2011

A multiobjective optimization approach to the operation and investment of the national energy and transportation systems

Eduardo Ibanez
Iowa State University

Recommended Citation

Ibanez, Eduardo, "A multiobjective optimization approach to the operation and investment of the national energy and transportation systems" (2011). *Graduate Theses and Dissertations*. 11948.
https://lib.dr.iastate.edu/etd/11948

A multiobjective optimization approach to the operation and investment of the national energy and transportation systems

by

Eduardo Ibanez

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Electrical Engineering

Program of Study Committee:
James D. McCalley, Major Professor
Dionysios C. Aliprantis
William Q. Meeker
Sarah M. Ryan
Leigh Tesfatsion

Iowa State University

Ames, Iowa

2011

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to express my thanks Dr. James McCalley for his help and good guidance throughout the graduate program and development of this work. I would like to thank my committee members for their efforts and contributions: Dr. Dionysios Aliprantis, Dr. William Meeker, Dr. Sarah Ryan, and Dr. Leigh Tesfatsion. I also would like to recognize the contributions from the faculty and students in the NETSCORE-21 research project.

I wish to acknowledge the financial support from the Department of Electrical and Computer Engineering at Iowa State University, the National Science Foundation[1], and the Government of Navarra. Without it, I would not have been able to take part in the graduate program.

I cannot find words to express my most sincere gratitude to my parents, family and friends for their continuous love and support.

Last but not least, I would like to dedicate this thesis to my wife Jessica and thank her for being my rock and helping me throughout the endless hours that led to the completion of this work. You make me feel like the luckiest man.

---

[1] This material is based upon work supported by the National Science Foundation under Grant No. 0835989.

CHAPTER 1. OVERVIEW

1.1 Motivation

Most U.S. energy usage is for electricity production and vehicle transportation, two interdependent infrastructures. The strength and number of these interdependencies will increase rapidly as hybrid electric transportation systems, including plug-in hybrid electric vehicles and hybrid electric trains, become more prominent. There are several new energy supply technologies reaching maturity, accelerated by public concern over global warming. The U.S. Energy Information Administration [1] suggests that national expenditures on electric energy and transportation fuels over the next 20 years will exceed $14 trillion, four times the 2010 federal budget [2]. Intentional and strategic energy system design at the national level will have very large economic impact.

The proposed work is motivated by a recognition that tools, knowledge, and perspective are lacking to design a national system integrating energy and transportation infrastructures while accounting for interdependencies between them, new energy supply technologies, sustainability, and resiliency.

The energy system is comprised by (but not limited to) electricity, natural gas, liquid fuels, nuclear, biomass, hydroelectric, wind, solar, and geothermal resources. Modeling of national freight and passenger transportation focuses on state-to-state travel; we consider both infrastructures (rail, highways, locks/dams, roads, ports, airports) and fleets (trains, barges, trucks, personal vehicles, airplanes, etc.), and there may be different kinds of fleets for each mode (e.g., diesel trains and electric trains or conventional and plug-in hybrid electric).

1.2 NETPLAN

The National Energy and Transportation Planning Tool (NETPLAN) is the implementation of a long-term investment and operation model for the transportation and energy system. An evolutionary approach with underlying fast linear optimization are in place to determine the best investment portfolios in terms of cost, resiliency and sustainability, i.e., the solutions that form the Pareto front.

Fig. 1.1 captures the scope of the modeling effort. The transportation and energy systems interact mainly at two stages: operation and investment. At the operational level each system needs to satisfy its demand with the existing capacity. However, operation of the two systems, and ultimately investment, are interdependent; while the transportation sector demands energy in the form of fuel, the energy sector requires the movement of raw bulk energy sources (e.g. coal or natural gas for thermal power plants). At the same time, the cost of meeting those reciprocal demands has an impact on final prices for energy and transportation.

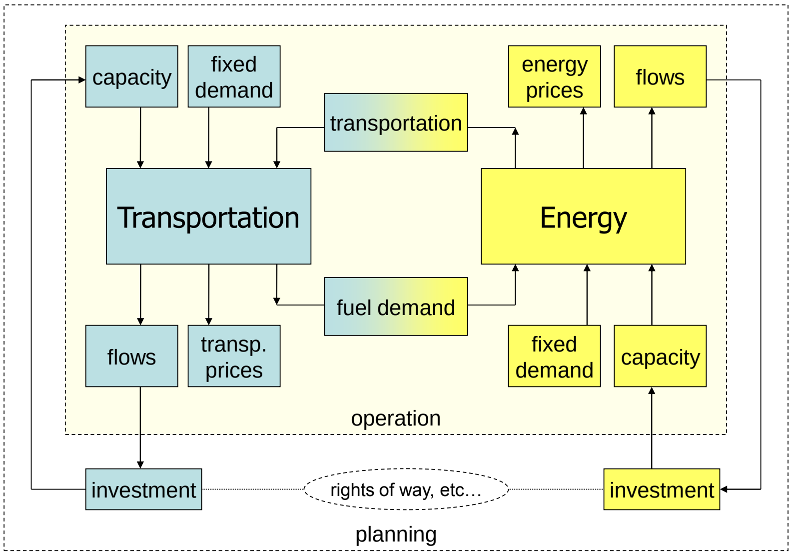


Figure 1.1: Proposed model that integrates the energy and transportation systems at two levels: operation and planning.

The ever-growing public need for energy and transportation creates the necessity to invest in new capacity. Given the potential for increased coupling between energy and transportation, it is apparent that better designs of both can be achieved if these designs are performed together. New investments will determine the behavior of the operational level. Additional dependencies can be expressed at this level. Consider for example the possible competing for right of way of new high-voltage transmission lines and new additions to the rail or highway systems.

This research is part of an ongoing project known as the "21st Century National Energy and Transportation Infrastructures Balancing Sustainability, Costs, and Resiliency" or NETSCORE-21 for short, funded by the National Science Foundation. Our goal is to identify optimal infrastructure designs in terms of future power generation technologies, energy transport and storage, and hybrid-electric transportation systems, with balance in sustainability, costs, and resiliency.

1.3 Objectives

This work proposes the definition of a modeling framework that will allow the study and optimization of the national energy and transportation portfolios. The utilization of multiobjective algorithms enables exploring and evaluating candidate investment portfolios in terms of cost, resiliency and sustainability. This optimization process is time and resource consuming so computational enhancements are presented and developed through decomposition and parallelization. More specifically, the objectives of this work can be summarized as:

1. Develop a modeling framework capable of representing the operation and investment of the energy and transportation systems.
2. Design a multiobjective optimizer that allows the exploration of the space of investments and identification of the Pareto front of solutions.
3. Study and develop resiliency metrics to test candidate portfolios
4. Enhance the computational performance of the mentioned models by use of decomposition and parallelization methods
5. Produce a software application capable of performing the previous tasks, while being simple to use and accessible for future development.

1.4 Literature Review

This section will review the tools and algorithms that are being used by industries and governments. Common notions for resiliency will be also introduced.

1.4.1 Planning tools

The following is a collection of modeling tools that are used to study the energy and transportation systems, both together and independently. An extensive description of the issues related to modeling and more information on methodology classification can be found in [3] and [4].

The only planning tools that expand beyond one sector and are most closely related to the scope of NETPLAN are the National Energy Modeling System (NEMS) [5] and the MARKAL/TIMES suite [6]. NEMS is an equilibrium model (it only evaluates proposed portfolios) which represents the petroleum, natural gas, coal, and electric supply and transport/transmission sectors together with energy/economy macroeconomic interactions and international energy activity. MARKAL/TIMES, an optimization model, represents resources associated with the petroleum, natural gas, coal, and electric sectors but not the transport/transmission associated with these sectors. In addition, whereas NEMS iterates to equilibrium through sequential solutions of each sector, MARKAL/TIMES minimizes total cost in a single optimization.

There are three main types of planning tools for electric infrastructure: reliability, production costing, and resource optimization. Reliability tools do not identify solutions but just evaluate them. Both deterministic and probabilistic tools exist and are heavily used in the planning process. Deterministic tools include power flow, stability, and short-circuit programs, providing answers for specified conditions. Probabilistic tools compute indices such as loss-of-load probability, loss of load expectation, or expected unserved energy, associated with a particular investment plan. A representative list of commercial-grade reliability evaluation models include CRUSE [7], MARS [8], TPLAN [9], and TRELSS [10].

Production cost programs have become the workhorse of long-term planning. These programs perform chronological optimizations, often hour-by-hour, of the electric system operation, where the optimization simulates the electricity markets, providing an annual cost of producing energy. Although production cost models make use of optimization for performing hourly dispatch, but not for selection of infrastructure investments. Production cost models are equilibrium/evaluation models. A representative list of commercial grade production cost models include GenTrader [11], MAPS [12], GTMax [13], ProMod [14], and ProSym [15], and Plexos [16]. Production cost programs usually incorporate one or more reliability evaluation methods.

Resource optimization models select a minimum cost set of generation investments from a range of technologies and sizes to satisfy constraints on load, reserve, environmental concerns, and reliability levels; they usually incorporate a simplified production cost evaluation, which includes a reliability evaluation. These models, as optimization models, identify the best generation investment subject to the constraints. However, at this point in time, most of these models do not optimize transmission investments. A representative list of resource optimization models includes EGEAS [17], Plexos [16], Strategist [18], and WASP-IV [19].

The U.S. Environmental Protection Agency (EPA) utilizes another resource optimization model, the Integrated Planning Model (IPM) [20], wich is a linear programming model of the electric sector used to evaluate the projected impact of environmental policies at a national level. It provides forecasts of least-cost capacity expansion, electricity dispatch, and emission control strategies for meeting energy demand and environmental, transmission, dispatch, and reliability constraints. IPM can be used to evaluate the cost and emissions impacts of proposed policies to limit emissions of sulfur dioxide (SO_2), nitrogen oxides (NO_X), carbon dioxide (CO_2), and mercury (Hg) from the electric power sector.

The Regional Energy Deployment System (ReEDS) model [21] is a multiregional, multiperiod linear programming model of capacity expansion in the electric sector of the United States. The model, developed by the National Renewable Energy Laboratory, performs capacity expansion but with detailed treatment of the full potential of conventional and renewable electricity generating technologies as well as electricity storage. The principal issues addressed include access to and cost of transmission, access to and quality of renewable resources, the variability of wind and solar power, and the influence of variability on the reliability of the grid.

Modeling on the transportation side include national freight forecasting models and tools. The U.S. Department of Transportation has developed a framework analysis that "integrates data from a variety of sources to create a comprehensive picture of freight movement among states and major metropolitan areas by all modes of transportation" [22]. A planning methodology with an interactive graphic interface is presented in [23]. Other transportation investment planning tools include the Highway Economic Requirement System (HERS-ST) [24] for highway investments, and RailDec [25] for rail investments. A statewide passenger travel forecasting model was developed by the transportation Research Board in [26].

1.4.2 Resiliency

The concept of resiliency is ubiquitous and can be found in the literature of many technical and non-technical fields. However, there is no consensus on its definition. In this section, we present definitions used within various disciplines.

In computer network systems, for instance, resiliency is defined as "the ability to provide and maintain an acceptable level of service in the face of faults and challenges to normal operation." Common elements of resiliency include the support of distributed processing and networked storage, as well as the ability to maintain service of communication technologies such as video conferencing, instant messaging, and other online collaboration. In a more general sense, resiliency in computer systems can be thought of as the ability to access applications and data as needed [27].

In the communications industry, resiliency is defined as the assurance that "the interoperability of communications systems is not affected by known or unknown circumstances of change." It is further defined as the ability to evolve and advance as new technologies and capabilities are developed. Alternatively, resiliency can be considered "a reflection of the flexibility of the system to respond to changes in operational requirement, or implementation strategies and technologies" [28].

Within the nuclear power industry, resiliency is defined as "the ability of an organization (system) to maintain, or recover quickly to, a stable state, allowing it to continue operations during and after a major mishap or in the presence of continuous stress." Common elements of a resilient nuclear power system include items such as continuous feedback and monitoring of critical systems, as well as a consistent plan of communications and syntax to minimize human error [29].

One definition of resiliency commonly used in the process control industry is the ability of a system to return to its original (or desired) state after being disturbed. According to this model, risk management is viewed as a central component to resiliency. Supply chain acceleration and deceleration, as well as a suitable mix of lean and agile processes, are also critical to building a resilient system. Ultimately, the goal of a resilient system in process control is to minimize output variability across all possible scenarios [30].

Within the aerospace industry, the general definition of resiliency is very similar to that used in other areas of engineering – resiliency is the ability to change when a force is enacted, as well as the ability to perform adequately or optimally while the force is in effect. Finally, resiliency also dictates the return to a predefined normal state when the force relents or is rendered ineffective [31].

There are fields beyond engineering that make extensive use of resiliency metrics and definitions within the scope of their societal roles. One such industry is health care. Resiliency-related terminology includes buffer capacity, flexibility vs. stiffness, margin, tolerance, and cross-scale interactions. Buffer capacity is defined as the size or kinds of disruptions the system can absorb or adapt to without a fundamental breakdown in performance or in the system's structure . Flexibility vs. stiffness is defined as the system's ability to restructure itself in response to external changes or pressures. Margin is defined as where the system is currently operating relative to one or another kind of performance boundary. Tolerance is defined as how a system behaves near a boundary – that is, whether the system gracefully degrades as stress increases or collapses quickly when stress exceeds adaptive capacity. Cross-scale interactions are defined as system interactions and effects that occur when changes are made on a microscopic or macroscopic level within the system [32].

Finally, we observe that in bulk power systems, the North American Electric Reliability Corporation (NERC) presents an important definition: “reliability is the degree to which the performance of the elements of that system results in power being delivered to consumers within accepted standards and in the amount desired. The degree of reliability may be measured by the frequency, duration, and magnitude of adverse effects on consumer service” [33]. Reliability of a power system is also referred to as “the probability of its satisfactory operation over the long run. It denotes the ability to supply adequate electric service on a nearly continuous basis, with few interruptions over an extended time period” [33].

A recent paper proposes independent definitions for resiliency and robustness, claiming that they cover different aspects of system performance under stress. Robustness is defined as “the ability of a system to maintain its function when it is subject to a set of perturbations of a given class, which may induce changes in its structure.” Then, resiliency is defined as “the ability of a system to (i) gracefully degrade its function by altering its structure in an agile way when it is subject to a set of perturbations and (ii) quickly recover it once the perturbations ceased” [34].

1.5 Thesis Organization

The present thesis is organized in 6 chapter. Chapter 1, *Overview*, describes the motivation behind this work as well as the objectives and literature review. Chapter 2, *NETPLAN*, introduces the principles that the modeling effort is built upon by describing how the operation and investment of the energy, transportation and passenger networks are treated. Chapter 3, *Mathematical formulation*, contains a formal description of the the principle behind NETPLAN and the algorithms used to perform multiobjective optimization, decomposition and parallelization. Chapter 4, *Model implementation*, lists the assumptions, networks, parameters used in the numerical example and the references used to obtain them. Chapter 5, *Study results*, presents the most important findings that are derived from the implementation of the national investment problem and summarizes the Pareto front of solution for investment portfolios. Finally, concluding remarks, list of contributions and directions of further research follow in Chapter 6, *Conclusions.*

CHAPTER 2. NETPLAN

This chapter describes the main concepts that form part of NETPLAN, the "National Energy and Transportation Planning" tool. The main objective is to be able to simultaneously evaluate different portfolios for those critical infrastructures and search for the best ones in terms of sustainability, resiliency and cost.

The model vision is captured in Fig. 2.1, where the two systems are represented, along with their interactions. The energy network is represented in the upper portion and it is responsible to supply the energy demand in the different subsystems (electric, natural gas, liquefied fuels). To this end, the available primary energy resources are transported and converted to the appropriate form of energy. The transportation network, pictured at the bottom, provides the appropriate infrastructure so that the movement of passengers and freight nationwide can be done in a cost-effective and timely way.

The two systems interact creating mutual demands. The transportation network requires fuel to accomplish the movement of freight and passengers around the country. This creates an additional loading of the energy system that must be satisfied. At the same time, the flow along the energy system requires the transportation of the so-called *energy commodities*, which increases the need for more transportation infrastructure. We define energy commodities for the purpose of this dissertation as those commodities that are transported by non-exclusive means and are used, at least partially, to satisfy energy demand or to be converted to other forms of energy. For example, natural gas is not considered an energy commodity since it is transported through a dedicated pipeline system, but coal is because is transported by barge or rail, the same modes of transportation utilized for grain or other commodities. Grain would only considered to be an energy commodity if we consider the conversion to biofuels. Otherwise,

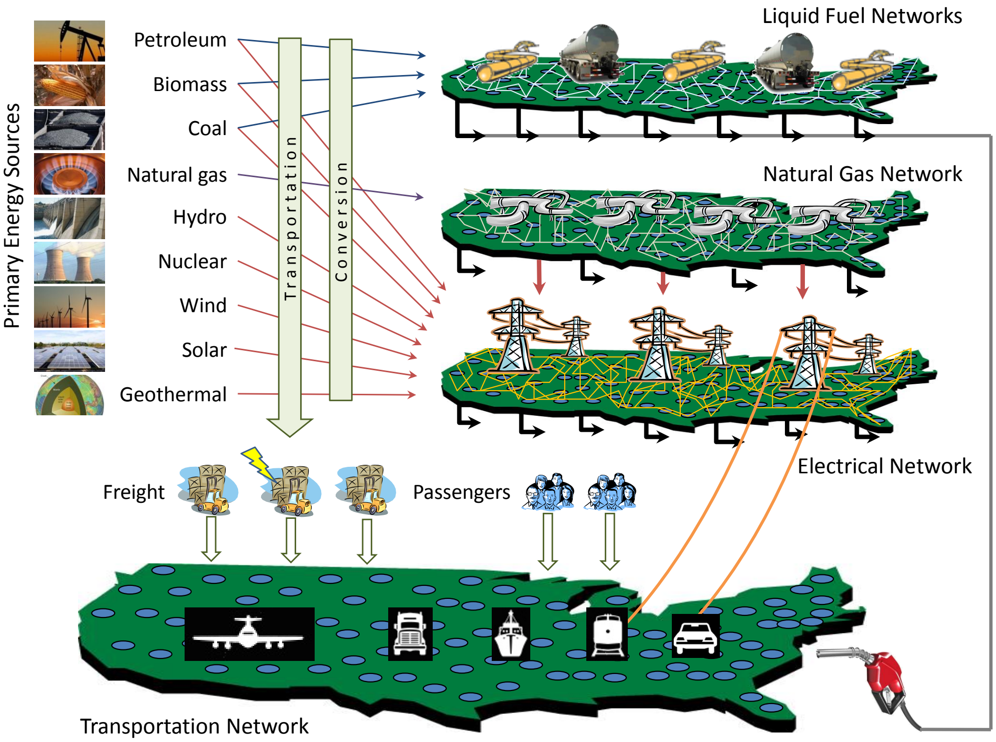


Figure 2.1: Conceptual representation of the energy and transportation systems.

it would just be a regular commodity.

The following sections describe in more detail the different portions of the modeling framework.

2.1 Energy system modeling

The energy system in NETPLAN does not only include the satisfaction of energy demand in form of electricity, natural gas or refined gasolines, but also the production and transportation from the raw energy sources and their conversion to the ultimate consumption form. Primary energy sources could be depletable (petroleum, coal, natural gas, uranium) or renewable (wind, sunlight, biomass, water in dams, tides). These sources are often converted through facilities such as power plants, wind or solar farms, electrolysis processes, or refineries. These processes lead to secondary forms of energy such as electricity, gasoline, diesel or hydrogen. It is also the case that resources are not available close to their ultimate source of consumption so mechanisms such as freight transport, pipeline systems or the transmission grid play an important role in the energy system [35].

This system can be modeled using a generalized network structure, in which energy is considered the only commodity and flows from sources to sinks. This structure enforces a number of concepts that are compatible to the energy system, such as the principle of conservation of energy or flow limits. In this case, flow limits act as the capacity of the different pieces of the infrastructure and by including these in the decision variables we can explore the possibility of investing on new infrastructure.

The following subsections describe in more detail how the network structure is applied to the energy system.

2.1.1 Nodes

Nodes are used to represent the points in the system where the conservation of energy (or flow) is enforced. We can classify them according to their function in the system. They can be *sources*, which serve to represent the production of raw energy forms; *transmission* nodes in which all the incoming energy is leaving the node; *storage* nodes, which are interconnected in time and allow the flow of energy between consecutive points in time; or *demand* nodes, where physical demand of energy like electricity or natural gas is enforced. It could be the case that a node can fall in more than one category, e.g., an electric transmission node can be part of the distribution of energy to its neighboring regions while satisfying its own demand.

Even though a different node could be used to represent system infrastructure at a very granular level, aggregation can be performed for facilities that share similar characteristics. A node is defined to represent both a geographical region and a type of energy. So if we were to define the natural gas and electricity demands at the same location we would need two separate nodes. This concept is better understood with the description of energy arcs.

The most important parameter associated which each node is the demand of energy, which is the difference between energy coming in and out of the node. For storage and transmission nodes, the demand is usually zero so that incoming energy equals the energy leaving the node. Finally, no demand is defined for source nodes since they are the origin of the energy that is consumed in the system.

Apart from enforcing energy conservation, other constraints can be modeled at the nodes. One example would requiring the generation that supplies energy to an electrical region to be able to meet the summer peak demand. The model could be enhanced to ensure that unused generation can cover spinning and non-spinning reserve requirements or that it has adequate ramping capabilities.

Finally, in order to avoid infeasibilities, demand nodes may not meet the pre-established demand. This situation is of course not desirable so the amount of unserved demand is penalized by giving it a very high cost per unit of energy not served. The advantage of this approach is that the optimization will produce a feasible result in cases when the system is compromised, i.e., when the system is undergoing stress through an event when calculating resiliency.

2.1.2 Arcs

Arcs are used to connect nodes and represent the available routes for flows. They are defined by the origin and destination nodes and can have costs and capacities associated to them. If the nodes belong to the same subsystem they represent transmission of energy geographically. Transmission lines, or natural gas or petroleum pipelines are examples of this. Arcs can also connect nodes from different networks to represent conversion between energy types, such as electric generators, electrolysis facilities or refineries.

Arcs can have costs associated with them that relate to the price of raw energy sources or operation and maintenance costs. Capacities can be enforced if the flow across an arc is limited by physical constraints. As in the case with nodes, arcs can be used to represent individual facilities but in most cases they are equivalent to an aggregate group of similar characteristics.

There is a number of parameters associated with each arc, although not all of them apply in each case. These parameters can assigned values, as it is done in Chapter 4, and are:

- Minimum flow. Usually is zero, but could be positive to represent situations such as electric transmissions contracts.

- Maximum flow, or arc capacity. It is a combination of the initial capacity (which can be set to gradually decrease as initial capacity is retired) and new investments.
- Investment cost, per unit of capacity added to the flow
- Minimum investment
- Maximum investment
- Cost, per unit of flow
- Efficiency parameter, which is a multiplier that is applied to the flow. It serves to represent, for example, losses in gas transportation and electric transmission. It can be used to transform units from one subsystem to another.

Arcs are defined to be unidirectional so energy can only flow from origin to destination. In order to represent the bidirectional nature of pipelines or transmission systems, the connection between nodes in these systems can be decomposed using two different and opposite arcs. Figure 2.2 shows a simple example between two nodes.

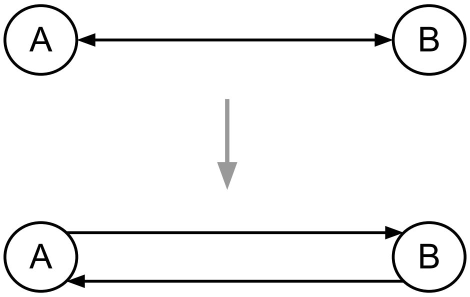


Figure 2.2: Transformation of bidirectional arcs into two directed arcs

2.1.3 Flows

Flows are the representation of energy moving along the system of arcs. Even though energy is the only commodity flowing in the system, different units are used for the different subnetworks: million short tons for coal, million cubic feet for natural gas, million of gallons for the petroleum network, and GW-hours for electricity.

Flows, along with investment variables, form the list of decision variables that will be optimized. A more thorough description can be found in Chapter 3.

2.1.4 Model dynamics

The energy network (as well as the transportation network) is a dynamic system that evolves other time. It is necessary to discretize it in order to be able to analyze it. This process is roughly equivalent to solving the problem at different instances in time.

When selecting a single time step, this has to be small enough to capture the dynamics of the smallest subsystem. However, this could mean that slower parts of the model will be represented with unnecessary detail, wasting computational resources. As a compromise, one can select different step sizes per subsystem in order to capture the essential dynamics in the model while using a more efficient approach.

A graphical description of this concept can be seen in 2.3 where three subsystems are represented. Three levels are used for time steps. The slowest one, the coal subsystem, is represented just for the top level step, e.g., yearly. The next level could be monthly and is applied to the natural gas system. The electric subsystem is the fastest and that is the reason why it is assigned the smallest time step, e.g. hourly.

Investment decisions are assumed to happen with the biggest time size. In the previous figure, investment would be consider in the same time frame that coal operational decisions are made.

2.2 Freight transportation modeling

The freight transport system is modeled as a multicommodity flow network where the flows are in the units of tons of each major commodity. According to the U.S. Bureau of Transportation Statistics [36] there are 23 commodities that comprise 90% of total ton-miles shipped in the country. Out of these, eight represent 55% and are in descending order: coal,

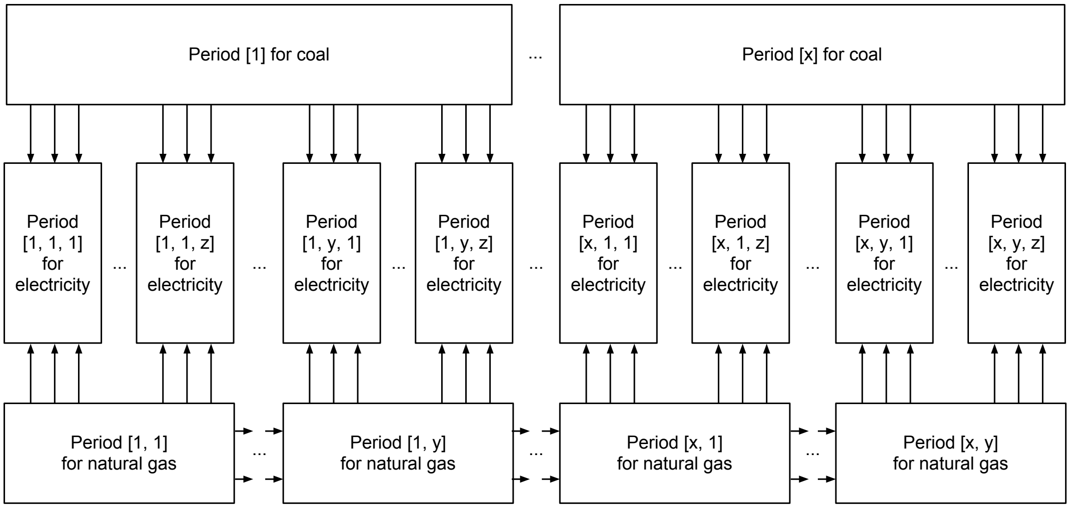


Figure 2.3: Model dynamics and multi-step approach

cereal grains, foodstuffs, gasoline and aviation fuel, chemicals, gravel, wood products, and base metals.

There are two fundamental differences between this formulation and that of the energy formulation. Whereas the energy problem must restrict energy flows of specific forms to particular networks (for example, natural gas or hydrogen cannot move through electric lines or liquid fuel lines), commodities may be transported by more than one transport mode (rail, barge, truck).

Also whereas energy movement requires only infrastructure (electric lines, liquid fuel pipelines, gas pipelines), commodity movement requires infrastructure (rail, locks/dams, roads, ports) and fleet (trains, barges, trucks), and there may be different kinds of fleets for each mode (e.g., diesel trains or electric trains).

For these reasons, the problem is stated as two multicommodity flows [37], one embedded inside the other. Commodities flow through the network formed by the different types of fleet available. At the same time, the units in those fleets travel along the network formed by the different infrastructures. An effective method to convert this situation into an ordinary network problem is captured in Fig. 2.4, where the flow from node A to node B is divided according to the types of infrastructures first and then into the different types of available fleets. Then one can apply capacity limits to the appropriate fictitious arcs.

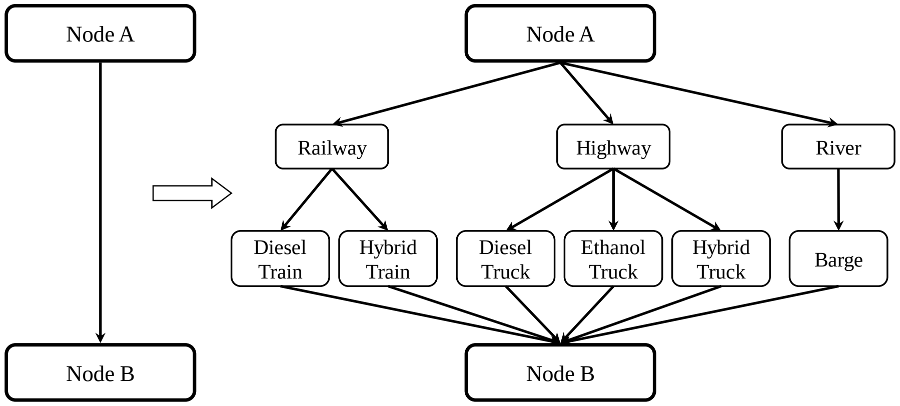


Figure 2.4: Decomposition of transportation arc in two steps: infrastructure and fleet

The size of the multicommodity flow problem increases at a great pace as more commodities and pairs of origin-destination nodes are introduced. Theoretically, each physical commodity sent from every node should be treated as an individual commodity in the model. The size of the problem can be reduced greatly by assuming that the routes for most of the commodities can be predetermined. This assumption preserves the major corridors in the system with the appropriate selection of routes. This feature can be disabled for certain commodities if necessary. One clear example is coal because the production and demand can vary greatly as the energy system is transformed.

