \mathbf{y})$ to the (MFMS) problem, we use the resulting spatial coverage (i.e., the total weight of the nodes served) and set it as the local service level parameter ϕ for the (MFPA); hence, we can contrast the policies for a setting with the same overall level of production. Fixing $\xi = 0.20$, we compare the policies for a hydrogen price range $p^H \in [2.95, 3.55]$. Figure 4-3 shows the behavior for the optimal number of plants and the expected yearly profits. The proportional allocation constraint has a significant effect in firm profitability, especially at low levels of adoption, where the firm is forced to forego service to significant portions of highly attractive markets, and divert output towards less attractive nodes. This effect is also evident in the fact that production under the market selection policy is highly focused on a small number of large markets for low hydrogen prices, while the spatial distribution of population centers induces opening a significantly larger number of production facilities for serving the same volume of demand under a proportional allocation policy. For instance, at a hydrogen price of 3.25 €/kg, the proportional allocation policy causes the optimal number of electrolysis plants to increase from 19 to 28 to achieve the same total expected production volume, with the firm suffering a 37.7% reduction in profits, and the regulator incurring 4.9% greater equipment subsidy costs (there is no subsidy of end-product price) due to the larger number of plants requiring fixed investment. In the cases where negative profits results from a proportional allocation policy being in place, the firm would simply decide not to enter the market and obtain zero profits; however the actual expected profits are shown to hint the existence of a threshold price for hydrogen for which a firm would enter a market regulated under a proportional allocation policy and a certain level of equipment incentives. For the case shown, the minimum price for feasible production under the proportional allocation policy is approximately 0.20 €/kg than the corresponding price from the market selection case.



(a)



(b)

Figure 4-3: Comparison of distribution policies as a function of hydrogen price: (a) Number of facilities; (b) Yearly profits. The continuous line shows the fraction of total demand covered for reference.

Chapter 5

Conclusions and Extensions

5.1 Summary of Results

This thesis has addressed the profit maximization problem faced by a firm with renewable energy generation capabilities, when it has the option to produce hydrogen through grid-connected electrolysis for distribution to consumers using hydrogen as an alternative fuel. We show the properties of the producer's profit function, and provide an exact method based on column generation for solving this problem under two alternative distribution policies.

As a way to ground our work in a practical context, we also present a numerical case study that extends previous research on a potential hydrogen supply chain for Spain. Our results indicate a prevalence of decentralized production in the case where the producer can select its markets for service, induced by the relatively long distances between demand points, which is accentuated in high market share and future efficiency scenarios. Additionally, at the current state of production technology, our computational tests indicate that an incentive scheme centered on reducing equipment and infrastructure costs for producers would result in a lower cost for the regulator than an end-product price subsidy, when full spatial and local coverage of demand are desired.

Our work, as far as we know, is the first network design model integrating energy supply and price uncertainty for a firm with dual wind generation and hydrogen production capabilities, and providing an exact solution method that is remarkably efficient and adaptable to existing energy market data. This allowed us to solve multiple scenarios to assess the trade-off between two dimensions of production incentives.

The ability to leverage generation and storage capabilities to maximize profits could prove crucial in increasing the feasibility of hydrogen as an alternative energy carrier, as well as improving the utilization of current (and future) renewable generation infrastructures. Thus, we consider this a valuable tool both for firms who aim to evaluate their network decisions, and also for policy-makers who wish to design fair incentive systems for the production of clean energy. We show how the model can be used to estimate the monetary cost (for the regulator) of emission reduction derived from clean energy delivery for transportation, as well as the cost of assuring equal adoption across a geographic area by means of a proportional allocation policy. We do not attempt to make specific policy suggestions, but rather provide a framework for stakeholders to have a more accurate representation of the outcomes induced by a set of policies, and enabling a proper comparison to other energy storage and delivery technologies.

5.2 Extensions

Many interesting directions arise for extending this work. First, other alternative distribution policies may be explored to balance profitability (for the producer), equity and social cost (for society), and incentive costs (for the regulator); for instance, the inclusion of minimum and/or differentiated service levels per location. Such policies might greatly affect the structure of the problem and require the use of methodologies differing from those proposed in this thesis.