The parameters for transportation demand are similar to those of the energy system arcs. The main differences are the lack of an efficiency parameter since the transportation network is supposed to be lossless, and the same units are used throughout the system. The parameters are:

- Minimum flow (which can be zero)
- Maximum flow, or arc capacity. It is a combination of the initial capacity (which can be set to gradually decrease as initial capacity is retired) and new investments.
- Investment cost for fleet and infrastructure, per unit of capacity added to the flow
- Minimum investment, for fleet and infrastructure

- Maximum investment, for fleet and infrastructure
- Cost, per unit of flow

2.3 Passenger transportation modeling

A full passenger model is not developed in NETPLAN and it is beyond the scope of this work to do so. However, an extension of the freight transportation model can be used as a simple representation of the passenger system. Such a model considers the cost of investing in a certain mode of transportation and the cost of operation during its lifetime, including fuel costs.

This is however a very simplistic representation of how the passenger system operates. A more mature model would incorporate some of the following features that can be observed in reality:

- The individual decisions that take place in the passenger model are far more numerous than in the energy and freight systems
- The travel patterns across those individuals vary greatly and appropriate aggregation should be made.
- Although cost is one of the factors in the decision making, it is not the only one. Others should be taken into account, such as total travel time, speed, comfort, independence, brands, and models.
- Air travel presents a challenge because, unlike land-based modes of transportation, the movements do not happen between adjacent regions. Thus, the arc-based approach for freight transport cannot be directly applied. On top of this routes tend to be radial and concentrated in major airports so travel is not guided using the shortest routes.

2.4 Interactions between energy and transportation

2.4.1 Fuel demand

The first interaction between the transportation and the energy systems is due to the fact that fuel is necessary for freight and passenger transportation.

This interaction is easy to understand with the help of Fig. 2.5. In this case the transportation network supports three modes of transportation and the energy network is comprised of natural gas, petroleum and electricity subsystems. The movement of freight using the petroleum vehicle creates a demand of fuel in the corresponding nodes in the petroleum network. Similarly, the electric vehicle causes energy demand on the electric network if it is utilized. The hybrid option follows the same pattern but it creates energy demand in more than one network.

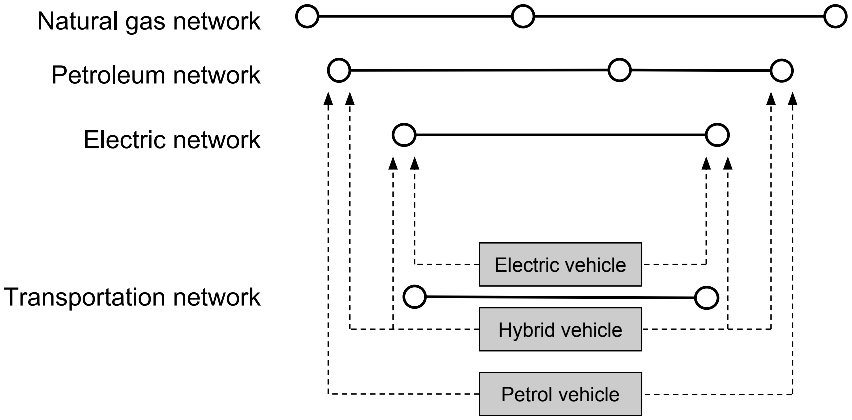


Figure 2.5: Fuel demand on the transportation network

It could be the case that the geographical definition of energy and transportation regions is not the same. Figure 2.6 depicts an example with a transportation network (nodes and arrows) on top of the fuel supply regions (in grey). As it can be seen there are four fuel regions but many more transportation nodes and arcs.

Transportation arcs completely contained in an energy region will get all the necessary fuel from said region, as it is the case for the arcs in the center of the picture. Otherwise, the fuel demand for any transportation arc that crosses the boundaries of two neighboring energy regions will be split among them.

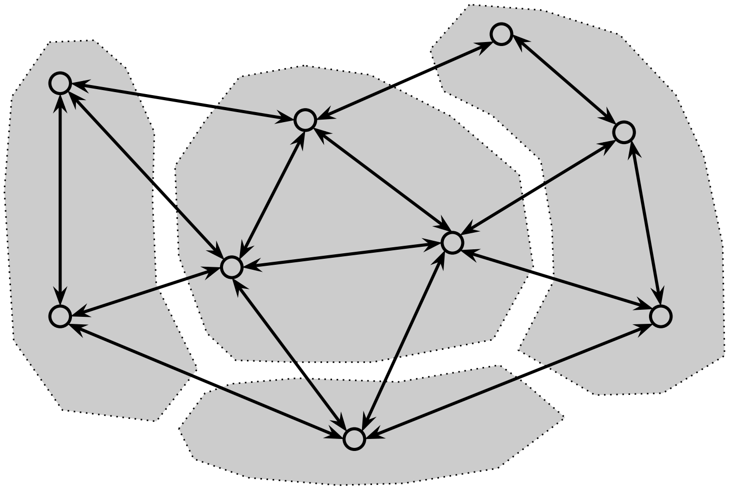


Figure 2.6: Transportation arcs and fuel supply regions

2.4.2 Energy commodities

The second interaction between systems corresponds to the need of transporting energy commodities. These are those commodities that are part of the energy system.

Fig. 2.7 captures the definition of these commodities. Energy networks are represented on the top and transportation on the bottom. In the coal network, coal is produced at the mines, transported and then converted to electricity. It is the movement from one coal node to another that is done using one or more modes of transportation.

However, transportation capacities and costs are taken into account in the transportation network. For that reason, the movement of coal in both network is synchronized. This is represented with the dashed line in the graph. This connection can be a one-to-one relation if tons is the unit used in the coal network. Otherwise, the heat content value of the type of coal being transported can be used so there is not a conflict with the units in the different networks.

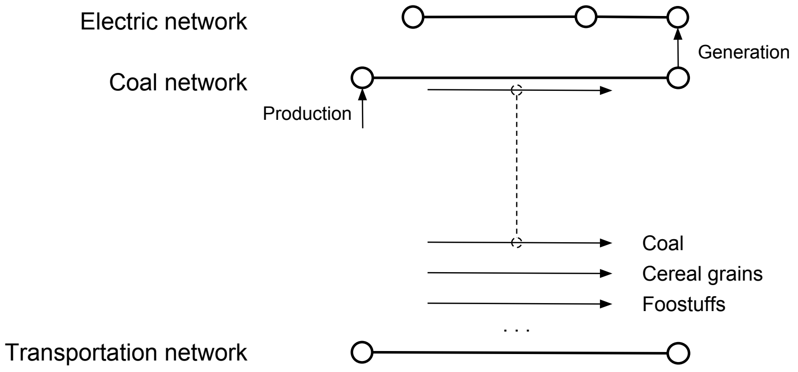


Figure 2.7: Transportation arcs and nodal energy demand

2.4.3 Interactions between investments

The previous interactions happen at the operational level. However, investment interactions interactions could exist between the energy and transportation systems like it was mentioned in the introduction of this dissertation and in Fig. 1.1. These interactions have not been introduced explicitly in NETPLAN but they should be taken into account in the preparation of the data or in future expansions of the framework.

One example of these would derive from the study of right-of-way needs for major infrastructure upgrades, such as interstate high-voltage transmission lines, railroads or highways. It could be the case that these infrastructures compete for said right-of-way or that the joint investment could be beneficial because of the proximity of one to another. Other constraints could take into account limited availability of other resources such as land or construction materials (steel, cement, silicon).

2.5 Summary of systems

Now that the different systems have been introduced, the main characteristics for each one of them is summarized. Table 2.1 contains the list of networks and for each one of them the following is specified: network flow type, commodities transported, units used, infrastructure, fleets, and where the node is specified.

Table 2.1: Summary of modeled systems

Network	Flow	Commodities	Units	Infrast.	Fleet	Demand
Energy	Single comm.	Electric	MWh	Electric	N/A	Nodes
		Natural gas		Pipeline		
		H_2, NH_3		Pipeline		
		Petroleum		Pipeline		
Energy comm. (coal)	Multicomm.	Bituminous	Tons	Rail	Diesel, elect.	Nodes
		Subbitmns		Barge	Diesel	
		Lignite		Highway	Diesel, hybrid	
Freight	Multicomm.	Grains	Tons	Rail	Diesel, elect.	Arcs
		Chemicals		Barge	Diesel	
		Gravel, etc.		Highway	Diesel, hybrid	
Passenger	Multicomm.	50 mph	People	Highway	Gasoline, elect.	Arcs
		100 mph		Rail	Diesel, elect.	
		200 mph		Rail	Electric	
		500 mph		Air	Petroleum	

Energy networks include electricity, natural gas, hydrogen, ammonia, and petroleum among others and all are expressed in energy units. Capacity infrastructure is enforced on the arcs and nodal demand is specified. On the other hand, freight transportation uses a multicommodity formulation with demand determined by arcs. Infrastructure and fleet are key in determining the maximum flows across those arcs. Similarly, the passenger network can be formulated using the multicommodity formulation taking into account infrastructure and fleets. The demand can be partitioned according to the speed required and that way distinguish between short and long trips, which are covered by different modes of transportation.

Energy commodities share some characteristics with both the energy and freight networks. In essence, energy commodities behave like freight transportation (multicommodity flow, use of infrastructure and fleet). However, the demand is defined by nodes and often is connected to other energy networks.

2.6 Metrics

This section will introduce the different metrics proposed to evaluate investment portfolios. The goal is to deliver metrics that integrate both the transportation and the energy networks.

2.6.1 Cost

There are two equally important components to the cost objective that can be differentiated: operational and investment costs. The latter usually receives the most attention perceived given the amount of capital that it requires over a relative short period of time. However, the operational component could have a larger impact on the total cost and the balance between of them must be considered.

Investment cost is calculated in dollars per unit of capacity, where capacity if given in MW for the energy side and tons/hour for the transportation side. The concept of overnight cost is used, which is the cost of financing the entire construction of a given infrastructure if only one payment was made at one point it time. This cost would include materials, labor, financial costs, intellectual property, and dismantling costs. Salvage value is taken into account with the appropriate inflation correction, should it exist.

Operational cost, as opposed to investment cost, is expressed in dollars per unit of energy produced or dollars per unit of mass transported. It could include, but is not limited to, some of the following: labor (operators, drivers), maintenance, byproduct disposal (nuclear waste, ash), or non-fuel materials (e.g. limestone in fossil fuel plants). Useful byproducts could potentially reduce the overall operational cost. There are two traditional operational cost components that are not assigned directly: fuel costs and amortization of the investment. The first is taken care of in the representation of the fuel sources and distribution networks with the appropriate interconnections, and the latter is included in the investment cost.

2.6.2 Sustainability

Sustainability is determined in terms of environmental impact and supply longevity. Four classes of environmental impacts are considered in relation to energy and transportation systems: net emissions, nuclear waste, water consumption (e.g., for biofuel production), and resource displacement (e.g., land usage). The most relevant emissions that result from energy and transportation systems are the emissions of four air pollutants (CO, NO_X, SO_2 and volatile organic compounds) [38], and green-house gas emissions (CO_2 and methane) [39].

For each environmental impact belonging to any of our four classes (emissions, waste, consumption, displacement) a linear expression is considered. Coefficients representing the impact per unit of flow need to be determined prior to the optimization process. Environmental consequences of investment can also be included and computed by analyzing the nominal impact of the life cycle of each infrastructure, such as emission during the processing of raw materials (steel, concrete) or during construction.

We also characterize supply longevity for a depletable resource (e.g., coal, gas, uranium) as the remaining years for the resource if used at the average rate over the simulation time. Sustainability expressions for water and land as depletable resources or air pollutants that should not exceed a predetermined threshold can be modeled as complicating constraints [35] that specify flow relationships between several arcs.

Greenhouse emissions are an essential part in the planning processes for energy and transportation, so an index [40] that compares solutions to projected trends has been developed. Other studies [41, 42, 43] rely on the evaluation of social costs associated with these emissions as well as trying to capture the time value of carbon. This index identifies not only the global reduction but also encourages trends that reduce emissions over time. Early cuts in emissions are preferred and accounted for, introducing the concept of “time value” of emissions reduction.

To calculate this index it is necessary to define a pessimistic scenario in which nothing is done to palliate greenhouse gas emissions and an optimistic case in which an aggressive reduction is accomplished. Other emission forecasts can be compared to those references with the following metric (2.1):

$$I_{GHG} = \frac{1}{n-k} \sum_{t=k}^{n} \frac{ActualEm(t) - LowEm(t)}{HighEm(t) - LowEm(t)} \tag{2.1}$$

where k and n are the first and last year considered for the index, $ActualEm$ is the emission forecast, and $HighEm$ and $LowEm$ are the yearly emissions for the pessimistic and optimistic cases, respectively. Because of lead times for new infrastructures, the first k years are not considered in the calculation of the index.

Figure 2.8 provides a representation of this approach. The top solid line represents the pessimistic scenario, with emissions increasing at the same rate as electric demand. The bottom line represents the optimistic case, with reductions in greenhouse emissions every year comparable to proposed 2010 bills in the U.S. House and Senate [44]. In between those lines we can see how the index takes different fractional values. The index is not limited to values between 0 and 1 because emissions could go beyond the reference cases.

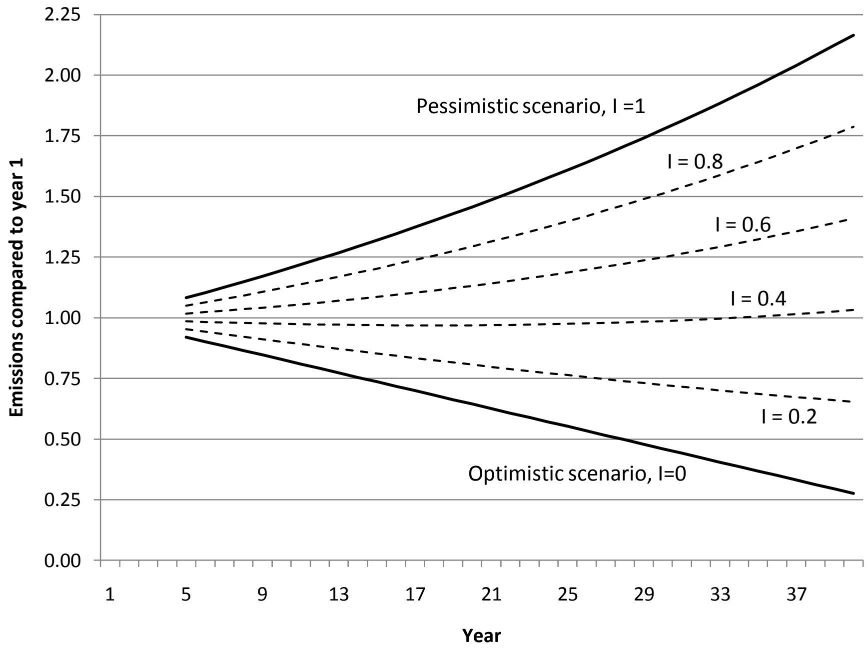


Figure 2.8: Greenhouse emission index for best, worst scenarios and fractions in between

2.6.3 Resiliency

In this section we provide a definition of resiliency which is appropriate for long-term investment planning [45]. This definition depends on three basic concepts: states, events and consequences, which are described first.

States The system state is loosely defined as the group of system attributes that completely characterize it at a particular time kT $(k = 1, 2, \ldots)$, where T is a duration of time for which the system is considered to be in a steady-state. Define topology as the physical infrastructure operable at that instant in time including, for example, electric generators and circuits, natural gas wells and pipelines, railways and trains, and highways, trucks, and cars. Further define operating conditions as the characterization of the way the physical infrastructure is used at that instant in time in terms of, for example, electric system loading and generation dispatch, natural gas supply and transport, and movement of commodities and passengers. Then we define state as a tuple consisting of specification of the topology and operating condition of the system.

Events System resiliency can only be assessed if one or more system changes occur or are simulated expecting to observe some kind of system response. Consequently, resiliency can only be considered relative to a defined system change or set of changes. Such changes may occur to the topology, to the operating conditions, or to both. For purposes of modeling and simulation, we conceive of such changes in terms of a first cause or source, which lies outside of our modeling framework, and an impact, which characterizes the changes in terms of the effect the source has on what is modeled. We refer to the combination of a particular source and its impact as an event. As an example, consider modeling the effect of a large scale natural disaster such as the effect of the 2005 Katrina/Rita hurricanes on the national energy system [46]. In this case, the hurricanes are the source and the loss of natural gas wells and pipelines together with loss of electric generation, transmission, and load in the gulf coast area is the impact.

An analyst or designer, in using simulation to consider resiliency, must therefore choose a set of events to simulate. This choice should be made based on the objective of the study. For example, in an operational study of a system, one wants to ensure the system will perform acceptably following the next event; in this case, the most likely events are of interest. In transportation systems, such events include accidents and weather-related closures for portions of the system. For electric transmission systems, NERC refers to such events as category B events [47]. NERC also specifies what is referred to as annual "extreme event testing" using so-called Category D events, very rare, but generally very severe, to "evaluate for risks and consequences." This is a type of resiliency evaluation, and it is in this spirit that we propose to evaluate resiliency of the national energy and transportation system using a limited set of very rare but very severe events. The Katrina and Rita hurricanes represent a good example of such an event, where a large amount of natural gas production was constrained for several weeks simultaneous with significant reduction of Mississippi River barge traffic and loss of many electric generation and transmission facilities in the area. Other events that could be simulated to assess resiliency of the energy and transportation systems include:

- 6 month loss of rail access to Powder River Basin coal,
- 1 year interruption of 90% of Middle Eastern oil,
- Permanent loss of US nuclear supply [48],
- 6 month interruption of Canadian gas supply,
- Earthquake in St. Louis [49] with major loss of transmission, rail, oil, and gas pipelines and extended interruption to Mississippi River barge traffic,
- 1 year loss of US hydro resources due to extreme drought,
- 1 year loss of US wind resources due to climate change effects,
- Sustained flooding in the Midwest that destroys crops, reducing the availability of biofuels, and interrupts key corridors of the east-west railroad system.

A common characteristic of these events is that their effects are typically exacerbated by some congestion. Recent effort has identified congested regions on the nation's electric transmission system, as illustrated in Fig. 2.9 [50]. Figures 2.10 [51] and 2.11 [52] compare current and forecasted future congestion of highway and railway, respectively, with orange-colored links in each map indicating highly congested paths. Events which include failure of such paths would constitute good candidate events for resiliency testing [53].

Ultimately the space of events to test is large and the selection can become a challenge. Finding the set of events that represent the highest risk can be achieved with various techniques, such as sensitivity analysis on the operational solution, Monte Carlo simulations, or specialized explicit formulations, such as the interdiction problem [54]. This set depends on the structure of the energy and transportation networks, which depends on the profile of the investment decisions.

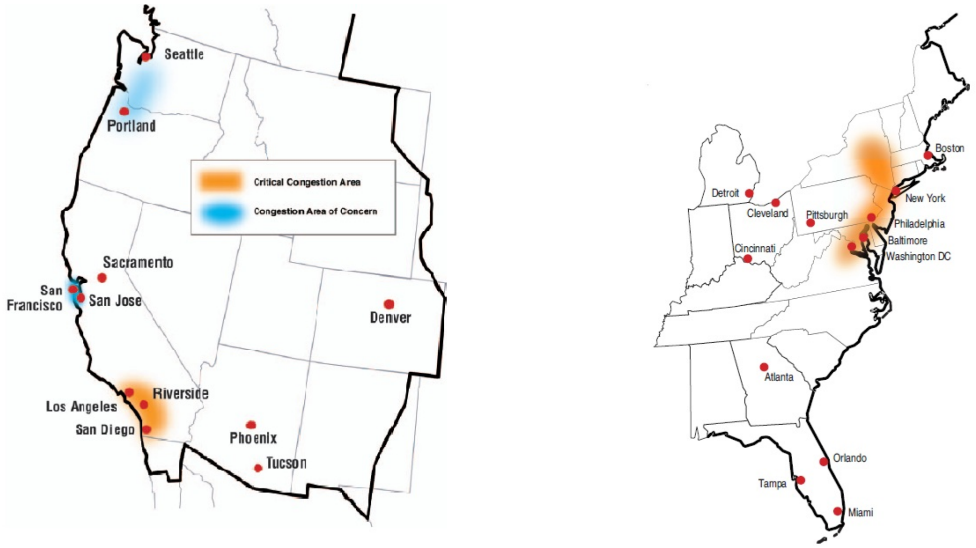


Figure 2.9: Congested areas for the Western and Eastern interconnections

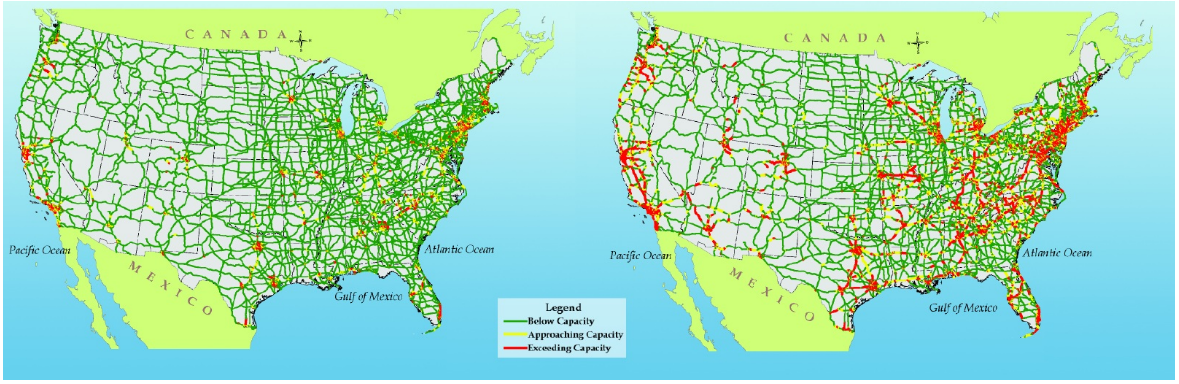


Figure 2.10: Highway level of congestion during peak in 1998 and 2020 (predicted)

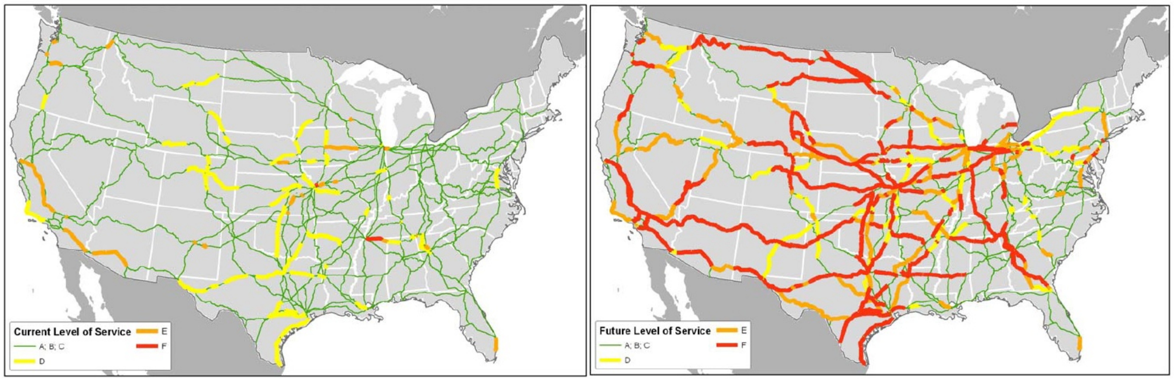


Figure 2.11: Rail levels of congestions in 2005 and 2035 (predicted)

Consequence The kind of events described above for which we are interested in simulating to evaluate resiliency will cause significant performance deviation which can be observed in NETPLAN. This is referred to as consequence and it can be measured in terms of increased operational cost with respect to the base case. Alternative and promising methods include the variation in nodal price per unit of energy or nodal price per unit of ton-miles integrated over time.

Variation in consequence may in turn have observable influence on many other aspects of society, including prices of manufactured goods, job loss, and gross national product. We define these other influences which we are not able to observe in our model as societal influences. Figure 2.12 illustrates these definitions together with those associated with events.

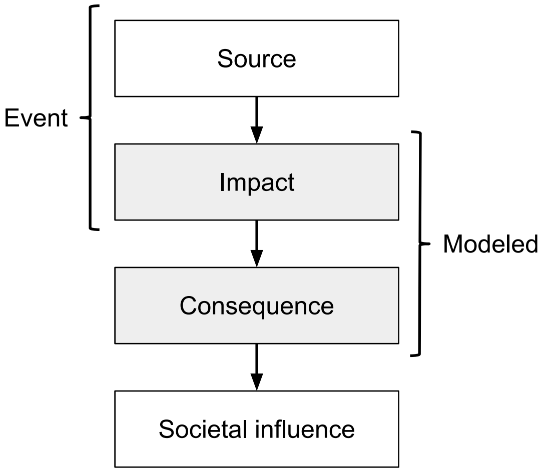


Figure 2.12: Relationship between events and consequences

After explaining the concepts of state, events and consequence, it is appropriate to formally define the concepts of resiliency and robustness in NETPLAN.

Resiliency and robustness Resiliency is the ability to minimize and recover from the consequences of an adverse event, whether natural or man-made, for a given state of the system. It is measure consequence in terms of system performance, as shown in Fig. 2.13, where we observe that maximum performance deviation, recovery time (duration following event initiation for performance measure to reach a steady-state value) and steady-state deviation, $\Delta P_{\infty,ij}$, play important roles in the evaluation of consequence.

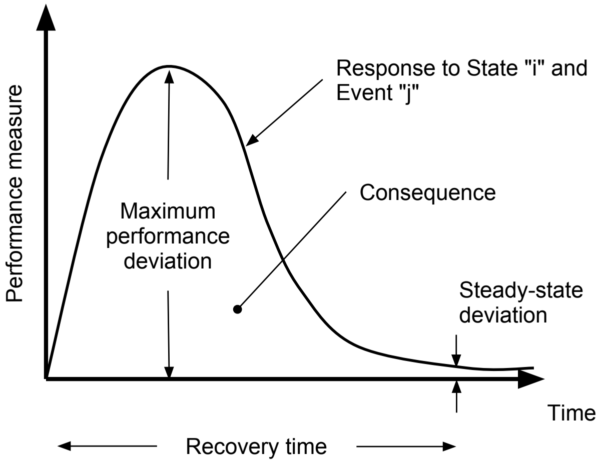


Figure 2.13: Illustration of resiliency measure for an event and state

Consider a set of events $(E_1, \ldots, E_I)$ and a set of states $(S_1, \ldots, S_J)$ that we deem relevant. Denote the performance measure as P and the consequence as C. Then, for event E_i occurring in state S_j, the consequence is given by:

$$C_{ij} = \int_0^\infty P(t)dt \tag{2.2}$$

as illustrated in Fig. 2.13. A system becomes more resilient with respect to an event E_i occurring in state S_j as C_{ij} decreases.

If $P_{max,ij}$ is the maximum performance deviation, $T_{R,ij}$ the recovery time, and $\Delta P_{\infty,ij}$ the steady-state deviation relative to performance previous to the event, then a simpler resiliency metric which serves as a reasonable approximation to (2.2) is

$$C_{ij} \approx (0.5 \times P_{max,ij} + \Delta P_{\infty,ij})T_{R,ij} \tag{2.3}$$

A system is robust with respect to an event if it is resilient for that event under all defined states. The calculation of consequence under different states, as illustrated in Fig. 2.14, yields a distribution from which one may extract an appropriate summary of the data (e.g. mean, median, spread, standard deviation). The consequence spread is interpreted as a measure of robustness.