Second, our work is based on a zonal pricing scheme for electricity, an assumption that could be relaxed to consider nodal (and possibly correlated) prices.

On the energy policy front, using the model for a more exhaustive analysis of pricing and regulation effects would be useful to further understand the effect of incentive schemes when technology and market parameters are subject to change. Such changes could arise from improvements in electrolyzer or compression/liquefaction technology, price or availability of water for electrolysis (a relevant matter in some geographic regions, Spain included), or long-term changes in the stochastic behavior of energy prices caused by the introduction of storage capabilities to the energy system; thus resulting in a framework for evaluating resource allocation in the development of clean energy carriers, including technology development, production infrastructure, consumer incentive programs, supply availability, and alternative carbon mitigation strategies.

A direct extension of our model in the field of policy design consists in incorporating carbon emission markets as part of the value functions of the parties involved. If the producer is allowed to receive credit for emission reductions due to displacement of fossil feedstocks from the transportation fuel mix, then the dependence on regulator incentives (subsidies) for sustainable development of a hydrogen production-distribution infrastructure could be reduced.

Üçtuğ et al. (2011) perform a feasibility analysis incorporating carbon trading for a hydrogen production plant using a methane reforming process. They show that carbon trading can be a cost-effective financial tool for hydrogen production plants during the first years of operation, when return of investment is low and the risk assumed by producers is high. Additionally, the results from Aflaki and Netessine (2012) indicate that carbon taxes by themselves might actually discourage investment in renewable energy capacity. Thus, it would be interesting to see the effect of integrating hydrogen (as an end product, rather than only for storage or load-shifting) in their framework to evaluate whether their perceived effect still holds.

Finally, we treat facilities as price-taking, which is reasonable when network congestion is not a significant factor (e.g., due to wind being more prevalent in off-peak periods). As part of the goal of a storage technology is to increase the penetration of renewable power in energy markets, it is likely congestion issues might merit consideration in high wind penetration markets.

Our intuition is that considering the location of generation facilities and network congestion would create a need for some production capacity closer to the wind farms, as such localized storage would punctually decouple the transmission system from the hydrogen distribution system during congestion periods. With wind farms usually located in sparsely populated areas, this localized production might be profitable to serve niche segments of the population, such as smaller urban areas near wind farms that are normally left out of pilot alternative fuel plans, creating a viable outlet for surplus generation in peak periods, while providing social value through adoption of a clean technology in areas otherwise ignored due to their relatively small demand.

Another consequence of relaxing the price-taking assumption is the feedback dynamics caused by an increase in renewable production due to the presence of storage technologies (hydrogen or otherwise) when penetration of renewables is high. Zhou et al. (2012) highlight this issue and suggest analyzing the combined effect of multiple wind farms and storage facilities on electricity prices. Extending our model to make electricity prices (a parameter affecting the location and capacity decisions) dependent on total volume of storage via hydrogen would certainly involve a more complex problem that would not be suitable for solution through the methodologies suggested in this thesis. Given part of the purpose of energy storage is increasing the viability of renewable generation projects, this relationship deserves to be at least considered.

The usefulness of our model goes beyond the field of energy systems. Consider a firm making a strategic production-distribution decision associated to an assortment of goods requiring a shared (possibly perishable) input, where the input itself can be sold to a market with price uncertainty, and where the end goods can be obtained through production processes shared by subsets of products. Such a setting bears some structural similarities to our problem. In that context, the location and capacity decisions relate to the selection of the transformation processes, while the allocation problem would relate to how much of each capacity –if any– would be assigned to each product, with transportation costs representing transformation costs for each process-product pair, and the market selection policy acting as a filtering mechanism for the firm to choose which end-products to produce. The model may be adapted to represent the prices of the different end-products, rather than a unique price as in our case. In the context of different production technologies for industries with inputs of expensive (precious metals for semiconductors), price-volatile (energy-intensive processes) or perishable nature (fresh produce with longer-lasting but lower-value end-products), such a model could certainly be applicable.