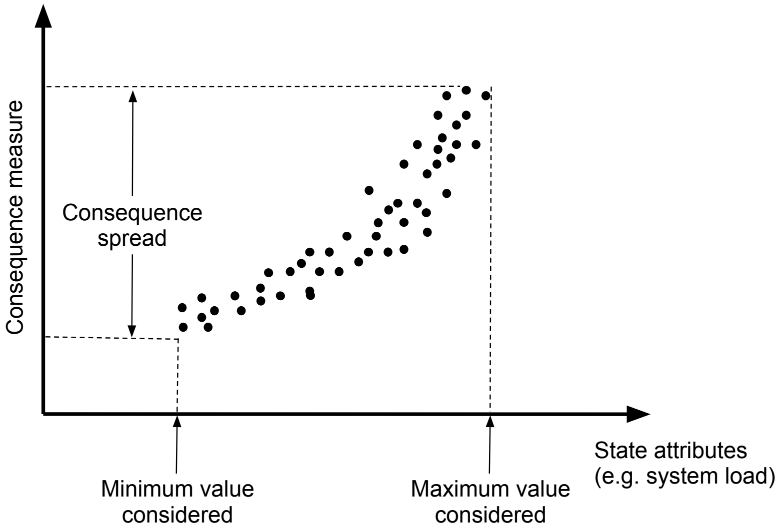


Figure 2.14: Illustration of resiliency across states

2.7 Small example

Even though the formal description of the mathematical model will be fully introduced in the next chapter, this section introduces a simple example in order to illustrate some of the capabilities of the proposed modeling principles [55]. The illustrated example (Fig. 2.15) features a high level representation of the energy and transport relations between the Midwestern and Eastern sections of the United States. These areas are also respectively referred to as "1" and "2". We consider the Midwest region to be delimited by the states between North Dakota, Wisconsin, Mississippi and Texas.

The Midwestern area is assumed to produce two types of commodities, coal and corn, which need to be transported to meet the East Coast demand. To simplify the model, it is assumed that coal is produced in the Illinois Basin with enough capacity to meet the thermal generation demand in the East Coast and the Midwest. To transport these commodities, two different infrastructures, railway and highway, are utilized. Only one type of fleet is accounted for in each infrastructure, train and truck, respectively.

Two different energy networks are considered. The diesel network is fed by the production in the Midwest and its mission is to fulfill the need for fuel from trains and trucks. Electricity

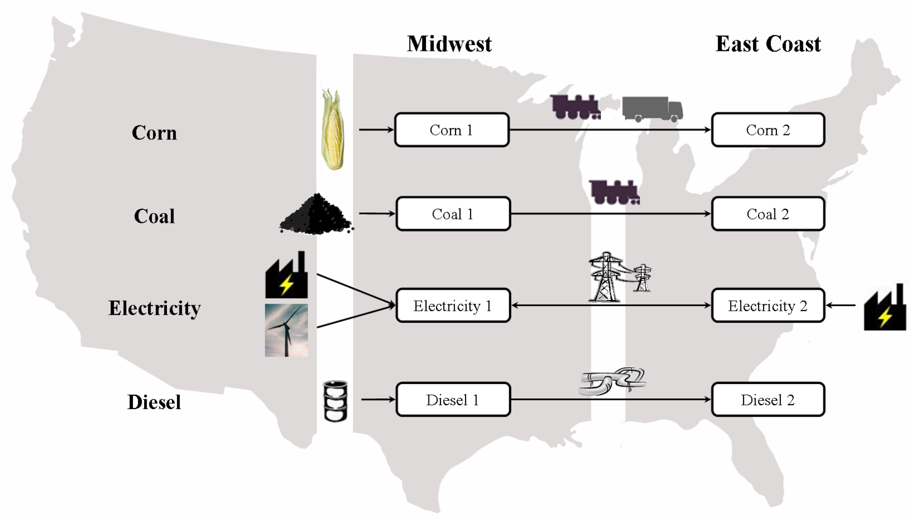


Figure 2.15: Small example layout displaying the two geographical regions

supply constitutes the second energy network. Both areas can produce electricity from thermal plants, which drive the demand for coal on the transportation side, and are connected by high-voltage power lines to allow energy trading. The Midwestern area has potential to use wind as a source for electricity, although there is no capacity installed at the beginning of the planning period, which lasts 40 years.

In order to capture all the components of the formulation, the previous set of physical nodes and arcs can be expanded as shown in Fig. 2.16. Note that columns represent the four networks and every row represents a node, corresponding to a physical region.

In the transportation networks the breakdown among different types of modes is represented between the Midwest and the East Coast. The transmission line in the electric network is replaced by two opposite directional arcs to ensure the non-negativity of the flows. The dashed lines represent the increase in demand on a node due to activity on other network, i.e., thermal units increase the demand for coal and the use of train and truck for transportation drive the demand for diesel. Based on this figure, we define the following as the set of variables and their corresponding units to represent the nodes around the two-node system for each year t.

- $e_{(CP1,CT1)}(t)$: Coal production in region 1 (MMBtu)
- $e_{(CT1,CT2)}(t)$: Coal transportation (as energy) between regions 1 and 2 (MMBtu)

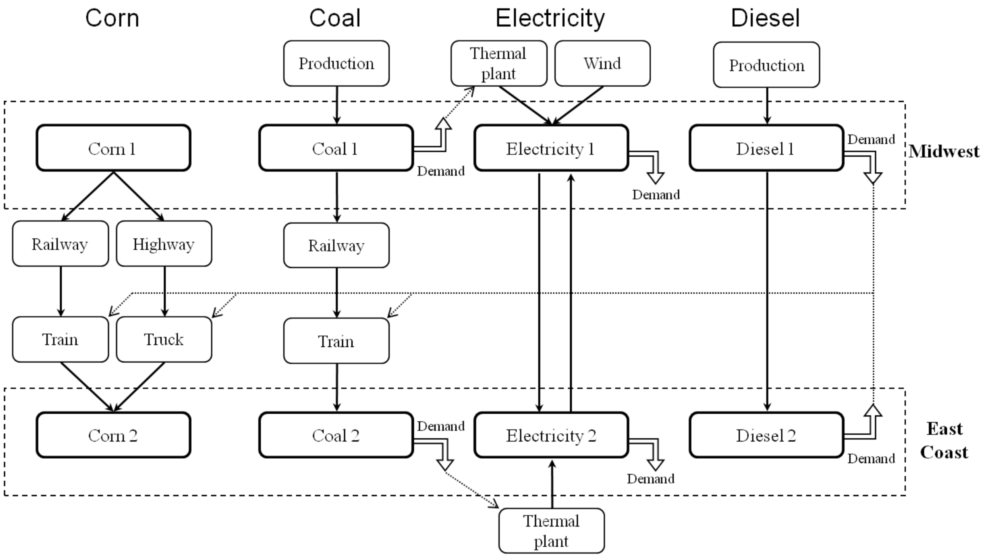


Figure 2.16: Example node and arc expansion for the two-node example

- $e_{(CT1,EC1)}(t)$, $e_{(CT2,EC2)}(t)$: Connection between node transportation node and coal generation for regions 1 and 2, respectively (MMBtu)
- $e_{(EC1,EL1)}(t)$, $e_{(EC2,EL2)}(t)$: Coal generation in regions 1 and 2 (GWh)
- $e_{(EW1,EL1)}(t)$: Wind generation in region 1 (GWh)
- $e_{(EL1,EL2)}(t)$, $e_{(EL2,EL1)}(t)$: Electric transmission between regions 1 and 2 (GWh)
- $e_{(DP1,DT1)}(t)$: Diesel production in region 1 (MMBtu)
- $e_{(DT1,DT2)}(t)$: Diesel pipeline between regions 1 and 2 (MMBtu)
- $f_{(1,2,C,r)}(t)$: Flow of coal (as freight commodity) between regions 1 and 2 by rail (ton)
- $f_{(1,2,O,k)}(t)$: Flow of corn between regions 1 and 2 by truck (ton)
- $f_{(1,2,O,r)}(t)$: Flow of corn between regions 1 and 2 by train (ton)

To simplify the analysis it is assumed that there is only capacity limits and investment in the following parts of the system: train and truck transportation, thermal plants, wind generation and electricity transmission. No investment on transportation infrastructure is considered. Also, there is no retirement of facilities or infrastructure so the initial capacity is assumed to be available throughout the simulation. We replace e in the variables above with $eCap$ and $eInv$ to respectively represent capacity and investment decision variables. Capacities and investments are expressed in power units. For freight transportation, fleet capacities and investments are defined as:

- $fleetCap_{(1,2,r)}(t)$, $fleetInv_{(1,2,r)}(t)$: Rail capacity and investment for year t between regions 1 and 2 by rail (ton/year)
- $fleetCap_{(1,2,k)}(t)$, $fleetCap_{(1,2,k)}(t)$: Truck capacity and investment between regions 1 and 2 by truck (ton/year)

The simulation is performed for 40 years with a time step of a year. Operational and investment parameters for energy networks are summarized in Table 2.2, while Table 2.3 contains the appropriate parameters of capacity, frequency of travel, costs and emissions for the available fleet.

Table 2.2: Electricity network parameters for small example

	Coal fired	Wind	Transmission
Initial capacity	225 / 180 GW	0 GW	50 GW
Max. investment	10 GW/year	3.5 GW/year	5 MW/year
Op. cost	1.7 \$/MWh	7 \$/MWh	2 \$/MWh
Investment cost	2.12 \$/MW	1.65 \$/MW	825 \$/kW-mile
Efficiency	9.95 MMBtu/MWh	0.35	1
CO_2 emission rate	55.77 lb/MWh	0	—

The electric demand is set to 141 GW for the Midwest region and 118 GW for the East Coast, with a growth of 1.5% every year. The amount of corn shipped from the Midwest to the East Coast equals 300 million tons, with a 0.5% yearly growth. The average distance between the two regions is set to 750 miles. Coal is produced in the Illinois Basin with a content of energy equal to 11,800 Btu/short ton and a cost of \$85 per short ton. Diesel production cost is set to \$31 per barrel and its endowed with a heat content of 143,500 Btu per gallon.

All costs are subject to a constant of 2% annual inflation, and a constant discount rate of 7% is used in the economic analysis. Salvage values are assigned for the investments close

Table 2.3: Transportation parameters for small example

	Train	Truck
Capacity	3200 ton/load	26 ton/load
Loads/year	68 loads	104 loads
Initial capacity	3500 trains	100,000 trucks
Operational cost	0.05 \$/ton-mile	0.16 \$/ton-mile
Fuel use	341 Btu/ton-mile	3357 Btu/ton-mile
CO_2 emission rate	0.2 lb/ton-mile	0.6 lb/ton-mile
Max. investment	50 trains	100 trucks
Investment cost	31 million \$/train	200,000 \$/truck

to the end of the simulation period. All investments suffer linear devaluation for a period of 15 years.

The following is a list of the different equations that form part of the model. Unless otherwise noted, one example is given for the first year in the simulation. We begin by stating the constraints that enforce that demand at the coal nodes:

$$e_{(CP1,CT1)}(1) = e_{(\mathrm{CT1,EC1})}(1) + e_{(CT1,CT2)}(1) \tag{2.4}$$

$$e_{(CT1,CT2)}(1) = e_{(CT1,EC1)}(1) \tag{2.5}$$

Next, electric demand is met at the electric nodes in regions 1 and 2:

$$e_{(EC1,EL1)}(1) + e_{(EW1,EL1)}(1) = (141\ \mathrm{GW})(8760\ \mathrm{h}) \tag{2.6}$$

$$e_{(EC2,EL2)}(1) = (118\ \mathrm{GW})(8760\ \mathrm{h}) \tag{2.7}$$

The following equations take into account the conversion from coal to electricity at thermal plants in regions 1 and 2:

$$e_{(EC1,EL1)}(1) = e_{(CT1,EC1)}(1)\ \frac{1\ \mathrm{GWh}}{9.95 \cdot 10^{-3}\ \mathrm{MMBtu}} \tag{2.8}$$

$$e_{(EC2,EL2)}(1) = e_{(CT2,EC2)}(1)\ \frac{1\ \mathrm{GWh}}{9.95 \cdot 10^{-3}\ \mathrm{MMBtu}} \tag{2.9}$$

Finally, demand of diesel is enforced at the corresponding nodes. The demand comes only from the use of truck and train transportation.

$$e_{(DP1,DT1)}(1) = e_{(DT1,DT2)}(1) + \frac{1}{2}(3.357 \cdot 10^{-3} \cdot 750 \text{ MMBtu/ton}) f_{(1,2,C,k)}(1) \tag{2.10}$$
$$+ \frac{1}{2}(0.341 \cdot 10^{-3} \cdot 750 \text{ MMBtu/ton})[f_{(1,2,C,r)}(1) + f_{(1,2,O,r)}(1)]$$
$$e_{(DT1,DT2)}(1) = \frac{1}{2}(3.357 \cdot 10^{-3} \cdot 750 \text{ MMBtu/ton}) f_{(1,2,C,k)}(1) \tag{2.11}$$
$$+ \frac{1}{2}(0.341 \cdot 10^{-3} \cdot 750 \text{ MMBtu/ton})[f_{(1,2,C,r)}(1) + f_{(1,2,O,r)}(1)]$$

The transportation of corn and coal has to be satisfied and the appropriate modes of transportation will be chosen. Notice how coal transport as freight is related to its equivalent in the energy system.

$$f_{(1,2,O,k)}(1) + f_{(1,2,O,r)}(1) = 300 \cdot 10^{6} \text{ ton} \tag{2.12}$$
$$f_{(1,2,C,r)}(t) = \frac{\text{ton}}{0.0118 \text{ MMBtu}} e_{(CT1,CT2)}(1) \tag{2.13}$$

Once the demands are established, capacities for electric generation and transmissions are enforced:

$$e_{(EC1,EL1)}(1) \leq eCap_{(EC1,EL1)}(1) \ (8760 \text{ h}) \tag{2.14}$$
$$e_{(EC2,EL2)}(1) \leq eCap_{(EC2,EL2)}(1) \ (8760 \text{ h}) \tag{2.15}$$
$$e_{(EW1,EL1)}(1) \leq 0.35 \ eCap_{(EW1,EL1)}(1) \ (8760 \text{ h}) \tag{2.16}$$
$$e_{(EL1,EL2)}(1) \leq eCap_{(EL1,EL2)}(1) \ (8760 \text{ h}) \tag{2.17}$$

These capacities depend on the initial capacity and investments and here is an example for year 10:

$$eCap_{(EC1,EL1)}(10) = (225 \text{ GW}) + eInv_{(EC1,EL1)}(2) + \ldots + eInv_{(EC1,EL1)}(10) \tag{2.18}$$
$$eCap_{(EC2,EL2)}(10) = (180 \text{ GW}) + eInv_{(EC2,EL2)}(2) + \ldots + eInv_{(EC2,EL2)}(10) \tag{2.19}$$
$$eCap_{(EW1,EL1)}(10) = (0 \text{ GW}) + eInv_{(EW1,EL1)}(2) + \ldots + eInv_{(EW1,EL1)}(10) \tag{2.20}$$
$$eCap_{(EL1,EL2)}(10) = (50 \text{ GW}) + eInv_{(EL1,EL2)}(2) + \ldots + eInv_{(EL1,EL2)}(10) \tag{2.21}$$

Investments must be below the maximum yearly allowed value:

$$eInv_{(\mathit{EC1},\mathit{EL1})}(t) \leq 10 \text{ GW} \tag{2.22}$$

$$eInv_{(\mathit{EC2},\mathit{EL2})}(t) \leq 10 \text{ GW} \tag{2.23}$$

$$eInv_{(\mathit{EW1},\mathit{EL1})}(t) \leq 3.5 \text{ GW} \tag{2.24}$$

$$eInv_{(\mathit{EL1},\mathit{EL2})}(t) \leq 5 \cdot 10^{-3} \text{ GW} \tag{2.25}$$

Fleet capacities are also enforced for rail and truck:

$$f_{(1,2,C,1)}(1) + f_{(1,2,O,r)}(1) \leq fleetCap_{(1,2,r)}(t) \tag{2.26}$$

$$f_{(1,2,O,k)}(1) \leq fleetCap_{(1,2,k)}(t) \tag{2.27}$$

The definition of capacity also depends on initial capacity and investments. For year 10, these would be defined for rail and truck as:

$$fleetCap_{(1,2,r)}(10) = (3500 \cdot 3200 \cdot 68 \text{ ton/year}) + \sum_{t=2}^{10} fleetInv_{(1,2,r)}(t) \tag{2.28}$$

$$fleetCap_{(1,2,k)}(10) = (100,000 \cdot 26 \cdot 104 \text{ ton/year}) + \sum_{t=2}^{10} fleetInv_{(1,2,l)}(t) \tag{2.29}$$

The maximum yearly investments for fleet are also determined:

$$fleetInv_{(1,2,r)}(t) \leq 50 \cdot 3200 \cdot 68 \text{ ton/year} \tag{2.30}$$

$$fleetInv_{(1,2,l)}(t) \leq 100 \cdot 26 \cdot 104 \text{ ton/year} \tag{2.31}$$

The cost objective is calculated by considering operational (2.32) and investment (2.33) costs for energy and operational (2.34) and fleet investment (2.35) costs for transportation.

$$\begin{aligned} CostOp^{E} = {} & (85 \cdot 0.0118 \text{ \$/MMBtu})\ e_{(\mathit{CP1},\mathit{CT1})}(1) + \ldots + \\ & + (1700 \text{ \$/GWh})\ e_{(\mathit{EC1},\mathit{EL1})}(1) + \ldots + \\ & + (1700 \text{ \$/GWh})\ e_{(\mathit{EC2},\mathit{EL2})}(1) + \ldots + \\ & + (7000 \text{ \$/GWh})\ e_{(\mathit{EW1},\mathit{EL1})}(1) + \ldots + \\ & + (2000 \text{ \$/GWh})\ e_{(\mathit{EL1},\mathit{EL2})}(1) + \ldots \end{aligned} \tag{2.32}$$

$$\begin{aligned} CostInv^{E} = {} & \frac{1.02}{1.07}\,(2120\ \$/\mathrm{GW})\ eInv_{(EC1,EL1)}(2) + \ldots + \\ & + \frac{1.02}{1.07}\,(2120\ \$/\mathrm{GW})\ eInv_{(EC2,EL2)}(2) + \ldots + \\ & + \frac{1.02}{1.07}\,(1650\ \$/\mathrm{GW})\ e_{(EW1,EL1)}(2) + \ldots + \\ & + \frac{1.02}{1.07}\,(825 \cdot 10^{6} \cdot 750\ \$/\mathrm{GW})\ e_{(EL1,EL2)}(2) + \ldots + \\ & + \frac{1.02}{1.07}\,\frac{\$31}{42 \cdot 0.1435\ \mathrm{MMBtu}}\ e_{(DP1,DT1)}(2) + \ldots \end{aligned} \tag{2.33}$$

$$\begin{aligned} CostOp^{T} = {} & (0.05 \cdot 750\ \$/\mathrm{ton})\ f_{(1,2,C,r)}(1) + \ldots + \\ & + (0.05 \cdot 750\ \$/\mathrm{ton})\ f_{(1,2,O,r)}(1) + \ldots + \\ & + (0.16 \cdot 750\ \$/\mathrm{ton})\ f_{(1,2,O,k)}(1) + \ldots \end{aligned} \tag{2.34}$$

$$\begin{aligned} CostInv^{T} = {} & \frac{1.02}{1.07}\,\frac{31 \cdot 10^{6}\ \$}{68 \cdot 3200\ \mathrm{ton/year}}\ f_{(1,2,r)}(2) + \ldots + \\ & + \frac{1.02}{1.07}\,\frac{200,000\ \$}{26 \cdot 104\ \mathrm{ton/year}}\ f_{(1,2,k)}(2) + \ldots \end{aligned} \tag{2.35}$$

CHAPTER 3. MATHEMATICAL FORMULATION

This chapter formally describes the models that implement the functions for NETPLAN. It begins with a description of the cost minimization problem, the introduction of the model parameters and the multiobjective optimization algorithms. Computational enhancements by decomposition and parallelization are presented, as well as the structure of the software that is able to solve the proposed problems.

3.1 General cost minimization formulation

The optimization problem associated with the investment and operation model can be conceptually described by (3.1),

$$\begin{aligned}
&\textbf{min} \quad CostOp + CostInv \\
&\textbf{subject to:} \\
&\textit{Meet energy demand,} \\
&\textit{Meet transportation demand,} \\
&\textit{Capacity constraints} \\
&\textit{Power flow constraints on electric transmission}
\end{aligned} \tag{3.1}$$

The objective is to minimize the combined energy and a transportation cost with constraints of meeting demands on energy and freight transport while following the capacity constraints. The operational characteristics of the model provide additional constraints on the system, such as the inclusion of power flow relations for electric transmission lines (we used the so-called "DC" power flow approximation for this purpose).

This chapter contains a rigorous description of the formulation needed to achieve the characteristics described above. The explanation of the formulation is preceded by an introduction to the nomenclature used.

3.2 Nomenclature

3.2.1 Decision Variables

$e_{(i,j)}(\mathbf{t})$: Operational flow of energy arc from node i to node j, for time step $\mathbf{t}$ (MWh)

$eInv_{(i,j)}(\mathbf{t})$: Capacity investment on energy arc from node i to node j, for time step $\mathbf{t}$. (MW)

$f_{(i,j,k,m)}(\mathbf{t})$: Operational flow of transportation arc from node i to node j for commodity k using transportation mode m during time step $\mathbf{t}$ (ton)

$fleetInv_{(i,j,m)}(\mathbf{t})$: Fleet m capacity investment for transportation arc from node i to node j during time step $\mathbf{t}$ (ton/hour)

$infInv_{(i,j,l)}(\mathbf{t})$: Infrastructure l capacity investment for transportation arc from node i to node j for time step $\mathbf{t}$ (ton/hour)

$\theta_i(\mathbf{t})$: Phase angle at node i, used to model DC power flow (radians)

3.2.2 Sets and networks

$\mathcal{N}^E$: Set of energy nodes

$\mathcal{N}^E_d \subset \mathcal{N}^E$: Subset of energy nodes where demand equations are enforced

$\mathcal{N}^E_p \subset \mathcal{N}^E$: Subset of energy nodes where peak demand equations are enforced

$\mathcal{A}^E$: Set of energy arcs

$\mathcal{A}^E_{DC} \subset \mathcal{A}^E$: Set of AC electric transmission arcs, which satisfy DC power flow equations

$\mathcal{N}^T$: Set of transportation nodes

$\mathcal{A}^T$: Set of transportation arcs

$\mathcal{A}_j^T \subset \mathcal{A}^T$: Subset of transportation arcs that create an energy demand at energy node j

$\mathcal{K}$: Set of commodities

$\mathcal{K}_e \subset \mathcal{K}$: Subset of commodities used by the energy system, e.g. coal

$\mathcal{M}$: Set of fleet or transportation modes

$\mathcal{M}_l \subset \mathcal{M}$: Subset of fleet that can use transportation infrastructure l

$\mathcal{M}_j \subset \mathcal{M}$: Subset of fleet that require energy from energy node j

$n^E_{(i,k)} \in \mathcal{N}^E$: Energy node that corresponds to the geographic location i and commodity k in the transportation system

3.2.3 Time and time steps

$\mathcal{T} = [1, \ldots, T_1] \times [0, \ldots, T_2] \times \ldots \times [0, \ldots, T_s]$: Definition of simulation time domain, which is divided in s time steps. Time step i is divided into T_i segments. For instance, simulation time could be divided in “years”, “months” and “days” with the appropriate number of divisions within each step

$\mathbf{t} = [t_1, t_2, \ldots, t_s] \in \mathcal{T}$: Time instance in the simulation domain

$\mathbf{t}_1$: Value of the top level step for time $\mathbf{t}$, i.e., t_1

$\mathcal{T}_\mathbf{t}$: Time that spans between the beginning of the simulation and time $\mathbf{t}$

$\Delta(\mathbf{t})$: Length of time step $\mathbf{t}$ (h)

3.2.4 Energy Parameters

$\eta_{(i,j)}(\mathbf{t})$: Efficiency of arc (i,j) during time $\mathbf{t}$ (unitless)

$lbe_{(i,j)}(\mathbf{t})$: Lower bound for flow in arc (i,j) for time $\mathbf{t}$ (MWh)

$ube_{(i,j)}(\mathbf{t})$: Upper bound for flow in arc (i,j) during time $\mathbf{t}$ due to the initial existing infrastructure (MW)

$lbeInv_{(i,j)}(\mathbf{t})$: Minimum allowed capacity increase in arc (i,j) at time $\mathbf{t}$ (MW)

$ubeInv_{(i,j)}(\mathbf{t})$: Maximum allowed capacity increase in arc (i,j) at time $\mathbf{t}$ (MW)

$eLife_{(i,j)}$: Expected lifetime for investments in arc (i,j) at time $\mathbf{t}$ (year)

$costOp_{(i,j)}(\mathbf{t})$: Operational cost for flow in arc (i,j) during time $\mathbf{t}$ (\$/MWh)

$costInv_{(i,j)}(\mathbf{t})$: Investment cost for capacity increase in arc (i,j) in network l, at time $\mathbf{t}$ (\$/MW)

$heatContent_k(\mathbf{t})$: Heat content of commodity k, at time $\mathbf{t}$ (MWh/ton)

$d_j^E(\mathbf{t})$: Fixed energy demand at node j during time $\mathbf{t}$ (MWh)

$peakD_j^E(\mathbf{t})$: Peak demand at electric node j during time $\mathbf{t}$ (MWh)

$cf_{(i,j)}(\mathbf{t})$: Capacity factor of generation between nodes i and j at time $\mathbf{t}$ (unitless)

$b_{(i,j)}$: Susceptance of transmission line between nodes i and j (p.u.)

3.2.5 Transportation Parameters

$ubFleet_{(i,j,m)}(\mathbf{t})$: Upper bound for total transportation flow for fleet m in arc (i,j) during time $\mathbf{t}$ due to the initial existing fleet (ton/h)

$lbFleetInv_{(i,j,m)}(\mathbf{t})$: Minimum allowed capacity increase in arc (i,j) for fleet m at time $\mathbf{t}$ (ton/h)

$ubFleetInv_{(i,j,m)}(\mathbf{t})$: Maximum allowed capacity increase in arc (i,j) for fleet m at time $\mathbf{t}$ (ton/h)

$fleetLife_{(i,j,m)}$: Expected lifetime of fleet m in arc (i,j) (years)

$ubInf_{(i,j,l)}(\mathbf{t})$: Upper bound for total transportation flow across infrastructure l in arc (i,j) during time $\mathbf{t}$ due to the initial existing infrastructure (ton/h)

$\mathit{lbInfInv}_{(i,j,l)}(\mathbf{t})$: Minimum allowed capacity increase in arc (i, j) for transportation infrastructure l at time $\mathbf{t}$ (ton/h)

$\mathit{ubInfInv}_{(i,j,l)}(\mathbf{t})$: Maximum allowed capacity increase in arc (i, j) for transportation infrastructure l at time $\mathbf{t}$ (ton/h)

$\mathit{infLife}_{(i,j,k,m)}$: Expected lifetime for infrastructure investment l in arc (i, j) (years)

$\mathit{costOp}_{(i,j,k,m)}(\mathbf{t})$: Operational cost for transportation flow in arc (i, j) in fleet m, commodity k and time $\mathbf{t}$ (\$/ton)

$\mathit{costFleetInv}_{(i,j,m)}(\mathbf{t})$: Investment cost for capacity increase in fleet m for arc (i, j) and time $\mathbf{t}$ (\$ h/ton)

$\mathit{costInfInv}_{(i,j,l)}(\mathbf{t})$: Investment cost for capacity increase in transportation infrastructure l for arc (i, j) and time $\mathbf{t}$ (\$ h/ton)

$\mathit{fuelCons}_{(i,j,m)}(\mathbf{t})$: Fuel consumption for transportation mode m for transportation arc (i, j) during time step $\mathbf{t}$ (MWh/ton)

$d^T_{(i,j,k)}(\mathbf{t})$: Fixed transportation demand of commodity k for arc (i, j) during time $\mathbf{t}$ (ton)

3.2.6 Auxiliary parameters

These parameters are calculated as a combination of decision variables and predetermined parameters.

CostOp^E: Operational cost from the energy system (\$)

$\mathit{CostInv}^E$: Cost due to the investment upgrades on the energy system (\$)

CostOp^T: Operational cost for transportation system (\$)

$\mathit{CostFleetInv}^T$: Investment cost on transportation fleet (\$)

$\mathit{CostInfInv}^T$: Cost on transportation infrastructure (\$)

$\mathit{eCap}_{(i,j)}(\mathbf{t})$: Capacity for energy arc from node i to node j, for time step $\mathbf{t}$. (MW)

$fleetCap_{(i,j,m)}(\mathbf{t})$: Fleet m capacity for transportation arc from node i to node j during time step $\mathbf{t}$ (ton/h)

$infCap_{(i,j,l)}(\mathbf{t})$: Infrastructure l capacity for transportation arc from node i to node j for time step $\mathbf{t}$ (ton/h)

$d_j^{ET}(\mathbf{t})$: Energy demand at node j during time $\mathbf{t}$ due to the transportation of commodities (MWh)

3.2.7 Other parameters and functions

r: Discount rate

P_E: Power base for the electric system

I[$test$]: Indicator function that equals 1 if $test$ is true and 0 otherwise

3.3 General formulation

The following formulation formally incorporates the modeling capabilities that have been described in previous chapters and allow the joint optimization of the transportation and energy networks.

The objective to be minimized (3.2) is the total cost of operating and investment of the combined system. We can distinguish five components in the objective: operational cost of the energy system (3.16), investment cost of the energy system (3.17), operational cost of the transportation system (3.18), and investment cost of the transportation system divided into fleet investment (3.19) and infrastructure investment (3.20).