Appendix A

List of demand points for case study

City, Province (1-25)	City, Province (26-50)
Madrid, Madrid	Jerez, Cadiz
Barcelona, Barcelona	Pamplona , Navarra
Valencia, Valencia	Fuenlabrada , Madrid
Seville, Seville	Almeria, Almeria
Zaragoza, Zaragoza	San Sebastian, Guipuzcoa
Malaga, Malaga	Santander, Cantabria
Murcia, Murcia	Burgos, Burgos
Bilbao, Viszcaya	Burgos, Burgos
Cordoba, Cordoba	Castello, Castello
Alicante, Alacant	Alcorcon, Madrid
Valladolid, Valladolid	Albacete, Albacete
Vigo, Pontevedra	Salamanca, Salamanca
Gijon, Asturias	Getafe, Madrid
L'Hospitalet De Llobregat, Barc.	Logrono, La Rioja
A Coruã, A Coruña	Huelva, Huelva
Granada, Granada	Badajoz, Badajoz
Vitoria, Alava	Leon, Leon
Badalona, Barcelona	Tarragona, Tarragona
Elx, Alacant	Cadiz, Cadiz
Oviedo, Asturias	Lleida, Lleida
Cartagena, Murcia	Marbella , Malaga
Mostoles, Madrid	Sta. Coloma De Gramenet, Barc.
Alcala De Henares, Madrid	Mataro, Barcelona
Sabadell, Barcelona	Jaen, Jaen
Terrassa, Barcelona	DosHermanas, Seville

Table A.1: List of demand points.

Bibliography

Aflaki, S., S. Netessine. 2012. Strategic Investment in Renewable Energy Sources. INSEAD Working Paper 2011/05/DS/TOM, June 2012.

Almansoori, A., N. Shah. 2006. Design and operation of a future hydrogen supply chain – snapshot model. *Chemical Engineering Research and Design* **84**(A6) 423–438.

Almansoori, A., N. Shah. 2009. Design and operation of future hydrogen supply chain: multi-period model. *International Journal of Hydrogen Energy* **34**() 7883–7897.

An, H., W. Wilhelm, S. Searcy. 2011. Biofuel and petroleum-based fuel supply chain research: A literature review. *Biomass and Bioenergy* **35**(9) 3763–3774.

Anderson, S.T. 2012. The demand for ethanol as a gasoline substitute. *Journal of Environmental Economics and Management* **63**(2) 151–168.

Arrow, K., T. Harris and J. Marschak. 1951. Optimal Inventory Policy. *Econometrica* **19**(3) 250–272.

Balas, E., M. Padberg. 1976. Set Partitioning: A Survey. *SIAM Review* **18**(4) 710–760.

Ball, M., M. Wietschel, O. Rentz. 2007. Integration of a hydrogen economy into the German energy system: an optimising modelling approach. *International Journal of Hydrogen Energy* **32**(10-11) 1355–1368.

Banzhaf, H.S., B.A. Chupp. 2010. Heterogeneous Harm vs. Spatial Spillovers: Environmental Federalism and U.S. Air Pollution. NBER Working Paper 15666, January 2010, revised August 2010.

Barnhart, C., E.L. Johnson, G.L. Nemhauser, M.W.P. Savelsbergh, P.H. Vance. 1998. Branch-and-price: Column generation for solving huge integer programs. *Operations Research* **46**(3) 316–329.

Barquín, J. 2006. Critically-congested Nash-Cournot equilibria in electricity networks. Presented at 9th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS'06), Stockholm, June.

Bento, N. 2008. Building and interconnecting hydrogen networks: Insights from the electricity and gas experience in Europe. *Energy Policy* **36**(8) 3019–3028.

Berman, O., D. Krass. 2002. Facility location problems with stochastic demands and congestion. Z. Drezner, H. Hamacher, eds. *Facility location: Applications and theory*. Chapter 11. Springer-Verlag, New York.

Berman, O., D. Krass, M.B.C. Menezes. 2007. Facility reliability issues in network p-median problems. *Operations Research* **55**(2) 332–350.

Bertsekas, D. (1999). *Nonlinear Programming*. Athena Scientific, Belmont, MA.

Black, M., G. Strbac. 2007. Value of bulk energy storage for managing wind power fluctuations. *IEEE Transactions on Energy Conversion* **22** 197–205.

Boyle, G. 2004. *Renewable energy*. Oxford University Press, Second Edition, New York.

Boyle, G. 2007. *Renewable electricity and the grid: The challenge of variability*. Earthscan Publications Ltd., October 2007.

Brey, J.J., R. Brey, A.F. Carazo, I. Contreras, A.G. Hernández-Díaz, V. Gallardo.2006. Designing a gradual transition to a hydrogen economy in Spain. *Journal of Power Sources* **159**(2) 1231–1240.

Carta, J.A., P. Ramírez and S. Velázquez. 2009. A review of wind speed probability distributions used in wind energy analysis – Case studies in the Canary Islands. *Renewable and Sustainable Energy Reviews* **13** 933–955.

Chen, C.-W., Y. Fan. 2012. Bioethanol Supply Chain System Planning under Supply and Demand Uncertainties. *Transportation Research Part E: Logistics and Transportation Review* **48**(1) 150–164.

Conley, T., C. Udry. 2010. Learning about a New Technology: Pineapple in Ghana. *American Economic Review* **100**(1) 35–69.

Cook, W., W. Cunningham, W. Pulleyblank, A. Schrijver. 1997. *Combinatorial Optimization*. Wiley-Interscience, First Edition.

Dagdougui, H.. 2012. Models, methods and approaches for the planning and design of the future hydrogen supply chain. *International Journal of Hydrogen Energy* **37**(6) 5318–5327.

Dantzig, G. and P. Wolfe. 1960. Decomposition Principles for Linear Programs. *Operations Research* **8**(1) 101–111.

Daskin, M.S. 1995. *Network and Discrete Location: Models, Algorithms, and Applications*. Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, New York, NY.

Dell, R.M., D.A.J. Rand. 2001. Energy storage — a key technology for global energy sustainability. *J. Power Sources* **100**(1-2) 2–17.

Denholm, P., R. Sioshansi. 2009. The value of compressed air energy storage with wind in transmission constrained electric power systems. *Energy Policy* **37**(8) 3149–3158.

De Wolf, D., M.-M. Ould Sidi, J. André, A. Simonnet, G. Maisonnier, J. Brac. 2009. Optimal Design and Dimensioning of Hydrogen Transmission Distribution Pipeline Networks. HEC Discussion Paper 2009 03/01, March 2009.

Drezner, Z., H. Hamacher. 2004. *Facility Location: Applications and Theory*. Springer-Verlag, New York.

European Commission. 2012a. Directorate-General for Energy and Transport. `http://www.ec.europa.eu/energy`.

European Commission. 2012b. HyWays: European Hydrogen Energy Roadmap. `http://www.hyways.de`.

European Commission. 2012c. Statistical Office of the European Union. `http://ec.europa.eu/eurostat`.

European Union. 2009. Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 on the promotion of the use of energy from renewable sources. `http://eur-lex.europa.eu/`.

Farrell, A.E., D.W. Keith, J.J. Corbett. 2003. A strategy for introducing hydrogen into transportation. *Energy Policy* **31**(13) 1357–1367.

Ganslandt, M., P.-J. Norbäck. 2004. Do mergers result in collusion? Working Paper, Research Institute of Industrial Economics, Sweden, June 14, 2004, retrieved October 4, 2009 from SSRN.

Garey, M.R., D.S. Johnson. 1995. *Computers and intractibility: A guide to the theory of NP-completeness*. W.H. Freeman and Co., San Francisco, CA.

Geidl, M., G. Koeppel, P. Favre-Perrod, B. Klöckl, G. Andersson, K. Fröhlich. 2007. Energy hubs for the future. *IEEE Power & Energy Magazine* **5**(1) 24–30.