The energy sector is represented by a set of nodes $\mathcal{N}^E$ and arcs $\mathcal{A}^E$. Each node represents an energy subnetwork in a geographic location. For example, for a particular location, there might be three energy nodes; one for electricity, one for gas, and one for petroleum. Energy arcs link these various nodes, both within the same subnetwork (representing transmission lines, natural gas pipelines, or petroleum pipelines) or between different networks, where conversion takes place (e.g., power plants).

$$\textbf{min} \quad \{CostOp^E + CostInv^E + CostOp^T + CostFleetInv^T + CostInfInv^T\} \tag{3.2}$$

subject to:

Meet energy demand at the appropriate nodes

$$\sum_i \eta_{(i,j)}(\mathbf{t}) e_{(i,j)}(\mathbf{t}) - \sum_i e_{(j,i)}(\mathbf{t}) = d_j^E(\mathbf{t}) + d_j^{ET}(\mathbf{t}) \quad j \in \mathcal{N}_d^E \tag{3.3}$$

DC power flow equations

$$e_{(i,j)}(\mathbf{t}) - e_{(j,i)}(\mathbf{t}) = b_{(i,j)}\Big(\theta_i(\mathbf{t}) - \theta_j(\mathbf{t})\Big)\, P_E\, \Delta(\mathbf{t}), \quad (i,j) \in \mathcal{A}_{DC}^E \tag{3.4}$$

Generation capacity must cover peak demand at electric nodes

$$\sum_i cf_{(i,j)}(\mathbf{t})\, eCap_{(i,j)}(\mathbf{t}) \geq peakD_j^E(\mathbf{t}), \quad j \in \mathcal{N}_p^E \tag{3.5}$$

Transportation demand for non-energy commodities

$$\sum_m f_{(i,j,k,m)}(\mathbf{t}) = d_{(i,j,k)}^T(\mathbf{t}), \quad k \in \mathcal{K} \backslash \mathcal{K}_e \tag{3.6}$$

Transportation demand for energy commodities

$$\sum_m f_{(i,j,k,m)}(\mathbf{t}) = heatContent_k^{-1}(\mathbf{t})\, e_{(n_{(i,k)}^E, n_{(j,k)}^E)}(\mathbf{t}), \quad k \in \mathcal{K}_e \tag{3.7}$$

Fleet upper bound for transportation flows

$$\sum_k f_{(i,j,k,m)}(\mathbf{t}) \leq fleetCap_{(i,j,m)}(\mathbf{t})\, \Delta(\mathbf{t}) \tag{3.8}$$

Infrastructure upper bound for transportation flows

$$\sum_k \sum_{m \in \mathcal{M}_l} f_{(i,j,k,m)}(\mathbf{t}) \leq infCap_{(i,j,l)}(\mathbf{t})\, \Delta(\mathbf{t}) \tag{3.9}$$

Decision variables:

$$\text{Energy flows:} \quad 0 \leq lbe_{(i,j)}(\mathbf{t}) \leq e_{(i,j)}(\mathbf{t}) \leq eCap_{(i,j)}(\mathbf{t})\, \Delta(\mathbf{t}) \tag{3.10}$$

$$\text{Energy capacity inv.:} \quad lbeInv_{(i,j)}(\mathbf{t}) \leq eInv_{(i,j)}(\mathbf{t}) \leq ubeInv_{(i,j)}(\mathbf{t}) \tag{3.11}$$

$$\text{Transportation flows:} \quad f_{(i,j,k,m)}(\mathbf{t}) \geq 0 \tag{3.12}$$

$$\text{Fleet inv.:} \quad lbFleetInv_{(i,j,m)}(\mathbf{t}) \leq fleetInv_{(i,j,m)}(\mathbf{t}) \leq ubFleetInv_{(i,j,m)}(\mathbf{t}) \tag{3.13}$$

$$\text{Infrastucture inv.:} \quad lbInfInv_{(i,j,l)}(\mathbf{t}) \leq infInv_{(i,j,l)}(\mathbf{t}) \leq ubInfInv_{(i,j,l)}(\mathbf{t}) \tag{3.14}$$

$$\text{Phase angles:} \qquad -\pi \leq \theta_i(\mathbf{t}) \leq \pi \tag{3.15}$$

where:

$$CostOp^E = \sum_{\mathbf{t}} \sum_{(i,j)} (1+r)^{-\mathbf{t}_1}\, costOp^E_{(i,j)}(\mathbf{t})\, e_{(i,j)}(\mathbf{t}) \tag{3.16}$$

$$CostInv^E = \sum_{\mathbf{t}} \sum_{(i,j)} (1+r)^{-\mathbf{t}_1}\, costInv^E_{(i,j)}(\mathbf{t})\, eInv_{(i,j)}(\mathbf{t}) \tag{3.17}$$

$$CostOp^T = \sum_{\mathbf{t}} \sum_{(i,j,k,m)} (1+r)^{-\mathbf{t}_1}\, costOp^T_{(i,j,k,m)}(\mathbf{t})\, f_{(i,j,k,m)}(\mathbf{t}) \tag{3.18}$$

$$CostFleetInv^T = \sum_{\mathbf{t}} \sum_{(i,j,m)} (1+r)^{-\mathbf{t}_1}\, costFleetInv_{(i,j,m)}(\mathbf{t})\, fleetInv_{(i,j,m)}(\mathbf{t}) \tag{3.19}$$

$$CostInfInv^T = \sum_{\mathbf{t}} \sum_{(i,j,l)} (1+r)^{-\mathbf{t}_1}\, costInfInv_{(i,j,l)}(\mathbf{t})\, infInv_{(i,j,l)}(\mathbf{t}) \tag{3.20}$$

$$eCap_{(i,j)}(\mathbf{t}) = ube_{(i,j)}(\mathbf{t}) + \sum_{\boldsymbol{\tau} \in \mathcal{T}_{\mathbf{t}}} eInv_{(i,j)}(\boldsymbol{\tau})\, \mathrm{I}[\mathbf{t}_1 - \boldsymbol{\tau}_1 \leq eLife_{(i,j)}] \tag{3.21}$$

$$fleetCap_{(i,j,m)}(\mathbf{t}) = ubFleet_{(i,j,m)}(\mathbf{t}) + \sum_{\boldsymbol{\tau} \in \mathcal{T}_{\mathbf{t}}} fleetInv_{(i,j,m)}(\boldsymbol{\tau})\, \mathrm{I}[\mathbf{t}_1 - \boldsymbol{\tau}_1 \leq fleetLife_{(i,j,m)}] \tag{3.22}$$

$$infCap_{(i,j,l)}(\mathbf{t}) = ubInf_{(i,j,l)}(\mathbf{t}) + \sum_{\boldsymbol{\tau} \in \mathcal{T}_{\mathbf{t}}} infInv_{(i,j,l)}(\boldsymbol{\tau})\, \mathrm{I}[\mathbf{t}_1 - \boldsymbol{\tau}_1 \leq infLife_{(i,j,k,m)}] \tag{3.23}$$

$$d^{ET}_j(\mathbf{t}) = \sum_{(a,b) \in \mathcal{A}^T_j} \sum_{m \in \mathcal{M}_j} fuelCons_{(a,b,m)}(\mathbf{t}) \sum_{k} f_{(a,b,k,m)}(\mathbf{t}) \tag{3.24}$$

The flows across these arcs, $e_{(i,j)}$, are part of the decision variables of the problem. These flows must be such that they satisfy the demand for energy (3.3), part of which is due to the fuel required to perform the movement of commodities in the transportation system (3.24). Energy flows are also required to meet lower and upper bound constraints (3.10). Upper bounds are determined by the capacity of the existing energy infrastructure, $eCap_{(i,j)}(\mathbf{t})$, and are a combination (3.21) of the initial capacity, $ube_{(i,j)}(\mathbf{t})$, and investment on upgrades, $eInv_{(i,j)}(\mathbf{t})$. The former are given parameters while the latter are also decision variables with their corresponding lower and upper bounds (3.11). Energy flows must also satisfy other constraints, such as DC power flow equations for electric transmission (3.4). Equation (3.5) ensures that generation for each electric region is able to cover the peak demand.

The transportation system is also formed by a set of nodes $\mathcal{N}^T$ and a set of arcs $\mathcal{A}^T$, although the nomenclature is slightly different. Here each node represents a unique geographic location, while the transportation flows (3.12) are referred to as $f_{(i,j,k,m)}$, where (i,j) represents the origin and destination nodes, k the commodity that is being transported and m the mode of transportation used.

These flows must satisfy the transportation demand, which is predetermined (3.6) for all commodities except for those that are energy related (e.g., coal, uranium). In that case transportation demands depend on the use of that commodity in the energy system (3.7). Transportation flows are constrained (3.8, 3.9) by the capacity (3.22) of the available fleet, $\mathit{fleetCap}_{(i,j,m)}(\mathbf{t})$, and the capacity (3.23) transportation infrastructure $\mathit{infCap}_{(i,j,l)}(\mathbf{t})$. Both can be increased by the their respective investments, $\mathit{fleetInv}_{(i,j,m)}(\mathbf{t})$ and $\mathit{infInv}_{(i,j,m)}(\mathbf{t})$, which have upper and lower bound constraints (3.13, 3.14).

3.4 Compact notation

A more compact version of the previous formulation can be produced using vectorial and matrix notation. First of all, the vector **capInv** includes all capacity investment variables, while the operational variables are grouped in vectors called $\mathbf{flows}_t$. The subscripts in this section represent the top level time step, t, to which a variable belongs, i.e., $\mathbf{t}_1$ following the notation from the previous section.

When using this notation we can reduce the model to the expressions (3.25).

$$\mathbf{min} \quad \begin{bmatrix} \mathbf{CostInv}^\top & \mathbf{CostOp}_1^\top & \mathbf{CostOp}_2^\top & \dots \end{bmatrix} \begin{bmatrix} \mathbf{capInv} \\ \mathbf{flows}_1 \\ \mathbf{flows}_2 \\ \vdots \end{bmatrix}$$

subject to:

$$\begin{bmatrix} -\mathbf{I} & \mathbf{0} & \mathbf{0} & \dots \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{A}_1 & \mathbf{0} & \dots \\ \mathbf{0} & -\mathbf{I} & \mathbf{0} & \dots \\ \mathbf{L}_1 & \mathbf{I} & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_2 & \dots \\ \mathbf{0} & \mathbf{0} & -\mathbf{I} & \dots \\ \mathbf{L}_2 & \mathbf{0} & \mathbf{I} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \mathbf{capInv} \\ \mathbf{flows}_1 \\ \mathbf{flows}_2 \\ \vdots \end{bmatrix} \leq \begin{bmatrix} -\mathbf{lbInv} \\ \mathbf{ubInv} \\ \mathbf{d}_1 \\ -\mathbf{lb}_1 \\ \mathbf{ub}_1 \\ \mathbf{d}_2 \\ -\mathbf{lb}_2 \\ \mathbf{ub}_2 \\ \vdots \end{bmatrix} \tag{3.25}$$

The objective value is (3.2), which is developed in expressions (3.16-3.20). The first two rows in the constraints are directly related to the minimum and maximum investment allowed (3.11, 3.13, 3.14).

Following that there are three rows that are repeated for each year. The first two correspond to operational constraints, such as demand balances (3.3, 3.6, 3.7) or DC power flow constraints (3.4). Such expressions are summarized in the matrix $\mathbf{A}_t$ and the vector $\mathbf{d}_t$. The next two lines correspond to the minimum and maximum operational flows from (3.8–3.10, 3.21–3.23). The matrix $\mathbf{L}_t$ accounts for the fact that only investments made prior to the operational year can account for available capacity.

3.5 Decomposition

Inspection of (3.25) reveals an underlying structure (3.26) that would allow the use of Benders decomposition methods [56]. This approach is not new to the field, and it is implemented in EGEAS [57], a heavily used investment software application used in the electric industry. This special structure is used to reduce the complexity of the linear problem and possibly use parallelization to increase solution speeds.

$$\begin{aligned} &\mathbf{min} \quad \begin{bmatrix} \mathbf{CostInv}^\top & \mathbf{CostOp}_1^\top & \mathbf{CostOp}_2^\top & \dots \end{bmatrix} \begin{bmatrix} \mathbf{capInv} \\ \mathbf{flows}_1 \\ \mathbf{flows}_2 \\ \vdots \end{bmatrix} \\ &\textbf{subject to:} \\ &\qquad \begin{bmatrix} \mathbf{C}_0 & \mathbf{0} & \mathbf{0} & \dots \\ \mathbf{C}_1 & \mathbf{D}_1 & \mathbf{0} & \dots \\ \mathbf{C}_2 & \mathbf{0} & \mathbf{D}_2 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \mathbf{capInv} \\ \mathbf{flows}_1 \\ \mathbf{flows}_2 \\ \vdots \end{bmatrix} \leq \begin{bmatrix} \mathbf{b}_0 \\ \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \end{bmatrix} \end{aligned} \tag{3.26}$$

Thus, the master problem can be formulated as (3.27), where the variables z_i are used to estimate the objective value of the subproblems

$$\begin{aligned} &\mathbf{min} \quad \mathbf{CostInv}^\top \cdot \mathbf{capInv} + z_1 + z_2 \dots \\ &\textbf{subject to:} \\ &\qquad \mathbf{C}_0 \cdot \mathbf{capInv} \leq \mathbf{b}_0 \end{aligned} \tag{3.27}$$

For each year $i = 1, 2, \dots, n$ we have the following subproblem (3.28) once the master has been solved

$$\begin{aligned} &\mathbf{min} \quad \zeta_i = \mathbf{CostOp}_i^\top \cdot \mathbf{flows}_i \\ &\textbf{subject to:} \\ &\qquad \mathbf{C}_i \cdot \mathbf{flows}_i \leq \mathbf{b}_i - \mathbf{D}_i \cdot \mathbf{capInv}^* \end{aligned} \tag{3.28}$$

These subproblems can be solved in parallel to further reduce computation time. Once determined, the solution status of each of these subproblems can be classified in one of the following three cases:

First Case: If subproblem (3.28) is infeasible, by a variation of the Farkas Lemma, there exists an extreme ray for its dual (3.29):

$$\begin{aligned} &\textbf{max} \quad \zeta_i = (\mathbf{b}_i - \mathbf{D}_i \cdot \mathbf{capInv}^*)^\top \mathbf{y}_i \\ &\textbf{subject to:} \\ &\quad \mathbf{C}_i^\top \cdot \mathbf{y}_i \geq \mathbf{CostOp}_i \end{aligned} \tag{3.29}$$

An extreme ray of (3.29) is a vector, $\Delta\mathbf{y}_i^*$, such that $\mathbf{C}_i^\top \cdot \Delta\mathbf{y}_i^* \geq \mathbf{0}$ and makes the dual infeasible since it meets $(\mathbf{b}_i - \mathbf{D}_i \cdot \mathbf{capInv}^*)^\top \Delta\mathbf{y}_i^* > 0$. A cut is then added to ensure feasibility of the subproblem

$$(\mathbf{b}_i - \mathbf{D}_i \cdot \mathbf{capInv})^\top \Delta\mathbf{y}_i^* \leq 0 \tag{3.30}$$

Second Case: If subproblem (3.28) and its dual (3.29) are optimal with solution $(\mathbf{flows}_i, \mathbf{y}_i)$ but $\zeta_i^* > z_i^*$, then z_i^* is not a realistic approximation of the subproblem objective value. A cut is then added to the master so that z_i should be more realistic

$$(\mathbf{b}_i - \mathbf{D}_i \cdot \mathbf{capInv})^\top \mathbf{y}_i^* \leq z_i \tag{3.31}$$

Third Case: If the subproblem (3.28) is optimal and $\zeta_i^* \leq z_i^*$ then z_i^* is a good estimation of the subproblem objective value, so no cuts are added to to subproblem.

Once all the subproblems have been solved the process stops if no cuts have been added in the last iteration, having obtained the optimal solution. Otherwise, if at least one cut has been added, the master problem is solved again with the added constraints. The new solution for **capInv** is then passed to the subproblems and these are solved again until no more cuts are necessary.

3.6 Multiobjective optimizer

A multiobjective optimization is proposed for NETPLAN based on an evolutionary algorithm to efficiently solve the problem described and produce an approximation to the Pareto front. Evolutionary algorithms are increasingly being applied in different fields, including engineering problems [58, 59]. As with the rest of the software, the multiobjective optimizer has been conceived in a modular fashion, to allow the parallel development of each one of its components. Two levels can be distinguished in this design, as depicted in Fig. 3.1.

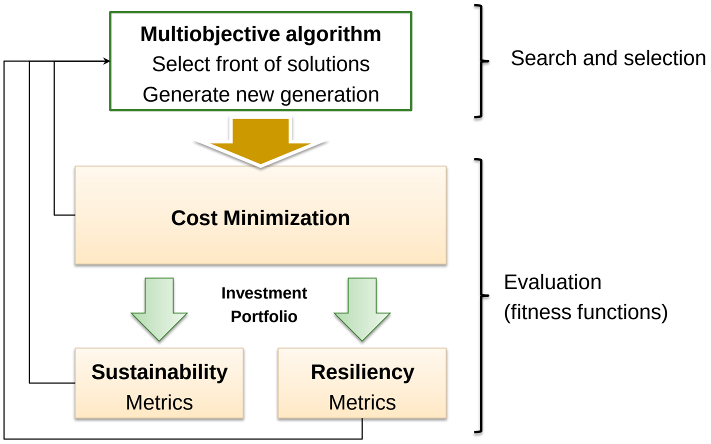


Figure 3.1: NETPLAN multiobjective approach

The high-level optimization is directed by the NSGA-II[1] algorithm [60], which is further explained in the next section. Each individual within a given population is characterized by a minimum level of investment that is to be enforced. Thus, a vector of minimum values for **capInv** is created for each candidate solution and it is evaluated by the fitness functions, represented here by the metric blocks.

The first step in the evaluation consists on solving the minimum objective cost problem solved while the the minimum investment constraints are incorporated. Previous sections capture the modeling of the energy and transportation systems for the cost minimization. This process produces the investment portfolio for the candidate solutions as well as the flows representing how the system is operated. Other economic factors, such as energy and transportation prices or economic opportunities [53], can also be extracted.

Once the portfolio is determined, metrics for sustainability and resiliency are calculated. These functions are based on the principles described in Chapter 2. For sustainability, simple linear functions are used to calculate total emissions or an emission index.

---

[1] Non-dominated sorting genetic algorithm II

Resiliency evaluation consists of the evaluation of the operational parts of the model, with the system stressed. This is performed by taking the vector of capacities **capInv** that was determined by the minimum cost problem and modifying it to simulate events. This usually involves reducing the capacity of affected arcs. This modified vector of capacities is applied to the pertinent operational subproblems (3.28) and the increase in cost of operating the system is calculated. This is performed for the predetermined number of events and the resiliency metric is computed as a function of these cost increases for all defined events. In the work reported in Chapter 5, the resiliency metric is calculated as the average cost increase across events.

The objective values for cost, sustainability and resiliency are then recovered and sent back to the high-level optimization in order to produce the next generation in search for the Pareto front. This loop could also be used in the future to guide the multiobjective optimizer towards a richer variety of solutions. For example, from the dual cost minimization problems one could identify more economic alternatives or, from the evaluation of resiliency, one could identify the weakest links in the system and reinforce them in future solutions. However, including these features would require major modifications in the multiobjective optimizer.

At this point, the use of minimum investment vectors facilitates shifting from the minimum cost solutions to others with better performance in terms of sustainability or resiliency. This approach is very simple to implement but, if needed, could be complemented with other methods such as subsidies for clean alternatives, taxes for polluting technologies (e.g., carbon tax) or preventing certain technologies to be used in favor of others.

3.7 The NSGA-II algorithm

The previous section presented the structure of the multiobjective optimizer to determine the best investment portfolios with respect to cost, resiliency and sustainability. This section introduces a description of the non-dominated sorting genetic algorithm II (NSGA-II) [60], which directs the high-level optimization in NETPLAN.

The main features of this approach with respect to other evolutionary algorithms are a fast sorting procedure, an elitist approach, lack of parameters, and diversity preservation. The first two ensure an improvement in the overall performance and speed of the algorithm. The third is desired since the optimal determination varies between problems. The last feature provides solutions that cover the approximation to the Pareto front much more uniformly.

Dominance and crowding-distance are two notions on which the algorithm is based and they are worth introducing prior to the description of NSGA-II:

Dominance A multiobjective optimization problem with conflicting objectives does not have a single solution. Instead, the Pareto front of solutions is defined as the set of solutions that are not dominated by any other. A solution dominates another one if all its objective values are equal or better and at least one of them is strictly better. Assuming that the objective values in Fig. 3.2 are to be minimized, all the filled circles are non-dominated.

It is also possible to determine the different *fronts* in a set of solutions. The first front is formed by the non-dominated solutions. The next front is formed by the solutions that are only determined by the first front, and so on. Define the non-domination *rank* as the front number that a solution belongs to, e.g., solutions in the first front have rank 1.

Crowding-distance To ensure the diversity of solutions the concept of crowding-distance is defined as an approximation of density estimation. The principle is that if a solution is close to other ones, it will score poorly. Crowding-distance is calculated taking into account the members of a single front.

The process begins by sorting the solutions for each objective. The extreme points (labeled “0” and “1” in Fig. 3.2) are assigned an infinite value for crowding-distance to ensure that these are preserved. Then for the remaining points the crowding-distance is calculated as the average distance of the adjacent solutions, e.g., the average of the box drawn around solution “i” in Fig. 3.2.

The NSGA-II algorithm can be applied to problems with constraints and a candidate solution could violate one or more than these constraints. The following is the sorting method

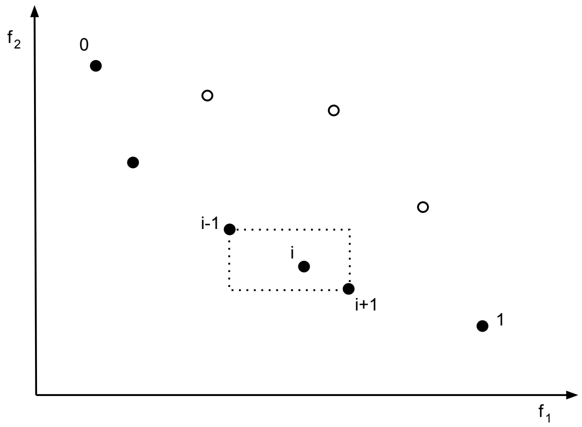


Figure 3.2: Dominance and crowding-distance calculation. Points market in filled circles are solutions of the same non-dominated front.

used to account for all of the above. Solution i is said to constrained-dominate a solution j if any of the following is true:

1. Solution i is feasible and solution j is not.
2. Solution i and j are both infeasible, but solution i has a smaller overall constraint violation.
3. If solutions i and j are feasible, then
   - (a) Solution i has a smaller rank than solution j.
   - (b) Solution i and j have the same rank, but solution i has a larger crowding-distance.

Now that all the necessary concepts have been defined it is time to introduce the NSGA-II algorithm. A graphical representation is included in Fig. 3.3.

Step 1 Randomly determine the first parent population, P_0, and evaluate it. P_0 has N elements.

Step 2 For the parent in the i-th generation, P_i, create child population C_i by binary tournament selection, recombination, and mutation operators. C_i has N elements.

Step 3 Combine P_i and C_i to create the mixed population, M_i, with size $2N$.

Step 4 Sort M_i by the method mentioned above. The different fronts $(F_1, F_2, F_3$, etc.) are organized in the list.

Step 5 Create the parent of the next generation, P_{i+1}, by copying the first N elements in the ordered mixed population. The solutions that are not copied are rejected.

Step 6 If the number of generations has been reached, STOP. Otherwise, increase generation number and continue with **Step 2**.

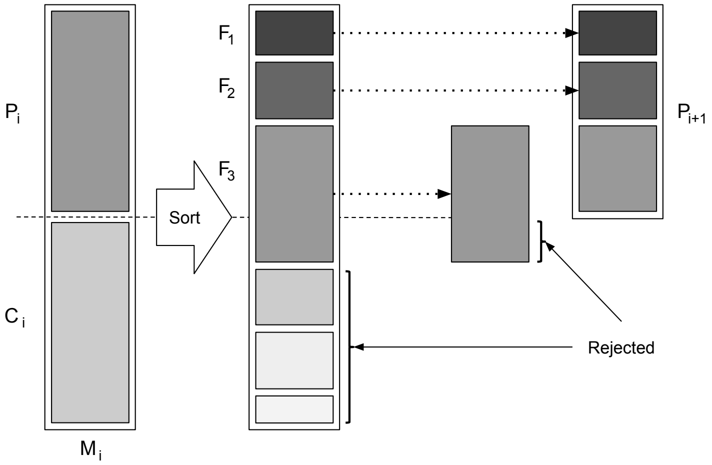


Figure 3.3: NSGA-II procedure

3.8 Parallel multiobjective optimizer

A parallelization scheme for the NSGA-II algorithm is presented in this section. This approach is being presently implemented and will allow for the deployment of NETPLAN on machines with multiple clusters, greatly reducing total computation time.

The sequential NSGA-II algorithm from the previous sections can be summarized with the flow chart shown on the left side of Fig. 3.4. In this algorithm the child population in the i-th generation is created by crossover and mutation methods from the previous best solution. This child population is evaluated and mixed with the current best solutions. Through tournaments the best selection is updated.

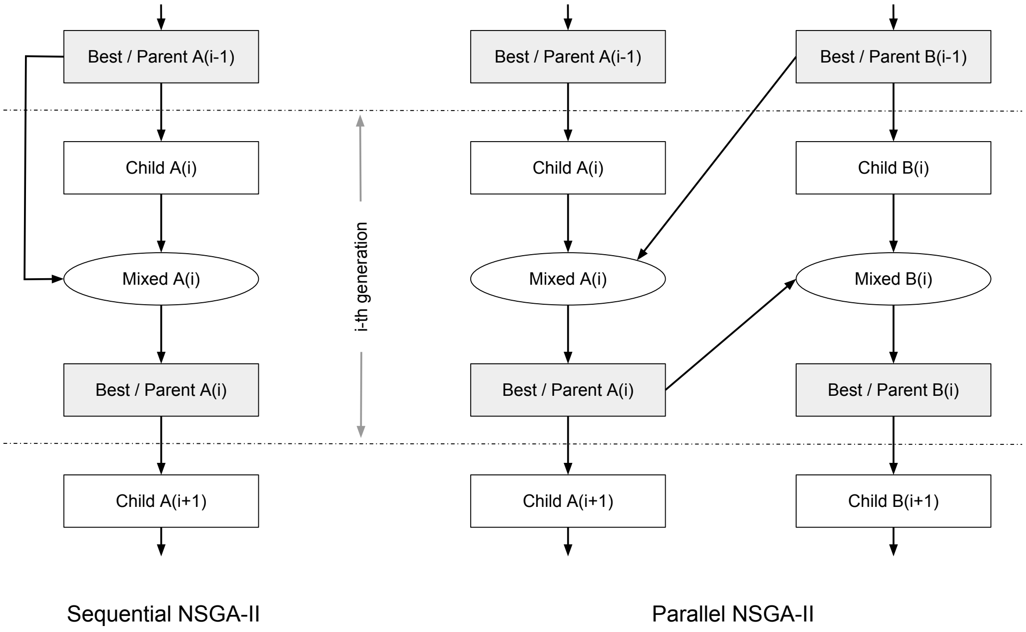


Figure 3.4: Flow charts for the sequential and parallel NSGA-II

The proposed parallelization preserves many of the functions of the sequential version so its implementation is straightforward. The flow chart in Fig. 3.4 complements the following description of the algorithm:

Step 1 Initialize Parent A1 and send elements to the queue to be evaluated

Step 2 Initialize Parent B1 and send elements to the queue

Step 3 Receive evaluation of Parent A1, create Child A2 from it and send elements to the queue

Step 4 Receive evaluation of Parent B1, create Child B2 from it and send elements to the queue

Step 5 Receive elements for Child A2, mix with Parent B1 and create Parent A2

Step 6 Create Child A3 from Parent A2 and send to the queue

Step 7 Receive elements for Child B2, mix with Parent A2 and create Parent B2

Step 8 If number of generations reached, STOP. Otherwise, create child B3, send to the queue and continue with **Step 5** for the next generation

If the number of available nodes in the parallel machine exceeds the number of desired elements per generation, this algorithm could be easily expanded to accommodate more parallel instances by connecting them in a circular fashion.

3.9 Software structure

The software implementation of NETPLAN is written in C++ and it utilizes ILOG CPLEX Concert Technology [61] to perform the linear optimization. A generic version of the NSGA-II evolutionary algorithm [60] is used as the base for the multiobjective optimization. This implementation is capable of performing most of the features described in this work. For more information on how to use the model and develop models, a small introduction in the form of user manual is included in Appendix B.

The software is designed with three objectives in mind: modularity, data-driven, and reproducibility. All these qualities are required to be the software used by the members of the NETSCORE-21 research project.

The source code is divided in a series of modules or libraries that are organized by functionality. This makes the code easier to understand and new features can be developed by different people in parallel.

The definition of the model to be studied (networks, parameters, objectives) is data-driven, that is, it is performed completely through input parameters and independently from the code. Because of this, a user can implement models without needing to understand the code behind NETPLAN. Also, models can be further explored by doing simple modifications in the input files.

The same set of input files always produces the same results so studies can be repeated and verified by different users. The basis for input and output files are the comma-separated value and MPS[2] formats. These file types are plain-text based and, thus, they can be easily

---

[2] Mathematical Programming System, file format for presenting and archiving linear programming.

modified and shared.