G. Giannakoudis, A. Papadopoulos, P. Seferlis, S. Voutetakis. 2010. Optimum design and operation under uncertainty of power systems using renewable energy sources and hydrogen storage. *International Journal of Hydrogen Energy* **35**(3) 872–891.

Goentzel, J. 2010. Delivering on the Promise of Green Energy. *Supply Chain Management Rev.* **14**(1) 10–17.

Greiner, C., M. Korpås, T. Gjengedal. 2008. Value of Combining Hydrogen Production with Wind Power in Short-Term Electricity Markets. Presented at the 2nd IEEE International Conference on Power and Energy (PECon 08), Johor Bahru, Malaysia, December.

Global Wind Energy Council. 2008. Global Wind Energy Outlook, Brussels. October 2008. `http://www.gwec.net`. Accessed 28 September 2009.

Hakimi, S.L. 1964. Optimum location of switching centers and the absolute centers and medians of a graph. *Operations Research* **12**(3) 450–459.

Hakimi, S.L. 1965. Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Operations Research* **13**(3) 462–475.

Hale, T., C. Moberg. 2003. Location science research: A review. *Ann. Operations Research*. **123**(1) 21–35.

Han, J.-H., J.-H. Ryu, I.-B. Lee. 2012. Modeling the operation of hydrogen supply networks considering facility location. *International Journal of Hydrogen Energy* **37**(6) 5328–5346.

Hirsch, R.L., R.H. Bezdek, R.M. Wendling. 2005. Peaking of World Oil Production: Impacts, Mitigation and Risk Management. U.S. Department of Energy, National Energy Technology Laboratory (NETL). February 2005.

Huang, Y., C.W. Chen, Y. Fan. 2010. Multistage optimization of the supply chains of biofuels. *Transportation Research Part E: Logistics and Transportation Review* **46**(6) 820–830.

Huang, Y., Y. Fan, N. Johnson. 2010. Multistage System Planning for Hydrogen Production and Distribution. *Networks and Spatial Economics* **10**(4) 455–472.

Huggins, R. 2010. *Energy Storage*. Springer, First Edition.

Hughes, J.E., C.R. Knittel, D. Sperling. 2008. Evidence of a shift in the short-run price elasticity of gasoline demand. *The Energy Journal* **29**(1) 113–134.

Ibrahim, H., A. Ilinca, J. Perron. 2008. Energy storage systems – Characteristics and comparisons. *Renewable and Sustainable Energy Reviews* **12** 1221–1250.

Jørgensen, C., S. Ropenus. 2008. Production price of hydrogen from grid connected electrolysis in a power market with high wind penetration. *International Journal of Hydrogen Energy* **33**(20) 5335–5344.

Kim, J., Y. Lee, I. Moon. 2008. Optimization of a hydrogen supply chain under demand uncertainty. *International Journal of Hydrogen Energy* **33**(18) 4715–4729.

Kim, J.H., W.B. Powell. 2011. Optimal Energy Commitments with Storage and Intermittent Supply. *Operations Research* **59**(6) 1347–1360.

Klose, A., A. Drexl. 2005. Facility location models for distribution system design. *Eur. J. Operations Research* **162**(1) 4–29.

Korpås, M. and C. Greiner. 2008. Opportunities for hydrogen production in connection with wind power in weak grids. *Renewable Energy* **33**(6) 1199–1208.

Labbé, M., D. Peeters, J.-F. Thisse. 1995. Location on Networks. M. Ball, T. Magnanti, C. Monma, G. Nemhauser, eds. *Handbooks in Operations Research and Management Science, Volume 8: Network Routing*. North-Holland, Amsterdam.

Larson, R., A. Odoni. 2007. *Urban Operations Research*. Second edition, Dynamic Ideas, Belmont, MA.

Levene, J.I., M.K. Mann, R.M. Margolis, A. Milbrandt. 2007. An analysis of hydrogen production from renewable electricity sources. *Solar Energy* **81** 773–780.

Lewis, D.J., B.L. Barham, B. Robinson. 2011. Are There Spatial Spillovers in the Adoption of Clean Technology? The Case of Organic Dairy Farming. *Land Economics* **87**(2) 250–267.