The implementation of NETPLAN is divided into three separate programs. Figure 3.5 shows the order in which these programs are executed (as white blocks) and their relationship to the input and output files (gray blocks).

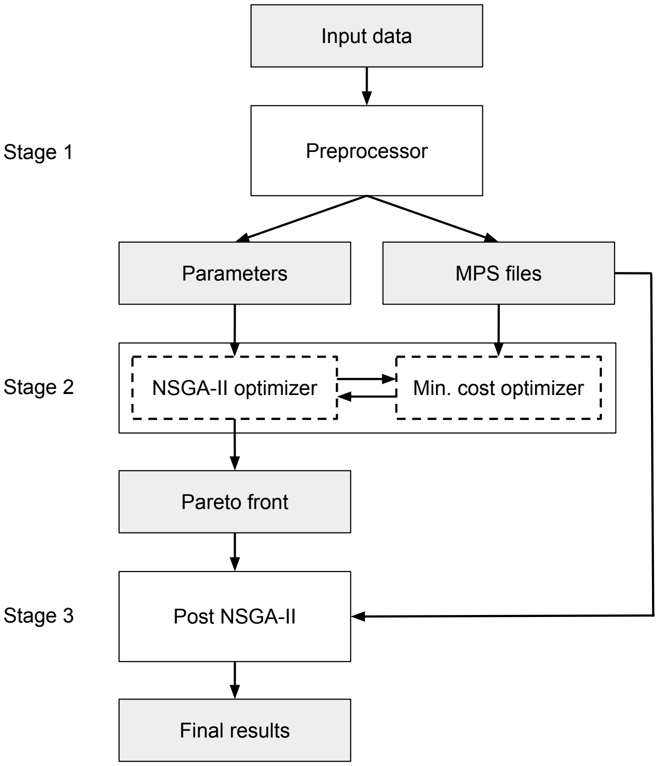


Figure 3.5: Software structure in NETPLAN

The stage 1 executable is called as the preprocessor and its only goal is to read the input files that define the model and create the appropriate file for the optimization, which is the second stage. The main steps followed are:

- Read node list file,
- Read arc list file,
- Read parameter files,
- Expand node list, according to their step size,

- Expand arcs list,
- Write the linear programs in MPS files,
- Write files with auxiliary parameters.

The NSGA-II optimization commands stage 2 and it alternates with the minimum cost optimizer to find the Pareto front of solutions. This process is done within the same program although a parallel implementation, introduced in the next section, is being developed but is not operational yet.

Once the optimizer has finalized the user is provided with the identification of the Pareto front of solutions. The resulting file contains the objective values and the minimum investments that characterize each solution. To obtain more information about these solutions a postprocessor for the NSGA-II, stage 3, is provided. This process could be expanded in the future to include analysis tools that provided additional insight similar to what was performed in the numerical results for this work.

3.10 Modeling questions

Modeling investment futures for the energy and transportation systems is an ambitious vision and there are always improvements that could be implemented. The main objective of this work in this sense is to create a framework and basic modeling approach for the problem. This section explores a number of assumptions that have been made and discusses how further features that would strengthen could be implemented.

3.10.1 Uncertainty

In its present form NETPLAN is a deterministic tool since the evaluation of investments are primarily based around one single forecast for model parameters. However, long-term planning involves uncertainty, which can be defined as "a potential deficiency in any phase or activity of the modeling process that is due to the lack of knowledge" [62].

NETPLAN already incorporates mechanisms to evaluate the resiliency and robustness of the system against disruptions. This methodology could be expanded to include not only large-consequence events but also smaller disturbances in terms of capacity loss or changes in other parameters (e.g., investment or operational costs, efficiencies, resource availability).

Another approach would be to convert the deterministic cost minimization into a full stochastic model. Such a formulation would increase the model size and thus computation times. However, the proposed decomposition methods would still apply in this case and they would help minimize the effect of the increase in the model complexity.

An alternative would be to consider a number of scenarios which span the space of possible futures. With the existing methods the optimal investment portfolios for each scenario could be found. The comparison of these results would provide insight into which facilities are most commonly used. Additionally, each portfolio could be evaluated for the rest of scenarios and then ranked according their performance across them.

3.10.2 Passenger formulation

As it was mentioned in Chapter 2, the passenger formulation has not been fully developed. The numerical example in this work presents a simple example where the investment decision for personal vehicles are divided between a regular gasoline vehicle and a plug-in hybrid vehicle.

It is not the objective of this thesis to develop a passenger model. However, the modeling framework developed for commodity transportation could serve as an excellent starting point for developing a modeling framework for passenger transportation. Such a model would be enhance to take into account, among others, the issues pointed out in Section 2.3: individual decisions, travel patterns, passenger preferences (cost vs. travel time, speed, comfort, independence), or air travel routes.

3.10.3 Carbon taxes or cap-and-trade

Carbon taxes, caps or cap-and-trade methods are methods that could be used to curb greenhouse emissions in the future. These are usually implemented through legislation and most planning tools provide the option of including them.

NETPLAN features the possibility of limiting national emissions on a yearly basis and this will be explored in Chapter 5. Implementing a carbon tax is equivalent to adding CO_2 emissions to the cost minimization problem with the appropriate cost per ton emitted. This feature is not currently implemented in NETPLAN but this would be simple to do since emissions are already being accounted for in the model.

These methodologies could be extended to other sustainability metrics in order to produce better solutions. In fact, emission limits or taxes could be made part of the decision variables in the multiobjective optimizer. With this change, the NSGA-II algorithm would search not only for the ideal investment portfolios but also for the most effective pieces of legislation.

3.10.4 Electric operational constraints

The introduction of certain renewable technologies (most notably wind and solar) could have an adverse effect on the performance on the system since they are intermittent resources. Numerous studies are being performed to understand what the system configuration should be to allow high penetrations of these technologies [63, 64]. On the other hand, these resources could be coupled with a wide deployment of electric storage facilities (batteries, compressed air, flywheels) to address variability.

A comprehensive study of these issues could provide some insight to model these interesting operational interactions. The result could be expressed in terms of operational constraints for the electric system, e.g., ensuring that minimum regulation reserve is met including the variability from renewables, or that available reserve can satisfy a determined ramp rate.

Another issue related to the operations would be the increased stress to which existing units could be exposed due to high variability. Coal generators are not designed to be engage in rapid cycling which they could be forced to do if the penetration of intermittent resources is considerable. This cycling could cause a higher rate of deterioration of the units and effectively reduce their projected life span.

3.10.5 Investment and DC power flow model

NETPLAN allows for the inclusion of DC power flow equations to simulate the behavior of electricity flow in the electricity network. As captured in (3.4) line susceptances, $b_{(i,j)}$, are crucial parameters in the definition of these equations. Unfortunately, investment on new power lines modifies the topology of the system and, thus, the susceptances. The present version of NETPLAN ignores those changes although one could obtain the solution, modify the susceptances accordingly and iterate until the process converges.

Relating susceptances directly with investment variables would result in a nonlinear problem, which would be significantly harder to solve. A formulation to include a single new connection between electric nodes is presented in [65] and utilizes a mixed-integer programming. This methodology can be expanded to accommodate additions to existing arcs, even considering more than one option.

Assume that for each electric arc $(i,j) \in \mathcal{A}^E_{DC}$ we consider S possible upgrades. We will use the binary variable $z_{(i,j),s}(\mathbf{t})$ to represent that upgrade $s_{(i,j)}$ is applied to arc (i,j) at time $\mathbf{t}$, with $s \in \{0,\dots,S\}$. Consider $s_{(i,j)} = 0$ to be the case in which no upgrade is made.

Upgrades are considered to be mutually exclusive, thus

$$z_{(i,j),0}(\mathbf{t}) + z_{(i,j),1}(\mathbf{t}) + \ldots + z_{(i,j),S}(\mathbf{t}) = 1 \tag{3.32}$$

is enforced for every time step.

The total cost of the $s_{(i,j)}$-th upgrade, $totalCostInv^E_{(i,j),s}$, can then be associated with its corresponding binary variable. The calculation of total investment cost in (3.17) can be updated to reflect this:

$$\begin{aligned} CostInv^E = \sum_{\mathbf{t}} \sum_{(i,j)} (1+r)^{-\mathbf{t}_1} \Big[& costInv^E_{(i,j)}(\mathbf{t})\, eInv_{(i,j)}(\mathbf{t}) + \\ & + \sum_{s=1}^{S} totalCostInv^E_{(i,j),s}(\mathbf{t}) z_{(i,j)s}(\mathbf{t}) \Big] \end{aligned} \tag{3.33}$$

We also define the binary variables $x_{(i,j),s}(\mathbf{t})$ to represent that the $s_{(i,j)}$-th upgrade has been performed by time $\mathbf{t}$, which is derived in the following way:

$$x_{(i,j),s}(\mathbf{t}) = \sum_{\boldsymbol{\tau} \in \mathcal{T}_{\mathbf{t}}} z_{(i,j),s}(\boldsymbol{\tau}) \tag{3.34}$$

Assuming that the s-th investment provides an increase in capacity $capInv^E_{(i,j),s}$, the capacity of an electric transmission arc in (3.21) would be calculated as follows:

$$eCap_{(i,j)}(\mathbf{t}) = ube_{(i,j)}(\mathbf{t}) + \sum_{s_{(i,j)}=1}^{S} x_{(i,j),s}(\mathbf{t})\, capInv^E_{(i,j),s} \tag{3.35}$$

The last equation to be replaced is the same one that originated the need for these changes in the model, equation (3.4). It is necessary to use a large number parameter, M, and a series of continuous variables, $U_{(i,j),1}(\mathbf{t})$. Assuming that the susceptance of arc (i,j) is $b_{(i,j),s}$ after upgrade $s_{(i,j)}$ has been made, (3.4) is replaced by the following equations repeated for each possible $s_{(i,j)}$:

$$e_{(i,j)}(\mathbf{t}) - e_{(j,i)}(\mathbf{t}) = b_{(i,j),s}\Big(\theta_i(\mathbf{t}) - \theta_j(\mathbf{t})\Big)\, P_E\, \Delta(\mathbf{t}) + (x_{(i,j),s}(\mathbf{t}) - 1)M + U_{(i,j),1}(\mathbf{t}) \tag{3.36}$$

$$0 \leq U_{(i,j),1}(\mathbf{t}) \leq 2(1 - x_{(i,j),s}(\mathbf{t}))M \tag{3.37}$$

Because of the way they are defined, for each arc at each point in time only one $x_{(i,j),s}(\mathbf{t})$ can take the value of 1. For that case, (3.37) results in $U_{(i,j),1}(\mathbf{t})$ being zero and (3.36) for that s-th value becomes

$$e_{(i,j)}(\mathbf{t}) - e_{(j,i)}(\mathbf{t}) = b_{(i,j),s}\Big(\theta_i(\mathbf{t}) - \theta_j(\mathbf{t})\Big)\, P_E\, \Delta(\mathbf{t}) \tag{3.38}$$

which is equivalent to (3.4) but with the modified value for susceptance. For the rest of possible investments that are not performed, (3.36) and (3.37) become

$$e_{(i,j)}(\mathbf{t}) - e_{(j,i)}(\mathbf{t}) = b_{(i,j),s}\Big(\theta_i(\mathbf{t}) - \theta_j(\mathbf{t})\Big)\, P_E\, \Delta(\mathbf{t}) - M + U_{(i,j),1}(\mathbf{t}) \tag{3.39}$$

$$0 \leq U_{(i,j),1}(\mathbf{t}) \leq 2M \tag{3.40}$$

which is equivalent to

$$-M \leq e_{(i,j)}(\mathbf{t}) - e_{(j,i)}(\mathbf{t}) - b_{(i,j),s}\Big(\theta_i(\mathbf{t}) - \theta_j(\mathbf{t})\Big)\, P_E\, \Delta(\mathbf{t}) \leq M \tag{3.41}$$

since $U_{(i,j),1}(\mathbf{t})$ is able to take values in the [0, 2M] range. Thus, with a value of M large enough, the equality in (3.39) is not enforced, which is desirable because we are dealing with line configurations that have not been selected. But M needs not be an infinite number since the flows are constrained by the capacity and the largest angle difference is 2π. In fact, M can be defined as

$$M = \max_{\mathbf{t},s} \Delta(\mathbf{t}) \Big[ube_{(i,j)}(\mathbf{t}) + capInv^{E}_{(i,j),s} + 2\pi P_E b_{(i,j),s} \Big] \tag{3.42}$$

This formulation added to the NETPLAN model would be able to handle the new line configurations with a mixed-integer linear program. It could be used to model new connections that are not in the system by setting the initial capacity for $s_{(i,j)}$ equal to zero and including the new upgrade choices as the rest of possible configurations.

CHAPTER 4. MODEL IMPLEMENTATION

The implementation of NETPLAN requires the definition of several networks to represent the energy and transportation systems. The set of nodes and arcs belonging to these networks are endowed with a large number of parameters to model the behavior of their real-world counterparts. Throughout this chapter, the values used for those parameters are introduced along with the list of sources from which they are obtained. Unless otherwise noted all data refers to year 2008, which is the most recent update for most of the sources at the time of producing this work.

Related efforts were carried out to represent the national energy system for 2002 [66], 2005 [35, 67], and 2007 [53]. The present work utilizes similar sources to characterize the coal, natural gas, and electric sectors, although with a more refined level of granularity. Coal transportation (and freight transportation in general) has been enhanced and passenger vehicles have been introduced. The previous efforts did not considered the investment problem and were limited to minimum-cost yearly simulations of system operations. SO_2 emissions and resiliency based on sensitivity were included in the previous studies, while this work introduces greenhouse gas emissions and large-consequence event simulations to evaluate resiliency. Multiobjective optimization is also featured in this work.

4.1 Energy system

The energy system in NETPLAN is represented in this case with four different, but interconnected, subsystems: coal, natural gas, electricity and petroleum.

4.1.1 Coal subsystem

In this implementation of the national energy and transportation systems, coal is the only commodity that has a direct connection with the energy system. The life cycle for coal is comprised of production, transportation and consumption. Transportation is obviously included in the transportation system modeling, which is described in more detail in Section 4.2. Coal production and consumption are treated as part of the energy network. In all cases, coal is expressed in year steps.

Geographically, the production and movement of coal is represented at the state level. However, since coal is part of the transportation network, the coal produced at each state would be treated as a separate commodity, increasing the complexity of the model exponentially. As a compromise, some geographical aggregation is done to reduce the number of represented commodities that reflect the major characteristics for coal.

The aggregation is based on the different coal regions proposed by the Energy Information Administration in its “Coal Transportation: Rates and Trends” report [68]. Those regions correspond to the major coal production hubs in the United States and are referred to as Appalachia, Interior and West. The geographical description of this areas and their subregions is described in Fig. 4.1 and Table 4.1.

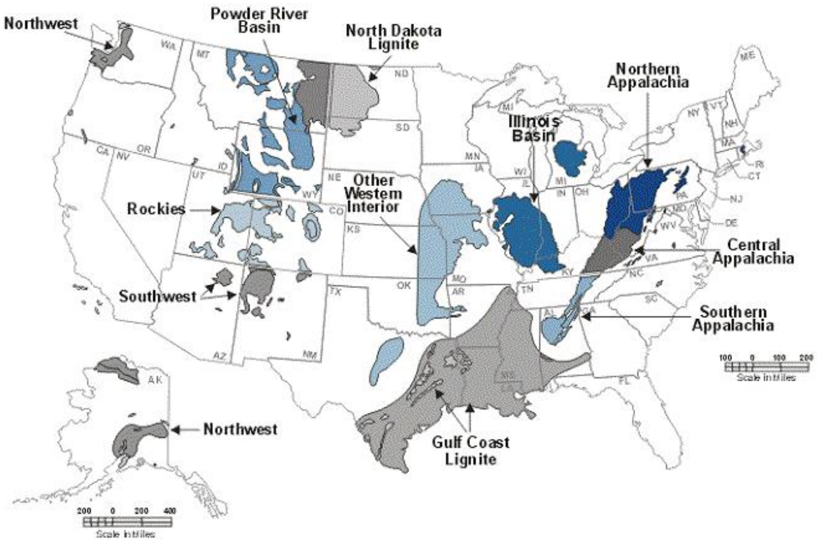


Figure 4.1: Coal regions defined by the Energy Information Administration (Source: [68])

Table 4.1: Coal regions, fields and the states they include

Coal Regions	Coal fields	States
Appalachia	Northern Appalachia	MD, OH, PA, Northern WV
	Central Appalachia	Eastern KY, VA, Southern WV
	Southern Appalachia	AL, TN
Interior	Illinois Basin	Western KY, IL, IN
	Gulf Coast Lignite	TX, LA, MS
	Other Western Interior	AR, IA, KS, MO, OK
West	Powder River Basin	WY, MT
	North Dakota Lignite	ND
	Southwest	AZ, NM
	Rockies	CO, UT
	Northwest	AK, WA

Taking into account these regions the “Monthly Cost and Quality of Fuels for Electric Plants” database (based on FERC Form 423 [69]) is analyzed. The database is summarized with four typical types of coal taking into account production costs, heat content, and sulfur and ash content. The parameters that characterize theses commodities are displayed in Table 4.2. Geographically, lignite is produced in the North Dakota and Gulf Coast fields and subbituminous in the remaining Western fields. The first type of bituminous coal is assigned to the Appalachian region and the second to the Illinois Basin and other Western Interior regions.

Table 4.2: Representation of coal types as commodities

Commodity name	Average cost ($/MMBTU)	Heat content (MMBTU/sh. ton)	Sulfur (% weight)	Ash (% weight)
Lignite	28.92	17.47	0.53	7.72
Bituminous (Appalachia)	27.69	23.86	1.39	11.11
Bituminous (Central)	29.54	19.52	0.90	6.44
Subbituminous	30.65	19.89	0.53	9.74

These four types of coals are treated as four separate commodities and, thus, they become part of the transportation network by assigning a demand node in every state.

As it was mentioned before, coal production is represented also at the state level by assigning a production node to each state. A node is connected between a fictitious coal supplying node and each of the state nodes, representing the node productions at the mines. The state production nodes are then connected to the their equivalent nodes in the transportation network. Average values for coal production capacities and costs are reported by EIA in its "Annual Coal Report" [70], more specifically in [71] and [72] and presented here in Table 4.3.

Table 4.3: Coal production capacities and costs

State	Capacity (Thousand short ton/year)	Cost ($/short ton)
Alabama	25,901	71.31
Arizona	3,500	36.28
Arkansas	3,500	40.41
Colorado	40,545	32.30
Illinois	43,129	40.41
Indiana	44,907	33.37
Kansas	3,500	40.41
Kentucky	151,622	51.33
Louisiana	3,500	28.92
Maryland	2,998	42.19
Mississippi	3,500	28.92
Missouri	3,500	40.41
Montana	49,332	12.40
New Mexico	29,750	36.28
North Dakota	32,900	13.34
Ohio	39,593	41.86
Oklahoma	1,899	47.72
Pennsylvania	75,427	49.65
Tennessee	4,811	48.94
Texas	40,553	28.92
Utah	27,042	26.91
Virginia	29,023	81.12
West Virginia	211,231	59.02
Wyoming	501,158	10.62

Between 90 and 95% of coal is consumed by the power generation sector [73] so no fixed demand is assigned to coal transport nodes. Instead, these are connected to the power plants that utilize coal as their primary generator fuel, which will be introduced in Section 4.1.3.

The volume of coal imports [74] and exports [75] are not significant with respect to the total amount of coal produced and consumed. For that reason international trade of coal is excluded from the model.

No investments or improvements are considered for the coal subsystem and it is assumed that the initial production capacity is available for the entire length of the simulation.

4.1.2 Natural gas subsystem

This section is a detail description of the natural gas subsystem, which is comprised of production, transshipment and consumption. The set of nodes and arcs are defined and their corresponding parameters (capacities, costs and efficiencies) are included.

A monthly step size is selected for natural gas in order to allow the model to capture the seasonal effects that appear in terms availability, prices and storage in the subsystem. A state-by-state representation of natural gas production and pipeline transshipment system is used. No upgrades or retirement of existing facilities are considered in the model.

Natural gas production is represented by connecting an arc from a fictitious supply node to each one of the state transshipment nodes. Two more off-shore production areas are considered and they correspond to the Gulf of Mexico (FG) and the Federal Pacific (FP) defined along the Gulf and Californian coast, respectively. Capacity production is estimated based on the "Distribution and Production of Oil and Gas Wells by State" report [76]. In that document, gas and oil wells are grouped into different groups based on their corresponding natural gas production capacity and the average for each group and state is given, so it is possible to estimate the aggregated production capacity for each of the regions, displayed in Table 4.4.

Yearly average wellhead prices displayed in the previous table are obtained from EIA [77]. But, as was mentioned before, monthly steps are going to be used to represent the natural gas network. Only national average wellhead prices are available on a monthly basis [78] as displayed on Table 4.5. An estimation of the monthly prices by state is done combining the national distribution is Table 4.5 with the average state prices from Table 4.4.

Fig. 4.2 is a good representation of the natural gas flows from the production areas towards the demand centers [79]. This picture along with the data in Table 4.4 contribute to the

Table 4.4: Natural gas production capacities

State	Capacity (MMcf/h)	Average cost ($/thousand cf)	State	Capacity (MMcf/h)	Average cost ($/thousand cf)
AL	32.38	9.65	NY	6.61	8.94
AR	73.21	8.72	OH	9.76	7.88
AZ	0.06	7.09	OK	245.89	7.56
CA	34.31	8.38	OR	0.11	5.33
CO	190.31	6.94	SD	1.46	7.94
KS	44.54	6.85	TX	1075.09	8.51
LA	187.32	8.73	UT	63.04	6.15
MI	18.31	5.63	VA	16.38	8.42
MS	15.49	8.80	WV	32.87	8.42
MT	14.67	7.50	WY	314.43	6.86
ND	11.62	8.55	FG	322.58	5.00
NE	0.42	6.22	FP	5.59	7.50
NM	178.20	8.40			

Table 4.5: Average national wellhead prices ($/thosand cubic feet)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
6.99	7.55	8.29	8.94	9.81	10.82	10.62	8.32	7.27	6.36	5.97	5.87

identify the Gulf of Mexico and Texas as the main supply areas. It is for this reason that strong hurricane seasons have an important effect on the national price of energy [46, 80].

It can also be observed that imports from Canada play an influential role in the natural gas system. Average import prices [81] in Table 4.6 also show a seasonal variation, similar to the national wellhead prices. The average import price was $8.57 per cubic feet for a total of 3.59 billion cubic feet in 2008.

Table 4.6: Average import prices ($/thosand cubic feet)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
7.60	8.27	9.01	9.58	10.74	11.62	11.65	8.48	7.38	6.66	6.43	6.57

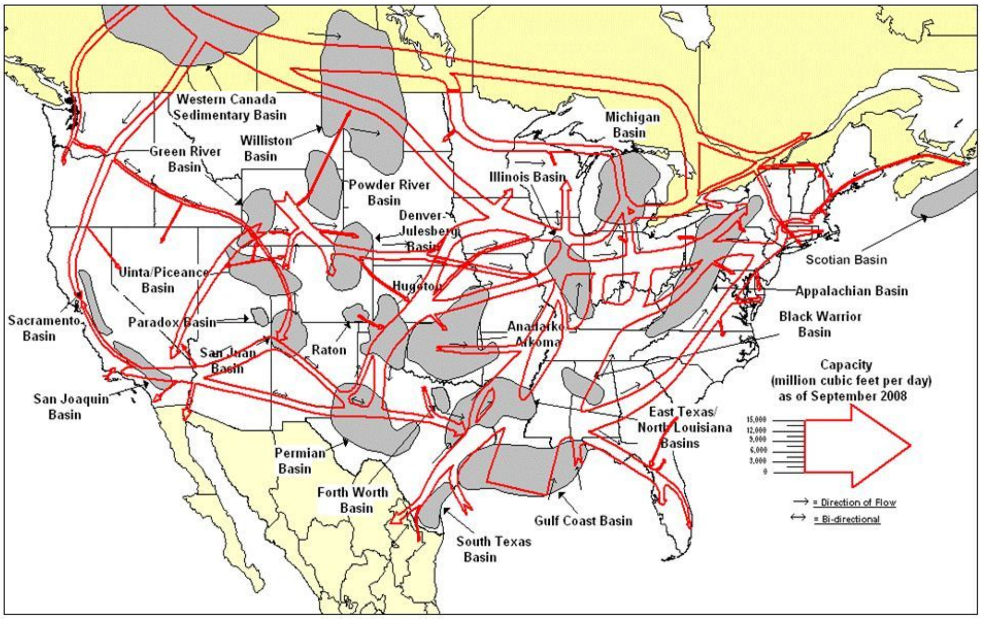


Figure 4.2: Natural gas transportation corridors (Source: [79])

Imported natural gas from Canada happens along most of the Northernmost states and for all neighboring Canadian regions. Import prices by point of entry (Fig. 4.7) are only available on a yearly basis [82]. To obtain monthly values this information is combined with Table 4.6 following the same steps taken to derive wellhead prices by state.

Table 4.7: Average import prices by point of entry

Canadian region	U.S. point of entry	Average price ($/thosand cf)
British Columbia	Washington	8.19
British Columbia	Idaho	7.88
Alberta	Montana	7.49
Saskatchewan	Montana	8.23
Manitoba	Minnesota	8.48
Ontario	Minnesota	8.50
Ontario	Michigan	8.59
Ontario	New York	8.94
Quebec	Vermont	9.74
Quebec	New Hampshire	9.72
New Brunswick	Maine	9.77

The main goal of the natural transmission gas system is to be capable of meeting the peak demand of its shippers. To meet this requirement the system, represented in Fig. 4.3, is composed of a combination of major transmission pipelines that bring the gas to the market areas and of a combination of underground natural gas storage sites and liquefied natural gas (LNG) peaking facilities located in the market areas [79].

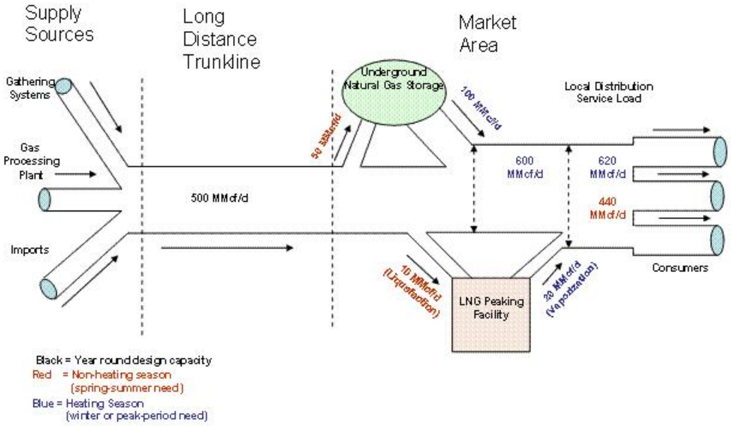


Figure 4.3: Generalized natural gas pipeline capacity design schematic (Source: [79])

The overall pipeline system configuration results in a sustained high usage level of main trunk lines through the year. For downstream facilities load factors are significantly higher during the peak-demand season and lower in the summer. This design minimizing is oftentimes referred to as peak-shaving.

Two-thirds of the lower 48 States are almost totally dependent upon the interstate pipeline system for their supplies of natural gas, as shown in Fig. 4.4. In 2008, more than 36 trillion cubic feet of natural gas (Tcf) was transported by interstate pipeline with the 30 largest interstate pipeline systems transported about 81% of the total [79].

To model the natural gas transshipment in NETPLAN the nodes corresponding to adjacent states are connected with bidirectional arcs. Interstate pipelines are reported on an individual level from EIA [83]. Their capacity is aggregated for each pair of neighboring states and applied to the corresponding arc in the model. This includes gas pipelines for Canadian imports, exports to Mexico and connecting to the Gulf of Mexico and Pacific off-shore regions. The resulting capacities are plotted in Fig. 4.5.

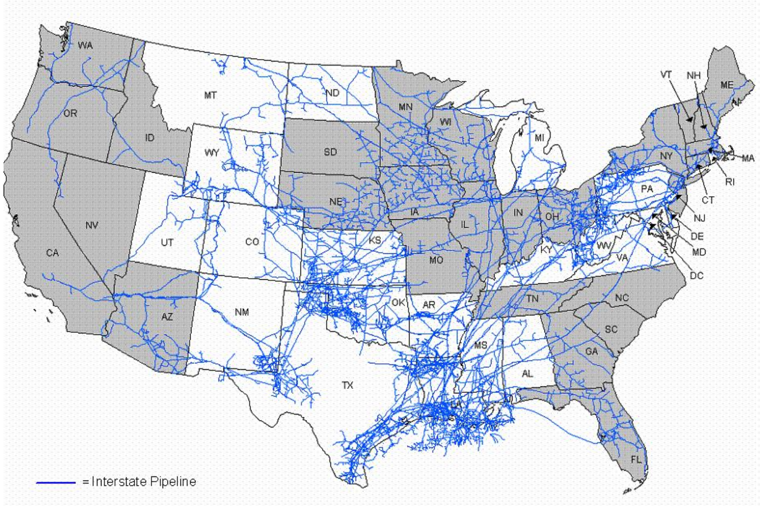


Figure 4.4: Natural gas pipelines and dependent states (Source: [79])

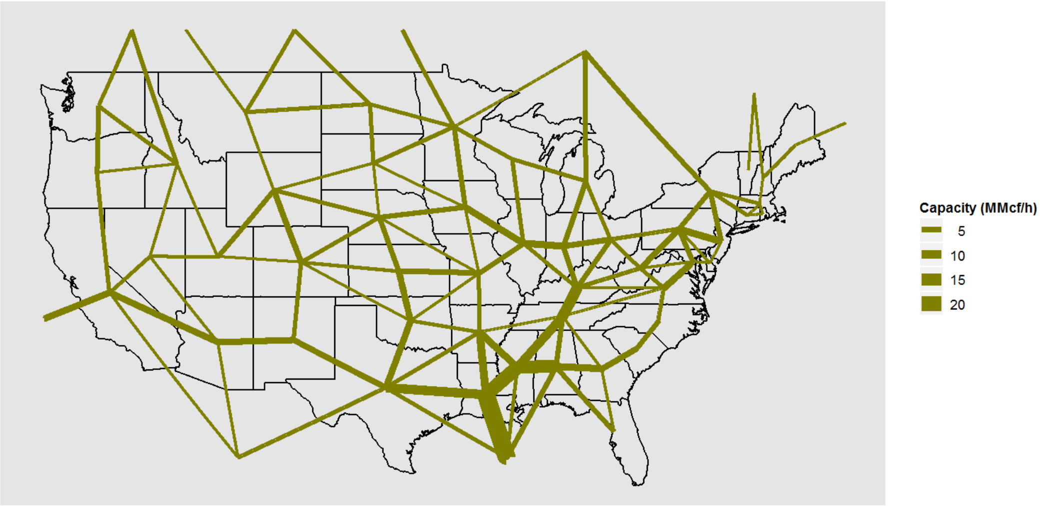


Figure 4.5: Modeled capacities for interstate and international natural gas pipelines

The last element modeled for the natural gas delivery subsystem is gas storage. Underground storage is an essential component of an efficient and reliable interstate natural gas transmission and distribution network. The size and profile of the transmission system often depends in part on the availability of storage as these are designed to accommodate the maximum peak-period requirements of consumption in the area. Then the transmission system is sized in terms of base load requirements [79].

There are mainly three types of facilities that are used to store natural gas underground: (1) depleted reservoirs in oil and/or gas fields, (2) aquifers, and (3) salt cavern formations. Figure 4.6 illustrates the location of those facilities in the continental United States.

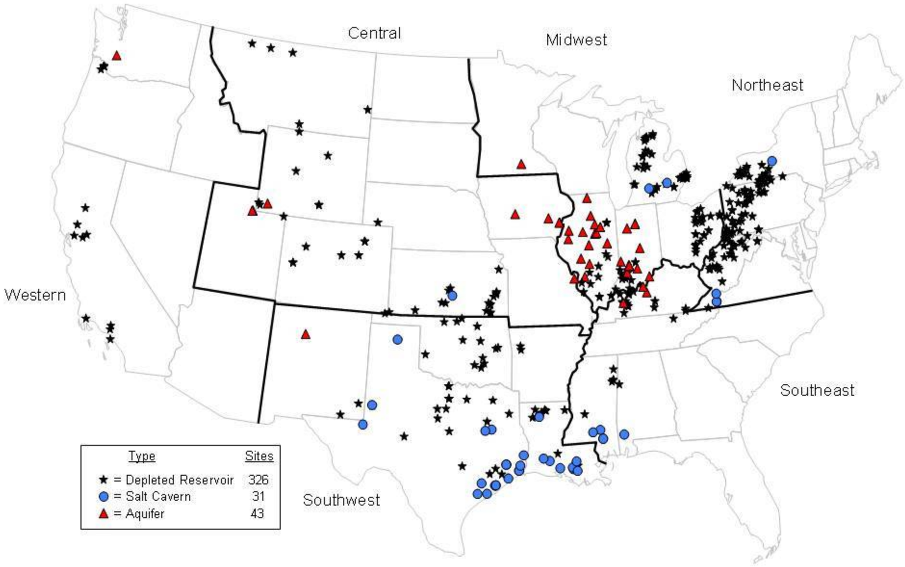


Figure 4.6: Natural gas storage sites (Source: [84])

The most common gas storage form is in depleted natural gas or oil fields that are close to consumption centers. Conversion of these fields from production is the simplest because of existing wells, gathering systems, and pipeline connections. Natural aquifers are predominant in the Midwestern United States and, although their geology is similar to depleted fields, their use in gas storage usually requires more base or cushion gas (amount necessary to provide enough pressure for withdrawals) and worse performance in terms of withdrawals and injections. On the other hand, salt caverns provide very high withdrawal and injection rates relative to their working gas capacity and base gas requirements are relatively low. The large majority of salt cavern storage facilities have been developed in salt dome formations located in the Gulf Coast states [84].

Natural gas storage is implemented in NETPLAN following the scheme laid out in Fig. 4.7. Each transshipment node is connected to a single storage node with two different arcs. These represent injections and withdrawals from the underground storage facilities. Then, storage nodes in the same geographical location and in consecutive time steps are connected through an arc to represent the natural gas that is carried over between time periods.

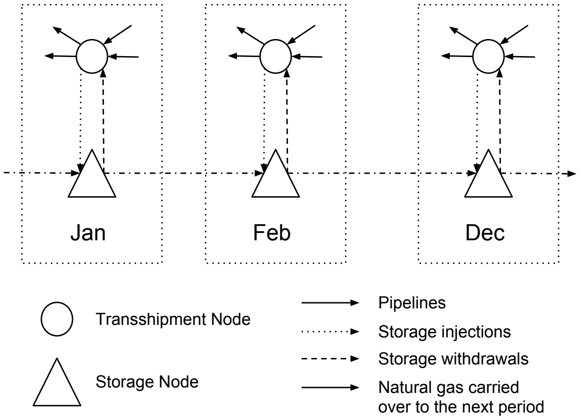


Figure 4.7: Natural gas storage nodes and arcs

Storage statistics are compiled and available for individual facilities in Form EIA-191A, “Field Level Underground Natural Gas Storage Data” [85]. From this data, the official monthly storage capacities and base storage requirements are obtained [86] and are here summarized in Table 4.8. Maximum injection rates into the storage nodes are also identified and all these quantities will serve as capacities for the different arcs in the storage system.

The last part of the natural gas subsystem corresponds to the demand and exports. In 2008, the total natural gas demand was 2,549 billion cubic feet. Roughly 20% was used by the power sector while the remaining 80% was destined to residential, commercial, industrial uses. Non-power demand is considered in NETPLAN as is applied to the appropriate natural gas transshipment nodes as fixed demand [87]. Monthly values are provided and the total sum by

Table 4.8: Natural gas storage capacities, base gas requirements and maximum injections by state

State	Capacity (Bcf)	Base gas (Bcf)	Injection (MMcf/h)
AL	26.90	3.64	66.67
AR	22.00	7.83	9.60
CA	510.70	212.35	286.16
CO	105.77	46.77	45.33
IL	977.68	680.24	255.67
IN	114.94	77.53	32.53
IA	284.75	197.05	50.00
KS	282.22	168.37	100.73
KY	220.36	123.32	79.51
LA	624.86	256.27	347.51
MD	64.00	46.68	16.67
MI	1062.34	396.63	724.63
MN	7.00	4.84	2.5
MI	190.55	85.86	222.81
MS	32.88	21.60	14.58
MT	374.20	178.50	12.91
NE	34.85	21.26	7.04
NM	80.00	28.27	14.58
NY	243.71	112.09	75.50
OH	572.48	346.22	203.21
OK	371.32	193.80	156.13
OR	29.41	11.18	21.46
PA	776.85	340.05	357.17
TN	1.20	0.34	0.83
TX	740.48	241.82	539.66
UT	129.48	64.91	21.96
VA	9.62	3.71	12.54
WA	39.29	21.48	47.92
VW	528.44	270.45	159.94
WY	111.17	64.97	12.57

state is represented in Fig. 4.8. For the 40 year simulation period a 1% annual load increase is assumed.

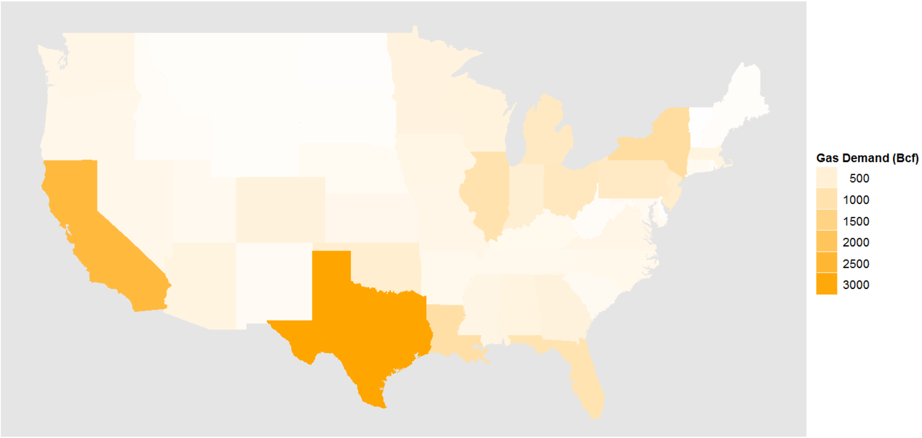


Figure 4.8: Total natural gas demand

To simulate natural gas demand from the power sector, transshipment nodes are connected to the nodes representing generators that use it as a fuel. To convert from volumetric quantities to energy units, an average heat content of 1.031 million Btu per thousand cubic feet is used [88].

In the flow diagram in Fig. 4.2 we can observe that exports to Mexico are a significant portion of the national natural gas subsystem. A demand node is assigned to simulate these exports as a demand and arcs are used to represent the pipelines that were shown in Fig. 4.5. Export data [89] is summarized in Table 4.9.

Table 4.9: Natural exports to Mexico (Bcf)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
39.88	37.45	31.02	28.11	25.14	30.36	30.18	34.59	26.51	28.45	25.82	27.89

4.1.3 Electric subsystem

This section is a detailed description of the electric subsystem for the United States that is modeled in NETPLAN. The analysis is done using monthly time step, and the area of study is divided using the supply regions used in NEMS [5] which are based on the regions defined by the North American Electric Reliability Corporation. These regions are presented in Fig. 4.9 and described in Table 4.10.

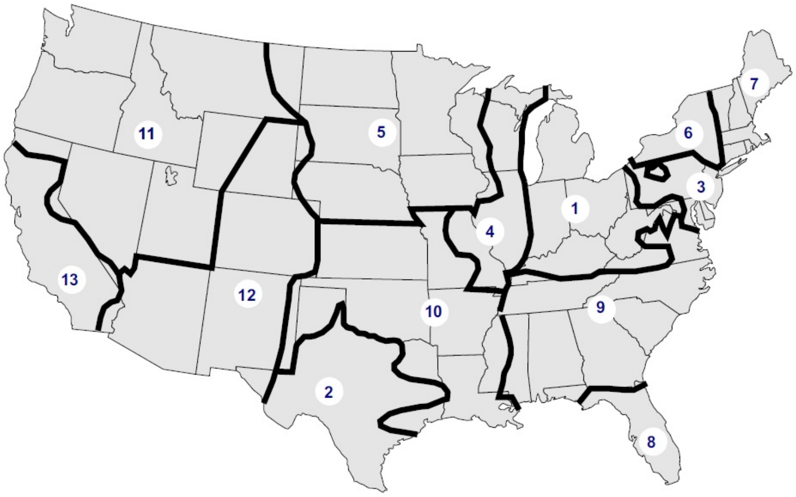


Figure 4.9: Electricity regions (Source: [5])

Table 4.10: Definition of electricity regions

Code	Acronym	Name
1	ECAR	East Central Area Reliability Coordination Agreement
2	ERCOT	Electric Reliability Council of Texas
3	MAAC	Mid-Atlantic Area Council
4	MAIN	Mid-America Interconnected Network
5	MAPP	Mid-Continent Area Power Pool
6	NY	New York Power Pool
7	NE	New England Power Pool
8	FL	Florida subregion of SERC
9	STV	Southeastern Electric Reliability Council (excluding Florida)
10	SPP	Southwest Power Pool
11	NWP	Northwest Pool
12	RA	Rocky Mountain and Arizona-New Mexico Power Areas
13	CNV	California-Southern Nevada Power Area

Electric demand data is in not available from any source that reported electric demand in a form that was compatible with the regions used in this study. Demand by region is calculated from data obtained at state level [90]. The resulting values along with the estimate peak load by region are summarized in Table 4.11.

In 2008, the United States had a net import of 33 TWh [91], less than 1% of the total load, which was slightly smaller than 3,600 TWh. For this reason, international trade is not considered in the model. If it was, the connection with Canada would be necessary since most of the imports come from that country.

Transmission lines between the electrical regions are represented. The maximum amount of power that can be transmitted across the region lines cannot be estimated by the sum of line capacities. Instead, the concept of area transfer capability (ATC) is introduced. ATC is defined as the ability of a transmission line to reliably transfer electric power in a network. Its definition depends not only on the physical limit of the individual lines but also on the topology of the network and its operating constraints.

Transfer capabilities are obtained from the model in NEMS [5] for the span of the simulation. The values for the first year are presented in Table 4.12. Transmission arcs are assumed

Table 4.11: Average and peak electricity demand

Region	Average demand (GW)	Peak Demand (GW)
ECAR	75.91	98.68
ERCOT	39.62	51.50
MAAC	25.74	33.46
MAIN	25.33	32.93
MAPP	23.01	29.91
NY	16.44	21.38
NE	14.05	18.26
FL	25.82	33.56
STV	70.62	91.81
SPP	32.73	42.55
NWP	28.25	36.73
RA	18.13	23.57
CNV	30.61	39.79

to be bidirectional so transfer capabilities are also valid for the directions opposite to those in the table. No upgrades to the transmission system are considered in this case.

Fifteen technologies are selected to represent the electrical generation in NETPLAN. Figure 4.13 presents the list of candidate generators, which is a balance between existing technologies at the beginning of the simulation and other technologies that could become significant. Operation and investment costs are provided in this table. Other operational parameters and emissions are summarized in Table 4.14. Although IGCC is one of the technologies, carbon capture and sequestration capabilities are not used as part of the model. Information for these parameters has been compiled and published in [92].

A capacity factor of 0.80 is assigned to all non-intermittent generation. Capacity factors generation for wind is assigned according to estimates for wind speeds [93] and varies from 0.1 to 0.4. Similarly, solar generation is provided with capacity factors between 0.1 and 0.25 that are estimated from solar power intensity values. The values used are presented in Table 4.15.

Generation investment is allowed for all the technologies, except for hydro. It is assumed that the great majority of hydro resources are already utilized at their fullest and existing generation will be maintained at the current level. New investment for the rest of technologies will

Table 4.12: Transfer capabilities for electric transmission

Origin	Destination	Transfer capability (GW)
ECAR	MAAC	7.50
	MAIN	3.69
	STV	6.89
ERCOT	SPP	0.97
MAAC	NY	3.42
	STV	4.47
MAIN	MAPP	1.62
	STV	5.16
	SPP	2.32
MAPP	SPP	1.81
	NWP	0.20
	RA	0.35
NY	NE	1.46
FL	STV	2.10
STV	SPP	1.26
SPP	RA	0.47
NWP	RA	2.59
	CNV	8.64
RA	CNV	6.88

have the objective of supplying the increase in new electricity demand and, more importantly, replace existing infrastructure that will gradually be retired.

A database resulting from Form EIA-860, "Annual Electric Generator Report," [94] contains a list of existing individual generators along with their capacity, type, location and years in operation. Aggregating this data by the defined electrical regions and utilizing the life times from Table 4.13 we can estimate how the initial capacity will be retired and, thus, know for each year what is the available generating capacity and what needs for new investments are. The evolution of this capacity is plotted in overall terms in Fig. 4.10.

Electric generation is, for the most part, a just-in-time system. Electricity being consumed at any particular moment is generated in that same instant. For that reason storage is not represented in this model.

Table 4.13: Electricity generation technologies and economic parameters

Techonolgy	Investment (MM$/MW)	Retirement (MM$/MW)	Fixed O&M ($/kW-yr)	Variable O&M ($/MWh)	Life time (years)
Pulverized coal	1.38	0.41	28.98	2.64	40
IGCC	2.06	0.62	46.82	2.26	40
NGCC	0.64	0.19	31.59	2.90	30
CT	0.42	0.13	14.05	2.79	30
Nuclear	2.43	0.73	48.81	4.09	60
Hydro	3.53	1.06	12.64	2.53	100
Inland Wind	1.21	0.36	18.68	7.17	25
Off-shore Wind	2.21	0.66	47.69	0.00	25
Oil	1.27	0.38	15.67	3.21	30
IPCC	2.55	0.76	117.58	16.12	30
Solar PV	5.28	1.58	11.51	0.00	30
Solar Thermal	3.50	1.05	50.33	3.68	30
Geothermal	2.06	0.62	152.69	0.00	50
OTEC	14.07	4.22	0.00	0.00	50
Tidal Power	4.74	1.42	38.10	0.00	50

Table 4.14: Operation and emissions by generation type

Technolgy	Heat rate $(\frac{MBtu}{kWh})$	Op. GHG $(\frac{lbCO_2e}{MMBtu})$	Op. NO_X $(\frac{g}{MMBTU})$	Op. SO_2 $(\frac{g}{MMBTU})$	Op. PM $(\frac{g}{MMBTU})$	Op. VOC $(\frac{g}{MMBTU})$
Pulverized coal	9.12	538.89	230.88	268.53	24.77	0.59
IGCC	8.43	507.10	166.01	294.84	19.50	0.59
NGCC	6.28	238.62	90.26	0.00	1.00	0.48
CT	8.84	325.73	97.98	0.00	1.50	0.48
Nuclear	10.32	0.00	0.00	0.00	0.00	0.00
Oil	9.36	473.68	2703.41	2408.58	49.90	1.04
Solar Thermal	15.26	0.00	0.00	0.00	0.00	0.00
Geo Thermal	31.69	72.43	0.00	9.03	0.00	0.00

Table 4.15: Capacity factors for wind and solar generation by region

Region	Wind	Solar
ECAR	0.3	0.15
ERCOT	0.4	0.2
MAAC	0.3	0.15
MAIN	0.5	0.15
MAPP	0.5	0.15
NY	0.3	0.15
NE	0.3	0.15
FL	0.3	0.22
STV	0.1	0.2
SPP	0.4	0.2
NWP	0.4	0.1
RA	0.2	0.25
CNV	0.3	0.22

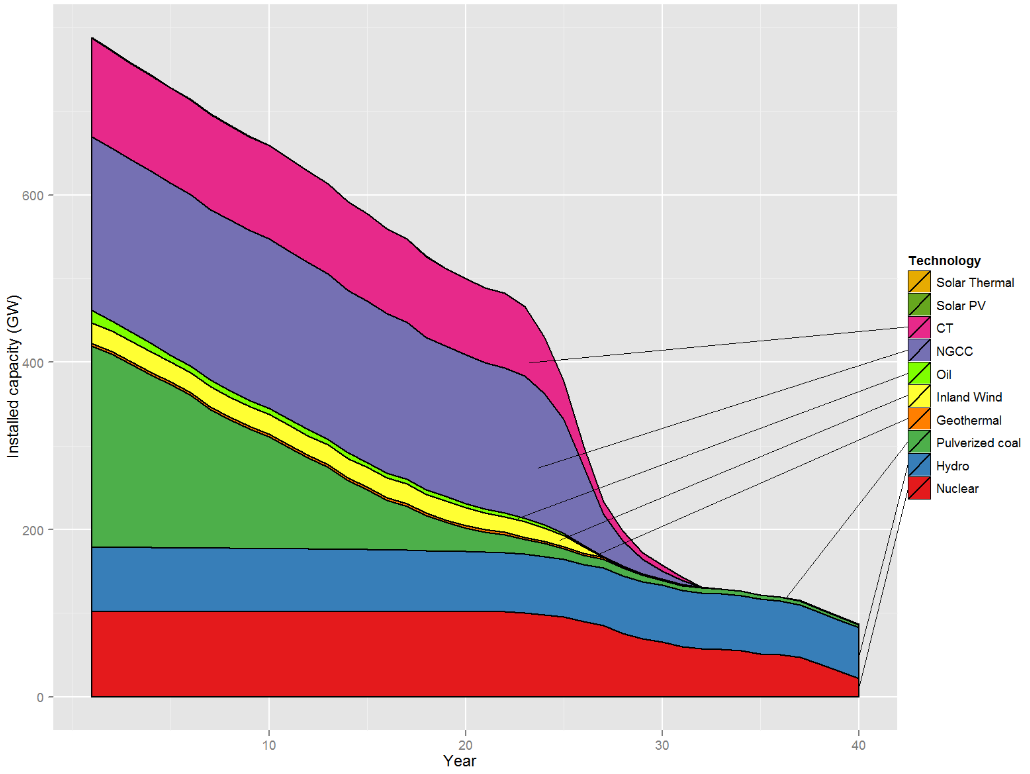


Figure 4.10: Evolution of initial electric capacity by technology

4.1.4 Petroleum subsystem

The petroleum supply network is a complex but vital part of the national economy. Not only is the source for the vast majority of transportation fuels, but the list of derivate products form part of everyday life. Unlike other energy subsystems, most of the petroleum consumed in the United States is imported, especially from countries belonging to the OPEC (Algeria, Angola, Ecuador, Iraq, Kuwait, Libya, Nigeria, Qatar, Saudi Arabia, United Arab Emirates and Venezuela) and others such as Canada, Mexico, Russia or the Virgin Islands [95]. As it can be seen in Fig. 4.11, the tendency over the last two decades has been to shift from national production to imports, although the trend seems to be interrupted in the most recent years [96].

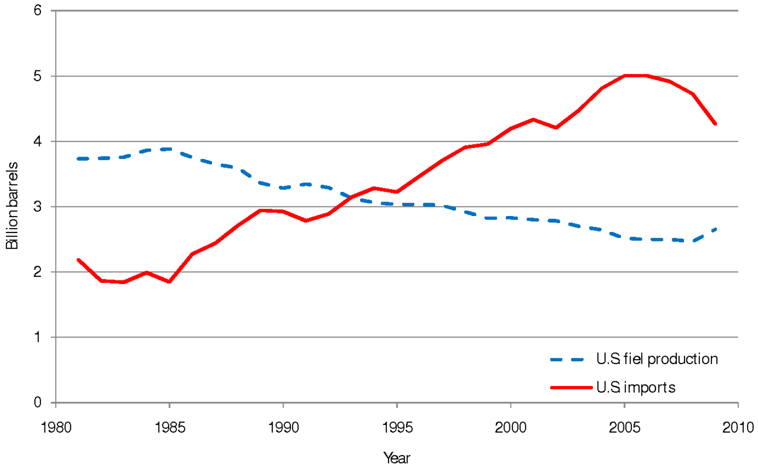


Figure 4.11: Historical data for crude oil production and imports (Source: [96])

Historically, crude oil have been rising continuously except for the financial crisis that started in 2007 as it can be seen in Fig. 4.12. In spite of the slow recovery of the economy, crude oil prices are rising again at a high pace [97]. Gas and diesel prices follow the trend set by crude oil [98] since this contributes between 50 to 60% of the price [97, 99].

For this numerical example, the petroleum network is simplified and it is only considered as a single node connected to an unlimited supply. A price of $4 per gallon is applied o the gasoline supply (used by passenger vehicles) and $3.8 per gallon of diesel (used by trucks and trains).

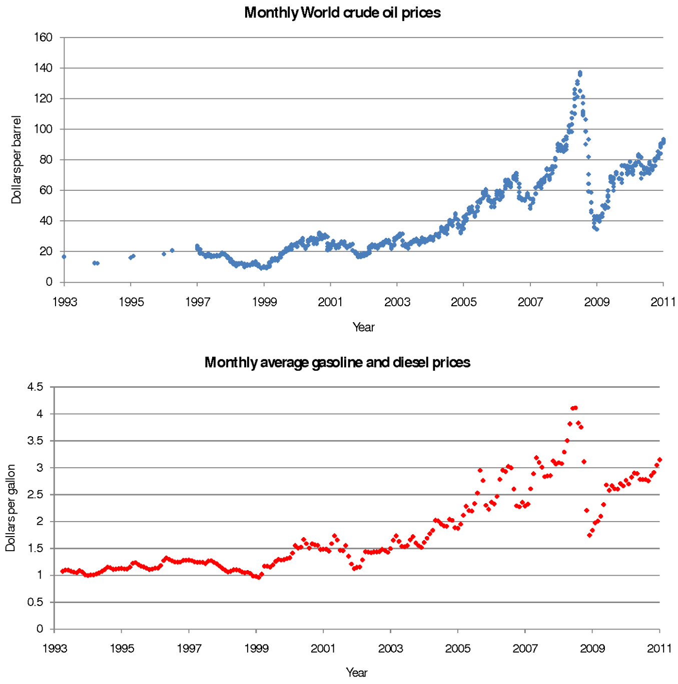


Figure 4.12: Historical data for crude oil and gasoline prices (Sources: [97, 99])

A more granular representation could be implemented if necessary, although publicly available data necessary is scarce when representing certain portions of the model. For instance, most of the data is available according to the five Petroleum Area Defense Districts (PADDs), represented in Fig. 4.13. Production and imports are rather well documented but there is not an official source for individual or aggregated pipeline capacities. The petroleum market module in NEMS, utilizes a linear model representation of the system [100] so it could be ported into NETPLAN in the future.

One of the primary determinants of green-house emissions for commodity and passenger transportation is the amount of carbon in the fuel. Thus, CO_2 emissions for those modes of transportation that utilize fuels derived from petroleum will be accounted for in this subsystem. The standard values for carbon emissions used by the Environmental Protection Agency are 2,421 and 2,778 grams per gallon of gasoline and diesel, respectively [101].

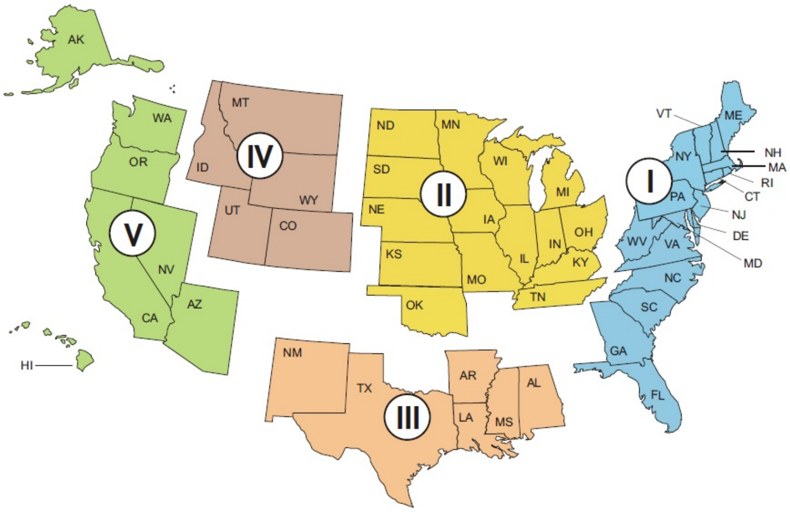


Figure 4.13: Petroleum area defense districts (Source: [100])

4.2 Transportation system

The transportation system in NETPLAN is based on the multicommodity flow formulation that was presented in Section 2.2. In this example, five commodities are used to represent the movement of goods around the country. Four of them correspond to the four types of coal that were introduced in Table 4.2. The fifth one is used to summarize the rest of commodities that do no have a direct connection to the energy system. A yearly time step is selected for the transportation system, and a node is used per state.

The projected transportation of commodities for the next 30 years is reported by the Bureau of Transportation Statistics through the "Commodity Flow Survey" [36]. The most significant commodities included in this document are base metals, basic chemicals, cereal grains, coal, gasoline, gravel, minerals, foodstuffs and wood products. These commodities are transported by rail, truck, pipeline and waterways.

Each of the different entries in the database of commodities consists on a pair of origin and destination states (could be the same or different), the commodity being shipped, the mode of transportation and a projection of annual shipments every 5 years for the next three decades. This set of data is, at first, incompatible with the way transportation demands are handled in NETPLAN. In the introduction and formulation it was noted that one of the inputs of the system was the definition of transportation loads for adjacent states. No route optimization is performed for commodities that are not energy commodities. In this numerical example, coal is the only energy variable. As it was previously mentioned, the use of route optimization for transportation commodities increases the size of the optimization model as the number of commodities and pairs of origin-destination states increases. Predetermining the routes reduces the size of the model, specially since it allows us to lump all not energy commodities into one.

A series of steps are necessary to transform the data from the "Commodity Flow Survey" into usable demand parameters for NETPLAN. The process is detailed below:

- With R and the *XML* package [102], a collection of routes with all the possible combination of states as origin and destination is obtained by using Google Maps API [103],
- Each entry in the "Commodity Flow Survey" data frame is divided into the segments according to the routes determined in the previous stage,
- For each pair of adjacent states the projected transportation demand is found by summing across the commodities and modes of transportation included in the study, i.e., commodities other than coal and for truck and rail only,
- Through interpolation and extrapolation, the demand is found for the years for which there is no data. For the projection beyond the end of the data it is assumed that the average increase between years 20 to 30 is maintained.