Lewis, D.J., B.L. Barham, K.S. Zimmerer. 2008. Spatial Externalities in Agriculture: Empirical Analysis, Statistical Identification, and Policy Implications. *World Development* **36**(10) 1813–1829.

Louveaux, F.V. 1986. Discrete stochastic location models. *Annals of Operations Research* **6**(2) 23–34.

Lübbecke, M.E., J. Desrosiers. 2005. Selected topics in column generation. *Operations Research* **46**(3) 316–329.

Martello, S., P. Toth. 1990. *Knapsack Problems: Algorithms and Computer Implementations*. Wiley Series in Discrete Mathematics and Optimization, John Wiley & Sons, Chichester-New York.

Melo, M.T., S. Nickel, F. Saldanha-da-Gama. 2009. Facility location and supply chain management - A review. *Eur. J. Operations Research* **196**(2) 401–412.

Menanteau, P., M.M. Quéméré, A. Le Duigou, S. Le Bastard. 2011. An economic analysis of the production of hydrogen from wind-generated electricity for use in transport applications. *Energy Policy* **39**(5) 2957–2965.

Minoux, M. 1987. A class of combinatorial problems with polynomially solvable large scale set covering/set partitioning relaxations. *RAIRO* **21**(2) 105–136.

Mirchandani, P.B. 1980. Locational decisions on stochastic networks. *Geographical Analysis* **12**(2) 172–183.

Mirchandani, P.B. 1990. The p-median problem and generalizations. P.B. Mirchandani, R. Francis, eds. *Discrete location theory*. Chapter 2. John Wiley & Sons, Inc., NewYork, NY. – include only if p-median heuristic is included

Nemhauser, G., L. Wolsey. 1999. *Integer and Combinatorial Optimization*. Wiley-Interscience.

The Organisation for Economic Co-operation and Development (OECD). 2008. Linking Renewable Energy to Rural Development: Drivers and Constraints. OECD Regional Development Policy Conference, Paris, 14-15 June 2012.

Ogden, J. 1999. Prospects for building a hydrogen energy infrastructure. *Annual Review of Energy and the Environment* **24** 227–279

Ogden, J., M. Steinbugler, T. Kreutz. 1999. A comparison of hydrogen, methanol and gasoline as fuels for fuel cell vehicles. *Journal of Power Sources* **79**(2) 143–168.

OMEL. 2011. OMEL Annual Report. `http://www.omel.es`. Accessed 29 June 2011.

Owen, A.D. 2004. Environmental Externalities, Market Distortions and the Economics of Renewable Energy Technologies. *The Energy Journal* **25**(3) 127–156.

Owen, A.D. 2006. Renewable energy: Externality costs as market barriers. *Energy Policy* **34**(5) 632–642.

Papapostolou, C., E. Kondili, J.K. Kaldellis. 2011. Development and implementation of an optimisation model for biofuels supply chain. *Energy* **36**(10) 6019–6026.

Parker, N., Y. Fan, J. Ogden. 2010. From waste to hydrogen: An optimal design of energy production and distribution network. *Transportation Research Part E: Logistics and Transportation Review* **46**(4) 534–545.

Rivier, M., I.J. Pérez-Arriaga. 1993. Computation and decomposition of spot prices for transmission pricing. 11th PSCC Power Systems Computation Conference (PSCC), Avignon, France, 30 August – 3 September, 1993.

Rogers, E. 1995. *Diffusion of Innovations*. Simon & Schuster, Fourth Edition, New York.

Rogers, J., S. Fink, K. Porter. 2010. Examples of Wind Energy Curtailment Practices, National Renewable Energy Laboratory Subcontract Report NREL/SR-550-48737, July 2010.

Rogner, H.-H. 1998. Hydrogen technologies and the technology learning curve. *International Journal of Hydrogen Energy* **23**(9) 833–840.

Saran, P., Goentzel, J., Siegert, C. 2010. Economic Analysis of Wind Plant and Battery Storage Operation using Supply Chain Management Techniques. *Proceedings of the 2010 IEEE Power & Energy Society General Meeting*.