The result is a collection of demands that will be enforced on the transportation system for each interstate arc. Coal transportation will contribute to the rest of the transportation demand. Figures 4.14 and 4.15 represent the geographical distribution of this load for the first and last year, respectively.

The selected transportation infrastructures are railways and highways. For mode, diesel rail and truck already exist in the system and electric rail is considered as an alternative. To decrease the number of calculations each arc is endowed with a single distance. To determine

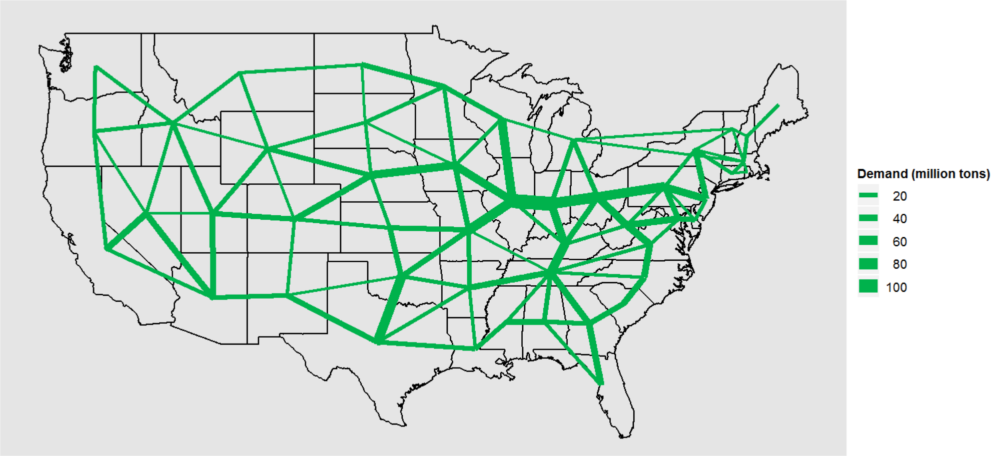


Figure 4.14: Transportation demand by arc for the first year

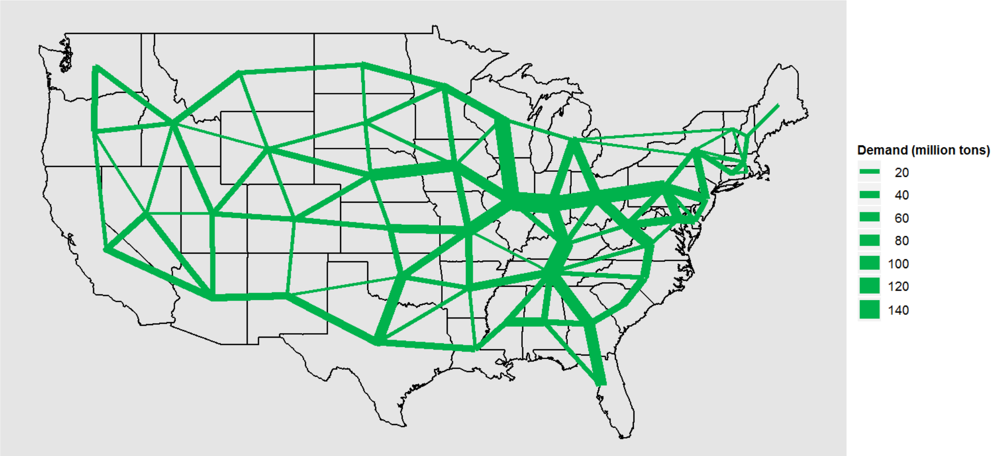


Figure 4.15: Transportation demand by arc for the last year

these distances, the weighted population is determined for each state. Then the distance is determined as the average of the distances by rail [104] and highway [105].

Existing capacity for rail and highway capacity is not easy to calculate. One should analyze the number of rail lines and highway lanes between adjacent states and estimate the capacity based on the maximum load that each arc could support with acceptable level of service. Alternatively, initial capacity is determined in this case based on the the initial transportation demand in the system. This is calculated taking into account the modes of transportation that we are modeling and account all commodities, including coal. The results are plotted in Figs. 4.16 and 4.17 for trucks and trains, respectively. Investment costs are determined from [52, 106].

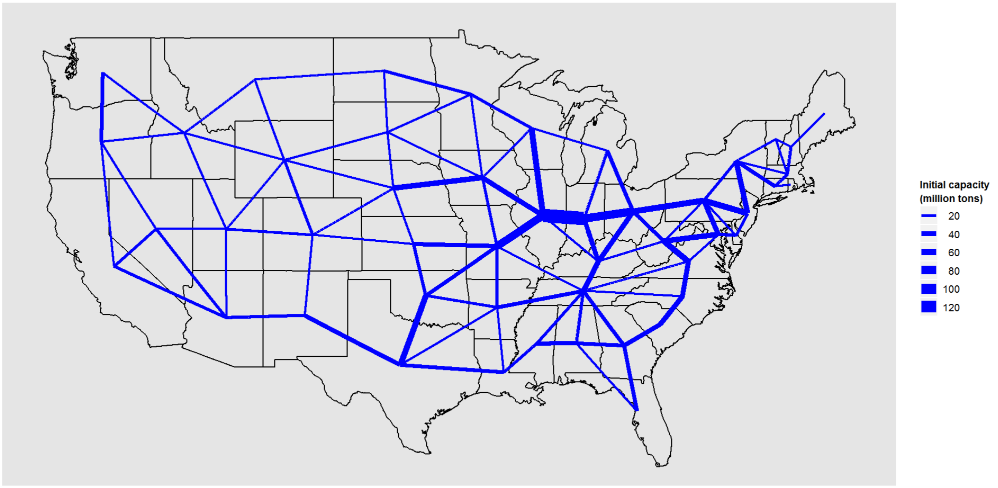


Figure 4.16: Truck capacity by arc

Fuel requirements are also present in Fig. 4.17. To obtain these, the characteristics of the terrain around the country are taken into account to estimate a proportion of plains, hills and mountains. With the use of fuel data from [107] the parameters shown in the picture are determined. The average fuel demand is 2.67 gallons for every thousand ton miles. The equivalent energy content is used for electric rail. This is a conservative estimate since electrified rail is more efficient than diesel rail in terms of tank-to-wheels performance. Electric motors are more efficient than internal combustion engines. To determine fuel consumption for truck

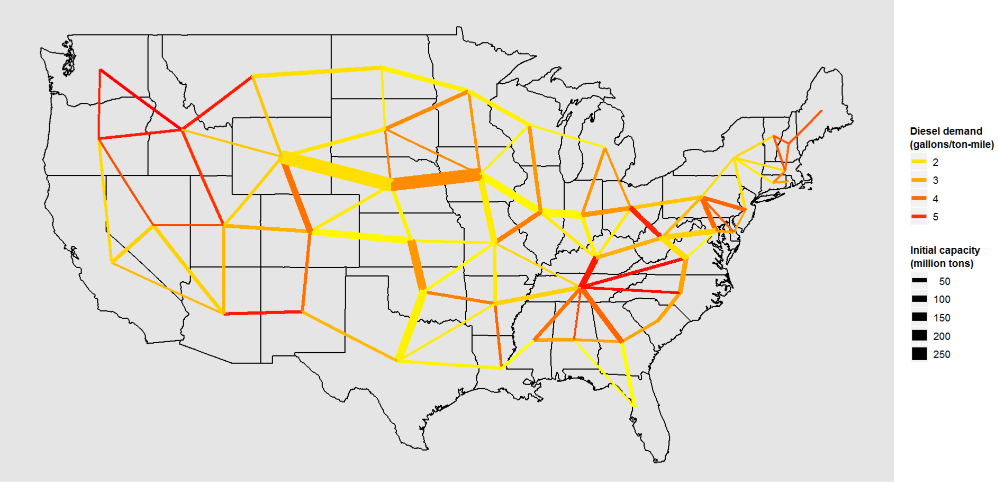


Figure 4.17: Initial rail capacity and diesel demand

by arc, the values for train are scaled taking into account that the average fuel consumption for truck is 7.57 gallons per thousand ton miles [108].

CO_2 emissions for all modes of transportation were accounted for both in the petroleum system and in the electric system. With these assumptions, green-house emissions are linearly related to the amount of fuel used.

4.3 Passenger vehicular transportation

As mentioned in Section 2.3, there are a multiplicity of factors that affect individual decisions in terms of what modes of transportation are to be selected. Operational and investment costs alone are not enough to determine these decisions, unlike the transportation of commodities. Further improvements in this field would be beneficial for the development of NETPLAN.

In spite of these limitations, a very simple model is implemented to characterize passenger vehicle transportation. It is assumed that we can predict the sales of new vehicles for the duration of the simulation. Two choices are available to cover the demand for new vehicles: a conventional gasoline and a plug-in hybrid electric vehicle, which has a 20-mile range on electric mode and then it switches to hybrid gasoline mode. The decision to purchase either vehicle will come from the optimization of the cost within the NETPLAN mode, which balances between purchase (investment cost) and fuels consumed during the lifetime of the vehicle (operational cost).

The demand of new vehicles is studied for each one of the thirteen electric regions presented in Section 4.1.3. According to the "Transportation Energy Data Book" [109] there were 205,239 million vehicles in the United States in 2008. These are assumed to be distributed among the 13 regions in proportion to the electricity demand. From [109] it can be derived that the average lifetime for passenger vehicles is 12 years so the existing number of vehicles is set to gradually phase out in that amount of time. The total number of vehicles is set to increase by 1% annually and, thus, we can calculate the number of vehicle sales for the duration of the simulation.

The gasoline model is assumed to have an average performance of 23 miles per gallon [110] and a price of $22,651 [111]. On the other hand, the hybrid vehicle performs at 40 miles per gallon during its gasoline operation and is estimated to cost $38,935 [112]. A methodology to calculate the energy necessary to power the all-electric mode in the PHEV is presented in [113]. According to it, the selected hybrid with uncontrolled charging at home would consume power according to Fig. 4.18.

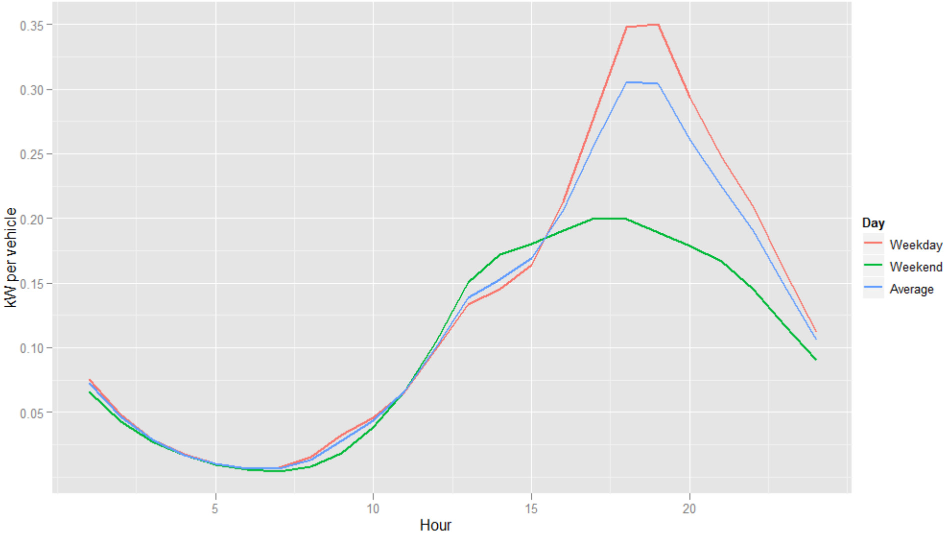


Figure 4.18: Power consumption for PHEV (Source: [113])

From Fig. 4.18 we calculate that each month a single PHEV would consume 87.1 kWh. This additional load is added to the load in its corresponding electric area. Similarly, assuming that each vehicle travels an average of 11,882 miles per year [109] and with the assumed mileages, gasoline consumption for regular and hybrid vehicles would be 516.6 and 297.1 gallons per vehicle per year, respectively. The gasoline demand is added to the node that represents gasoline demand.

As it was the case with freight transport, greenhouse emissions related to gasoline consumption were accounted for in the representation of the petroleum network as well as the electricity production subsystem.

This representation of the interactions between passenger transportation and energy demand is simple but at the same time it shows the potential importance of a fully developed model, which would include more choices for personal vehicles, more technologies to be used as fuels, interstate travel and other modes of transportation like rail or air. All this would be accompanied with the appropriate modeling framework to better represent the decision making of the average traveler.

4.4 Metrics

The selected objective for this implementation of NETPLAN are the overall cost, greenhouse gas emissions and resiliency.

The cost term takes into account operation and investment in the system. The cost parameters described in the previous sections are assumed to remain constant throughout the span of the simulation. An inflation rate of 2% is applied across the board. Discount rate is set to 2%. Investment cost will come from enhancing the generation, transportation and vehicular portfolios.

CO_2 emissions are tracked for electric generators, commodity transportation and passenger fleets. The emission index introduced in Section 2.6.2 is used. The upper and lower references for emissions are set by a 1% annual increase and decrease of the first year emissions.

Resiliency will be tested by applying four separate events to the system. Initially, an event was defined as the loss of a type of generation but these resulted in a large number of calculations that were not very significant. With these information, however, it was possible to group generation technologies in three groups that, when they are subject to failure, have similar effect on the increase in operational cost. The fourth event was selected to complement the other three as it is a failure in the natural gas system and not the electric system. The effect of this event could extend to the electric system through gas turbines and combine cycle units. The events used are the following:

- Loss of 70% of nuclear energy power
- Loss of 70% of coal generation (pulverized coal and IGCC)
- Loss of 70% of installed capacity for renewables (wind, off-shore wind, solar photovoltaic and thermal, geothermal, ocean tidal and ocean thermal)
- Loss of 50% of natural gas production and pipelines in the Gulf of Mexico

Each event includes the loss of capacity at year 20 and at year 30 and gradual recuperation during the year after each of them. This is equivalent to applying 8 separate events, but reduces the required number of computations and plots.

4.5 Model summary

This section summarizes the scope and size of the model to conclude the chapter. Table 4.16 presents the number of arcs and nodes utilized for each one of the different systems.

Table 4.16: Nodes and arcs by system

System	Type	Size
Coal	Production	24 nodes
	Demand	49 nodes
Natural gas	Production	25 nodes
	Demand	50 nodes
	Pipelines	108 arcs
	Import pipelines	9 arcs
	Storage	30 nodes
Electricity	Generation	203 arcs
	Demand	13 nodes
	Generation	203 arcs
	Transmissions	19 arcs
	Import transmission	8 arcs
Petroleum	Gasoline	13 nodes
	Diesel	13 nodes
Freight	Transportation	95 arcs
	Coal demand	49 nodes
Passenger	Vehicles	13 arcs

CHAPTER 5. STUDY RESULTS

This chapter describes the results obtained from NETPLAN with the modeling assumptions described in the previous chapters. First, the minimum cost solution is analyzed and then a sensitivity analysis is performed. Then, the solutions for the multiobjective optimizer are presented and the different solutions from the Pareto front are compared in terms of differences and similarities.

5.1 Introduction

The implementation of the minimum cost solution results is a linear program with 748,394 variables and 472,920 constraints. The problem is solved in a Iowa State University server with a 1.6 GHz processor and 24 GB of RAM memory. C++ libraries for ILOG CPLEX 12.2 [61] are used to solve the linear programs.

Solution times for the full problem without decomposition average 17 minutes. Applying the decomposition method described in Section 3.5 does not produce a speed-up in solution time and it is, in fact, several times slower. In this case, CPLEX includes an intelligent presolver that further increases the speed of the full case.

However, the biggest reason for this difference is that it is not possible to use a “hot start” (use a previous solution to initialize the current problem) for the subproblems. This is due to the fact that version 12.2 of CPLEX has a known-problem with memory management [114] that caused the most efficient implementation of Benders decomposition to fail and the cuts had to be implemented in a different way.

In smaller scale models for which memory was not an issue the solution times with decomposition methods were, on average, 50% smaller than those for the full size problem. Further speed up could be achieved by solving the multiple year subproblems in a parallel fashion, although this effort is not part of this work.

One measurable advantage of using Benders decomposition is memory usage. When the full model is loaded in memory it occupies 10.2 GB as opposed to 5.3 GB for the series of master and subproblems. This could prove critical for much larger implementations of NETPLAN if resources are limited.

5.2 Minimum cost solution

In this section the portfolio for the minimum cost case with the values presented in Chapter 4 will be presented. This case will later be referred as the base case.

The minimum cost solution is solved with a total cost equal to 52.3 trillion dollars. CO_2 emissions equal 116.5 billion tons, and the calculated emission index is 0.304. This index, introduced in Section 2.6.2, indicates that the emissions are closer to the lower band of 1% annual decrease than to the "bussiness as usual" case.

Resiliency calculations are captured in Table 5.1 and show that the average operational cost increase for the energy and transportation system under the selected events is 36.56 billion dollars. All the events have a significant impact on the system with the loss of nuclear energy being the most severe. More details about these results will be given throughout this section.

Table 5.1: Impacts of events on operational cost

Event	Cost increase (billion $)
Loss 70% nuclear energy	78.43
Loss 70% coal generation	27.98
Loss 70% renewables	28.37
Loss 50% natural gas in the Gulf	11.47
Average (metric)	36.56

Table 5.2 includes some general statistics that have been collected for the year 2008 and the comparable simulated results. Observed data was obtained from [78, 115, 116, 117, 118] and all simulated results have the same order of magnitude.

Table 5.2: Actual versus simulation results

Value	2008 actual data	Simulation	Difference
Coal production (million short tons)	1,041	747	-28%
Natural gas production (Bcf)	15,618	9,370	-40%
Natural gas production and imports (Bcf)	19,602	16,377	-16%
Motor gasoline demand (billion gallons)	132.2	129.3	-2.2%
Train diesel demand (million gallons)	3886	5479	41%
Electric CO_2 emissions (million tons)	2605	1673	-36%
Transportation CO_2 emissions (million tons)	1968	1315	-33%

Coal and natural gas production are underestimated, mainly due to the fact that hydro production is overestimated, as shown in Fig. 5.1, where electric generation by type is compared to the actual values [119]. This problem indicates that the simulation of hydro facilities should further be enhanced to include the fact that full generation is not always possible due to the variability in available stored water through the seasons.

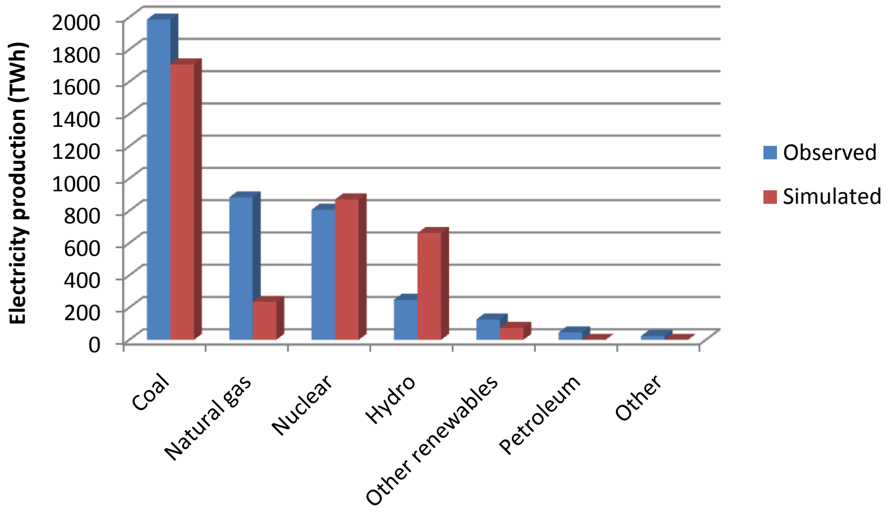


Figure 5.1: Electricity generation by type

National natural gas production is lower than observed since imports from Canada are overestimated. Import prices in Table 4.7 are more competitive than those in the major production states (Table 4.4), largely due to the fact that 2008 was a particularly active year for hurricanes in the Southern states [80], which drove prices up. Furthermore, in the model, natural gas imports are only constrained by pipeline capacities crossing the border so some extra work could be necessary to better represent the availability of natural gas from Canada and any constraints on its import.

The operational model for the energy system was tested in [120, 53], so with the appropriate changes to data input the results would be improved to better represent the observed data.

Gasoline demand was estimated with a simple model and resulted in a good approximation on the real consumption. Diesel demand from rail was overestimated by 40%. One of the factors that contributes to this difference could be the way interstate connections were defined between the centers of population of the states, which causes the number of total miles (and thus the number of ton-miles transported) to be larger than the actual mileage for rail.

Carbon dioxide emissions were underestimated for both electric generation and transportation. In the first case, this is largely due to the reduced use of coal and natural gas as fuels for power generation. For transportation, there are emitting sources that are not being accounted for in this model, such as intrastate transportation through trucks or passenger transit.

The remaining parts of this section will revolve around the presentation of the portfolio and evolution of the system for the simulation span. First, we consider the evolution of generation capacity and energy generated in Figs. 5.2 and 5.3, respectively. To better understand the evolution of capacity, Fig. 5.4 represents annual investments on power generation technologies.

The most interesting features in those figures is the strong investment that is performed in nuclear, geothermal and wind. Nuclear becomes the largest producer of electricity displacing a significant portion of coal generation. IGCC also becomes an important player in the generation mix, and pulverized units still remain but on a smaller scale. Geothermal and wind complete the major portion of energy generated, which is also complemented by small portions of combined cycle. Figure 5.3 shows that the electricity production increases in steps and is due to the way demand is defined.

There is a dip in capacity between years 25 to 28 in Fig. 5.2, which is consistent with a massive the retirement of existing natural gas combined cycle and combustion turbine units that can be appreciated in Fig. 4.10. This decrease is compensated by heavy investments

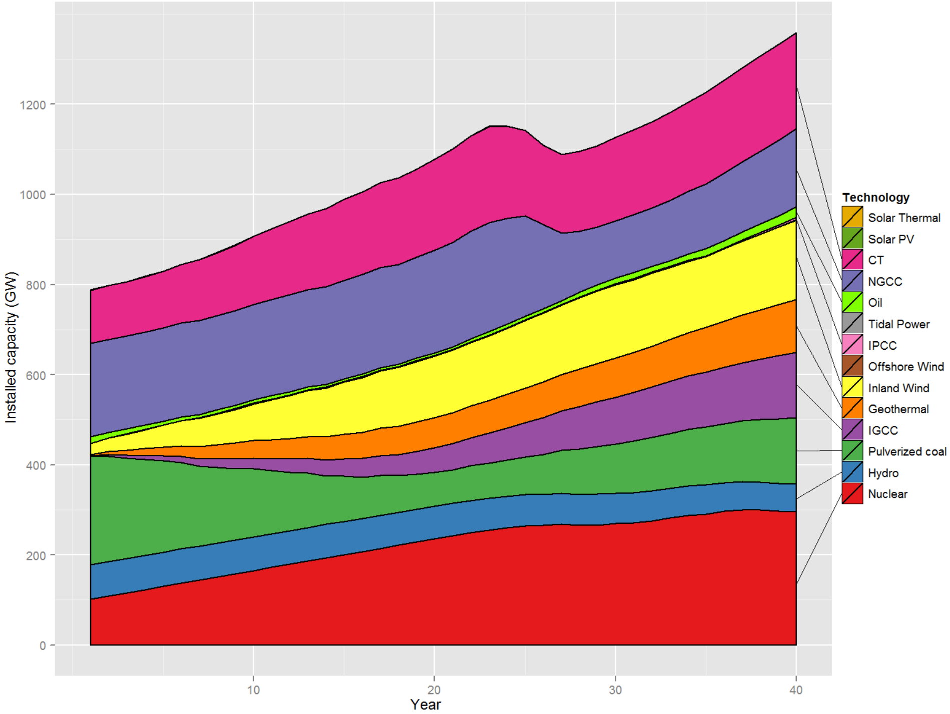


Figure 5.2: Total electricity capacity

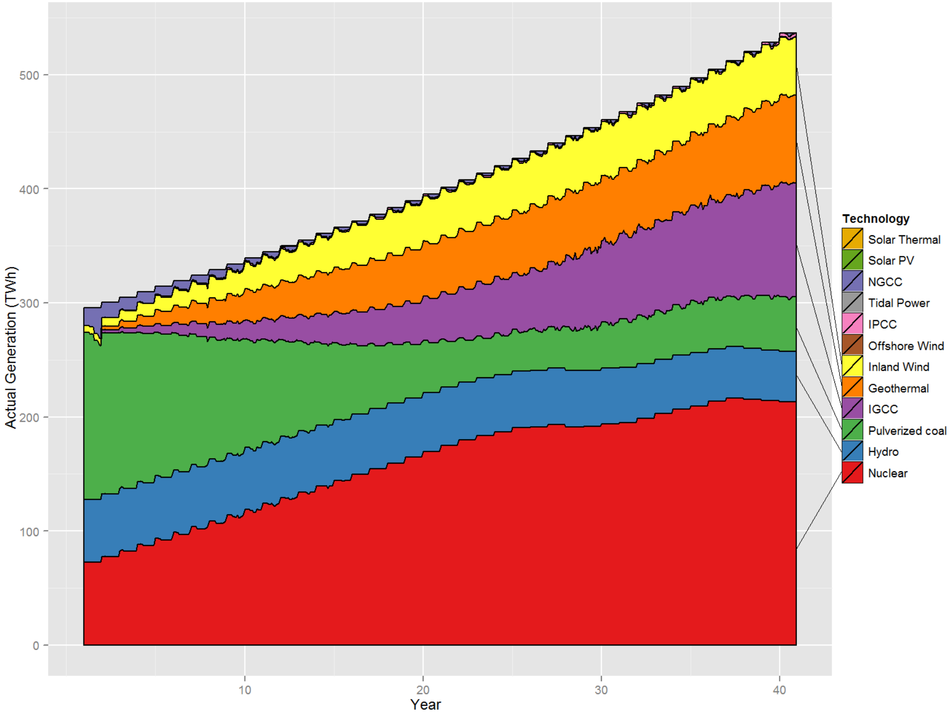


Figure 5.3: Energy generated

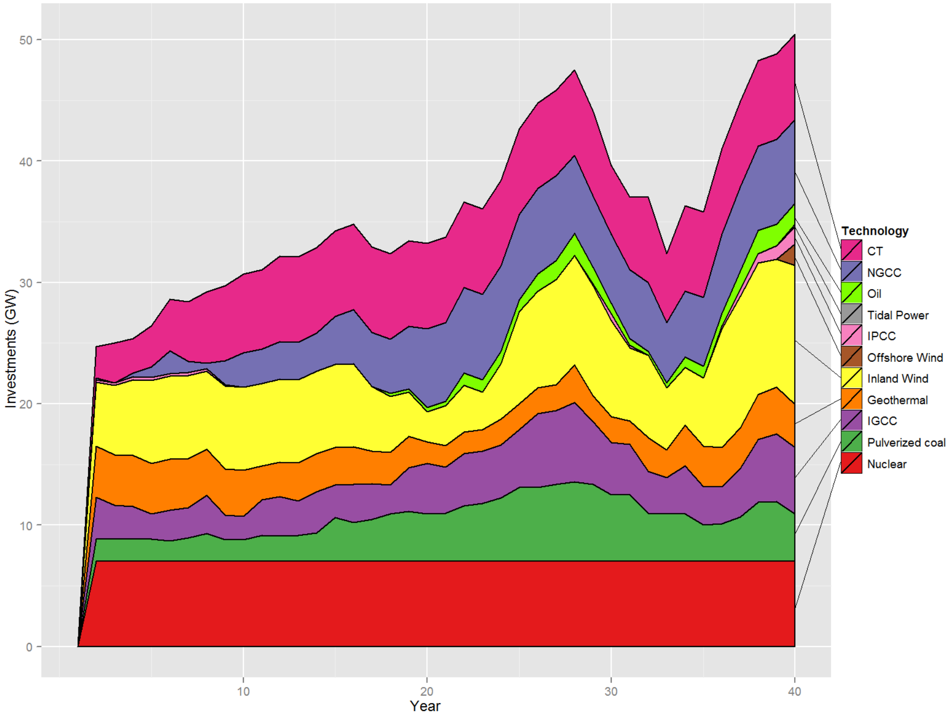


Figure 5.4: Electricity investment

(specially for wind) around year 28, according to Fig. 5.4.

It is interesting to notice that investments in natural gas combined cycle and combustion turbines are quite high even though they provide very little energy. This investment is driven by the necessity of covering the peak demand since natural gas generation are the cheapest technologies to invest on. Production costs are quite high and explains the lack of energy produced from these generation technologies as the system evolves. A more realistic use of these facilities would be possible if a refined model was used to represent the daily variations in electric load. In the present model, the use of monthly loads favors the use of base generation. NETPLAN has modeling capability to address this issue by segmenting the load duration curve, but doing so is computationally expensive and was therefore nor utilized in the work reported in this dissertation [53].