Schrijver, A. 2003. *Combinatorial Optimization: Polyhedra and Efficiency*. Springer, First Edition.

Schwoon, M. 2006. Simulating the Adoption of Fuel Cell Vehicles. Working Paper FNU-59, Research unit Sustainability and Global Change, Hamburg University, revised February 2006.

Schwoon, M. 2006.Learning-by-doing, Learning Spillovers and the Diffusion of Fuel Cell Vehicles. Working Paper FNU-112, Research unit Sustainability and Global Change, Hamburg University, revised June 2006.

Sioshansi, R. 2010. Welfare impacts of electricity storage and the implications of ownership structure. *The Energy Journal* **31** 189–214.

Snyder, L.V. 2006. Facility location under uncertainty: a review. *IIE Transactions* **38**(7) 537–554.

Stoft, S. 2002. *Power System Economics*. Wiley-IEEE Press, First Edition.

Sundararagavan, S., E. Baker. 2012. Evaluating energy storage technologies for wind power integration. *Solar Energy* **86**(9) 2707–2717.

Turner, J.A. 1999. A Realizable Renewable Energy Future. *Science* **285** 687–689.

Tzimas, E., P. Castello, S. Peteves. 2007. The evolution of size and cost of a hydrogen delivery infrastructure in Europe in the medium and long term. *International Journal of Hydrogen Energy* **32**(10-11) 1355–1368.

Üçtuğ, F.G., M. Melikoğlu, C.A. Naseh. 2011. Feasibility Analysis of Carbon Trading for Hydrogen Production from Natural Gas. SET2011, 10th International Conference on Sustainable Energy Technologies, Istanbul, 4-7 Sep. 2011.

U.S. Department of Energy. 2007. National Renewable Energy Laboratory, Fact Sheet NREL/FS-560-42284, October 2007. `http://www.nrel.gov`.

U.S. Department of Energy. 2009a. U.S. DOE webpage, Distributed Energy Program: Compressed Air Energy Storage. `http://www.energy.gov/energysources/hydrogen.htm`. Accessed 28 September 2009.

U.S. Department of Energy. 2009b. U.S. DOE webpage, Energy Sources: Hydrogen. `http://www.eere.energy.gov/de/compressed_air.html`. Accessed 28 September 2009.

U.S. Department of Energy. 2009c. U.S. DOE webpage, Hydrogen Analysis (H2A) Project. `http://www.hydrogen.energy.gov/h2a_analysis.html`. Accessed 26 January 2010.

U.S. Department of Energy. 2012. Market Brief: Status of the Voluntary Renewable Energy Certificate Market (2011 Data). National Renewable Energy Laboratory, Technical Report NREL/TP-6A20-56128, September 2012. `http://www.nrel.gov`.

Wald, M. 2007. "Utility will use batteries to store wind power". The New York

Times, September 11, 2007. `http://www.nytimes.com/2007/09/11/business/11battery.html`. Accessed November 7, 2009.

Wang, M. 2002. Fuel choices for fuel-cell vehicles: well-to-wheels energy and emission impacts. *Journal of Power Sources* **112** 307–321.

Wang, Y.-W., C.-R. Wang. 2010. Locating passenger vehicle refueling stations. *Transportation Research Part E: Logistics and Transportation Review* **46**(5) 791–801.

Weaver, J., R. Church. 1983. Computational Procedures for Location Problems on Stochastic Networks. *Transportation Science* **17**(2) 168–180.

Welch, A. Venkateswaran. 2009. The dual sustainability of wind energy. *Renewable and Sustainable Energy Reviews* **13** 1121–1126.

Yang, C., J. Ogden. 2007. Determining the lowest-cost hydrogen delivery mode. *International Journal of Hydrogen Energy* **32**(2) 268–286.

Zhou, Y., A. Scheller-Wolf, N. Secomandi, S. Smith. 2012. Managing Wind-based Electricity Generation in the Presence of Storage and Transmission Capacity. Tepper Working Paper 2011-E36, July 2012.