Investments by region show interesting trends (Fig. 5.5). Nuclear and natural gas generation present consistent levels of investment across areas. The investment on wind, geothermal, pulverized coal and IGCC are the main competitors for the remaining investment. The distribution is mainly directed by the efficiency assigned to wind, which was presented in Table 4.15. For this reason, ERCOT, MAIN, MAPP, Florida and SPP see big wind penetrations. This pattern is not so clear for geothermal energy and a revision of the data should take into account the higher potential of the Southwest regions in terms for this generation technology.

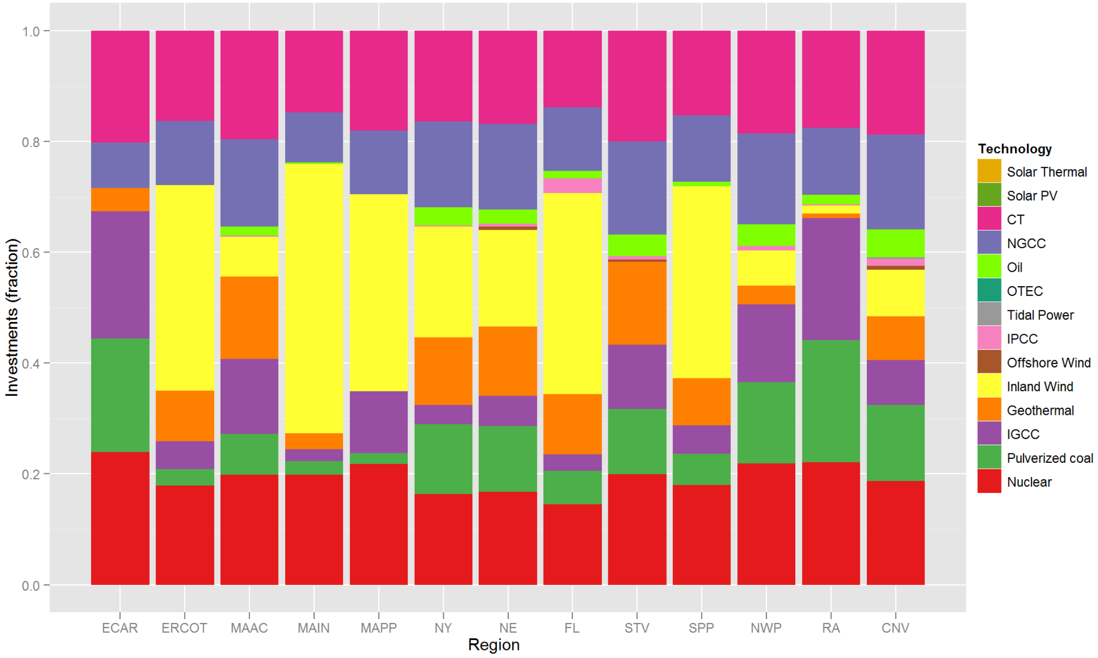


Figure 5.5: Electricity investment by region

The behavior in the electricity generation has a direct impact on the production of coal and natural gas. Coal demand for the 40 year period is pictured in Fig. 5.6 and we can see that the demand begins to decrease as other forms of generation (nuclear, wind, geothermal) start replacing pulverized coal units. As electricity demand increases in the second half of the simulation so does the investment in pulverized coal and, especially, IGCC facilities, which drive the demand of coal. Even though pulverized coal units are cheaper, IGCC are more efficient and that offsets the different between the two. It is also interesting to see that the origin for coal evolves over time and depends on coal quality but also on the location where it is produced. There is an increase between the first and the second year, which is facilitated by the investment on transportation as we will see later. This means that transportation capacity is not sufficient to distribute the necessary coal. The introduction of transportation by barge could help alleviate this problem.

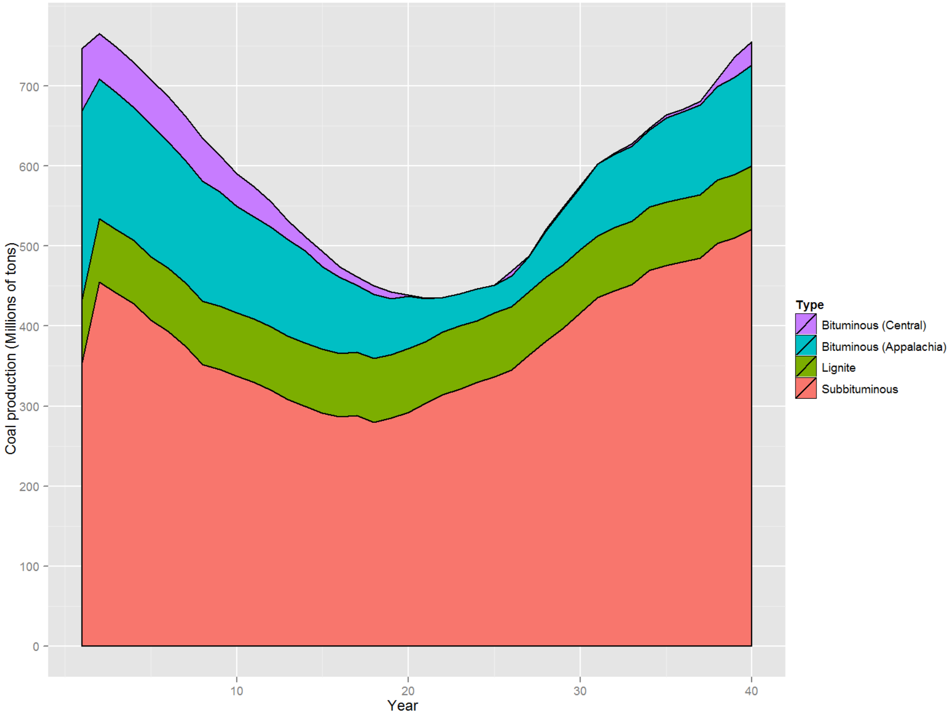


Figure 5.6: Coal production

Natural gas demand has a much simpler evolution as shown is Fig. 5.7. It begins with a small dip in demand corresponding to the shift from natural gas generation to coal. After that, natural gas demand is mainly driven by the non-electric demand, which has an annual increase of 1%. This figure also shows how the imports are always at the maximum level allowed by the pipelines across the border.

Electricity prices can also be obtained from the model via the dual variables at the electric transmission nodes. In general, the predictions are lower than observed retail prices, which average $67 per MWh in 2008 [121]. Predictions do not take into account taxes.

As in the case of optimal power flow models, electricity prices are determined by the incremental cost of the most expensive generation. To further illustrate the time variations shown in Fig. 5.8 the prices for two regions, New York and New England, are plotted along with the actual generation in Figs. 5.9 and 5.9, respectively. Both regions are connected by electric transmission and when the capacity of that line is not binding prices are the same for

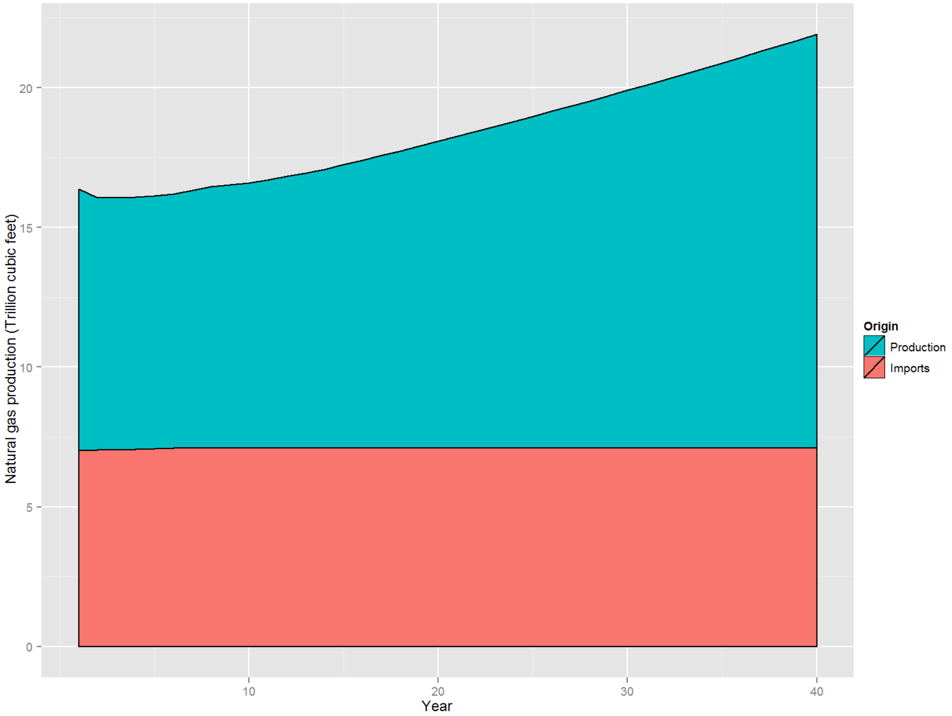


Figure 5.7: Origin of natural gas

both regions. Before year 8 the transmission line is at its limits and prices in New England are higher than in New York. Electricity prices in New England show a monthly jumps during the initial years, which are due to natural gas cost variation that fuels the combined cycle units.

As the previous graphs suggested, the electric sector i beings completely transformed along the simulation span and since the objective is minimization of cost it is logical that prices tend to decrease and stabilize. Figure 5.11 presents the average price for each decade by region. The price gradient in years 1 to 10 becomes less evident as the generation mix becomes more uniform across regions. Prices in the Northeast are set lower due to the abundance of energy from hydro generation (which is being overestimated in the model), while they are the highest in Florida and the Northeast, due to transmission limitations.

The passenger portfolio is dominated by gasoline vehicles for the entire simulation as seen in Fig. 5.12. Similarly, freight transportation is based entirely in truck and diesel rail. Figure 5.13 shows that electric transportation is not part of the solution. Later in the chapter it will be shown that this lack of electricity-powered transportation is due to the fact that gasoline

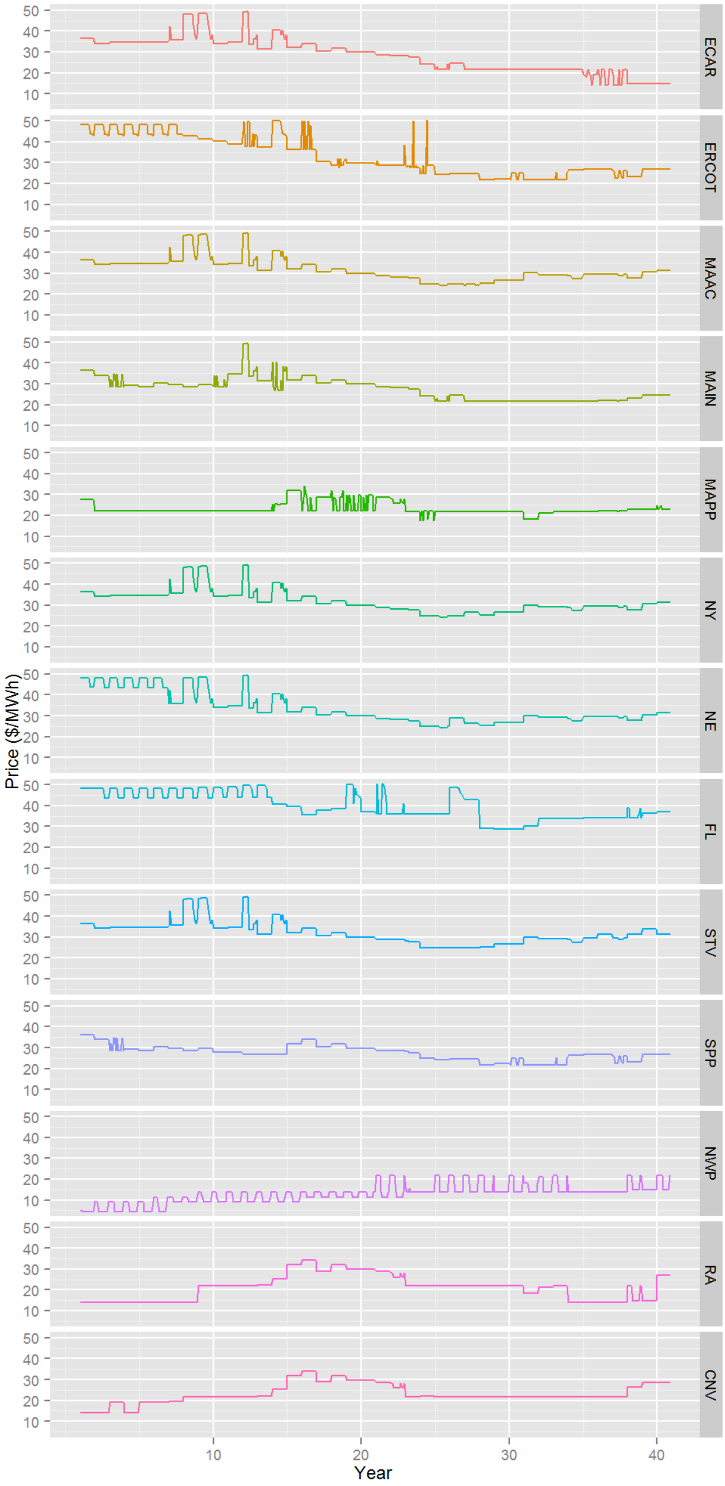


Figure 5.8: Electricity prices by region

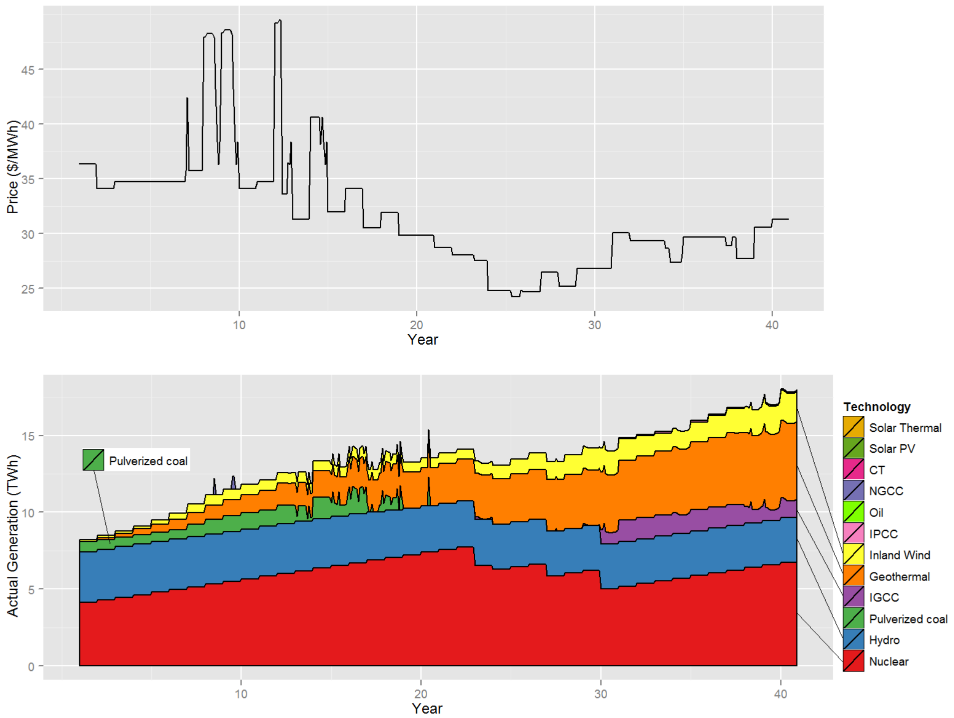


Figure 5.9: Prices and electricity generation for New York

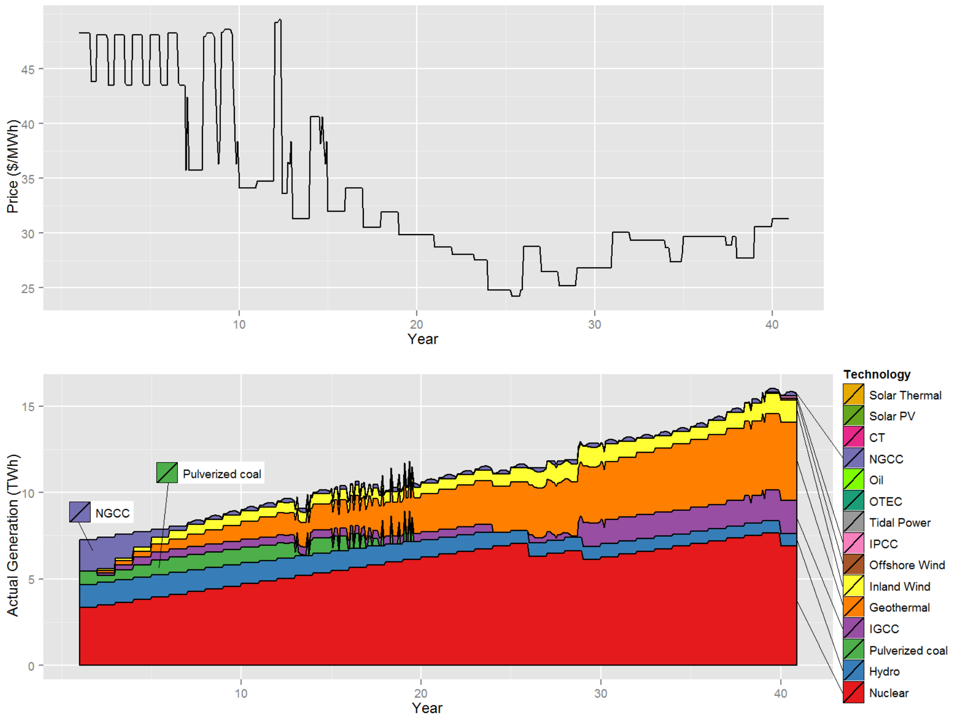
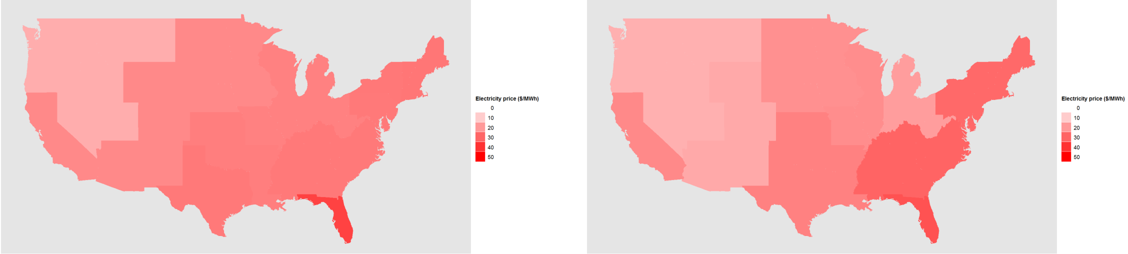


Figure 5.10: Prices and electricity generation for New England

(a) Years 1–10 (b) Years 11–20



(c) Years 21–30 (d) Years 31–40

Figure 5.11: Average electricity price by state

prices are comparably lower than those for electricity. Higher gasoline prices gradually shifts the investment towards electrified rail and hybrid portfolios.

Transportation investment presents some interesting features in Fig. 5.14. In the second year there is a spike in the investment, which is motivated by coal transportation as discussed earlier. This investment also displaces some of the transportation by truck, according to Fig. 5.13. The rest of the simulation shows that investment is done progressively until it suddenly drops and begins to rise again. This drop in investment is due to a drop in the global demand for transportation, which is determined by the input data.

Gasoline and diesel demand in Fig. 5.15 are consistent with the modes of transportation chose. Gasoline production is overwhelmingly larger than diesel for rail.

The evolution of carbon emissions in Fig. 5.16 relates directly to the operational aspects observed earlier. Rail produces a very small portion of the emissions, which are dominated by the power and passenger transportation sectors. CO_2 emissions due to passenger increases along with the use of vehicles while electric generation emissions are directly related to coal-

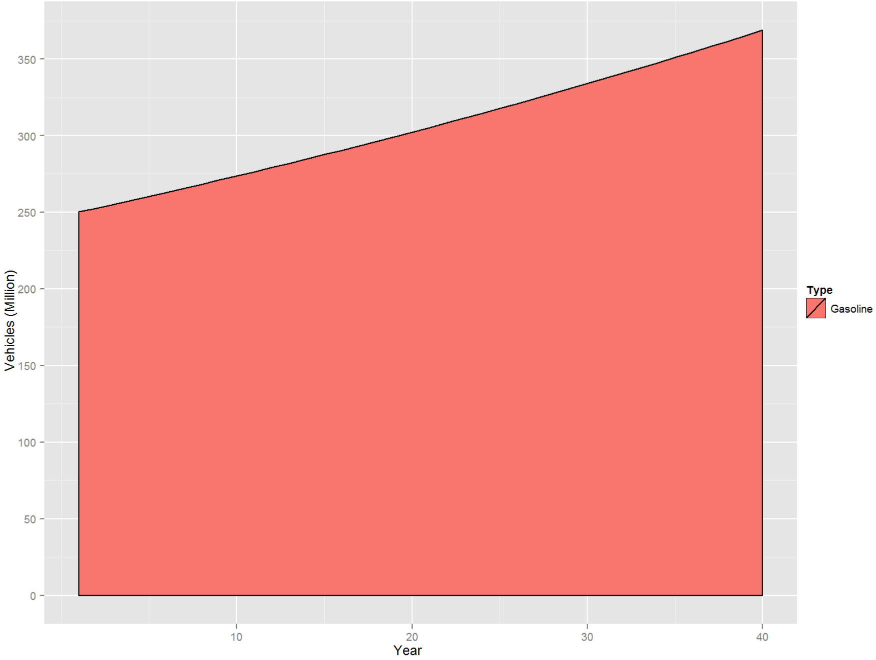


Figure 5.12: Passenger fleet portfolio

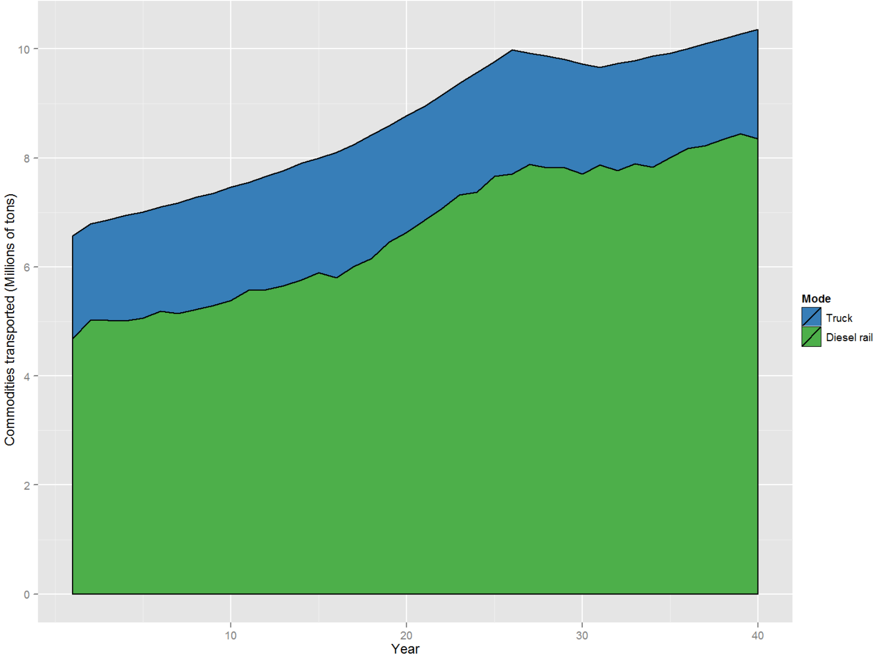


Figure 5.13: Transportation fleet usage

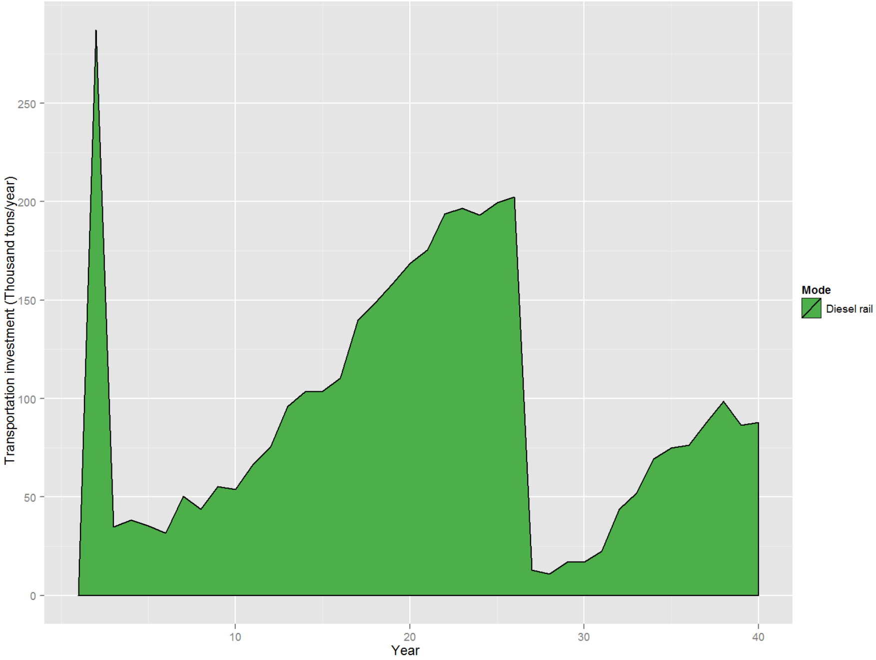


Figure 5.14: Transportation fleet investment

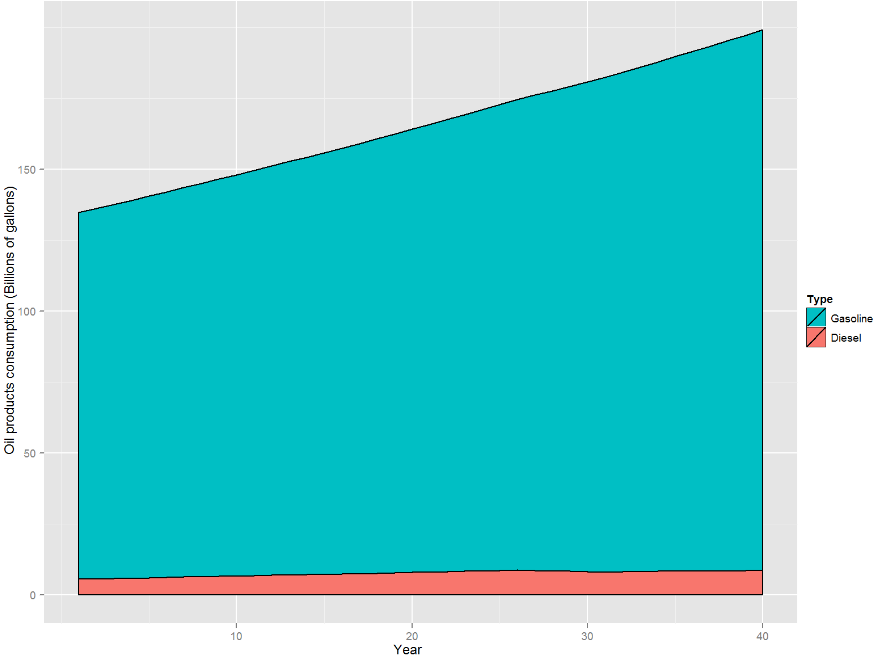


Figure 5.15: Gasoline and diesel demand

based technologies and, thus, with coal produced (Fig. 5.6).

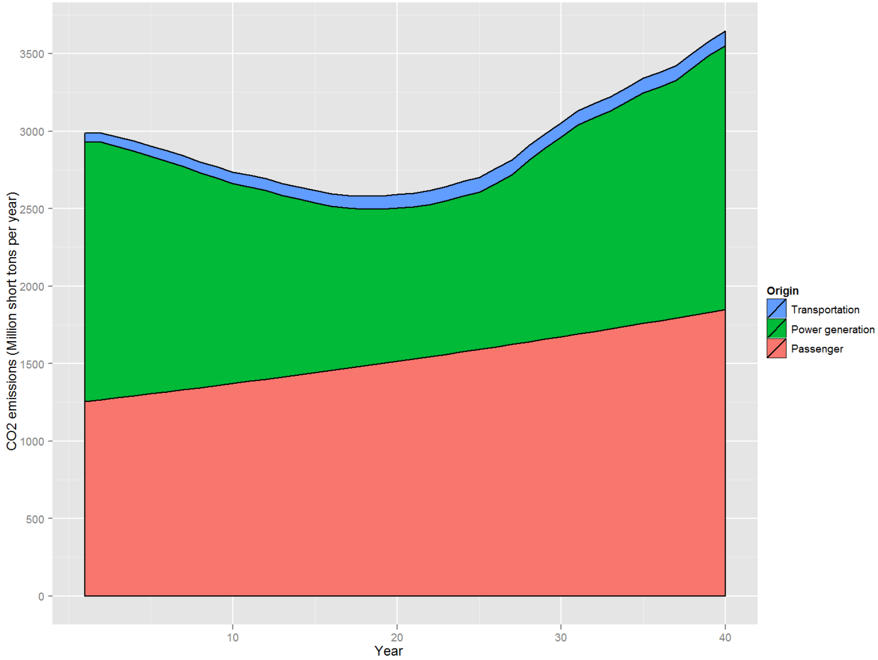


Figure 5.16: Evolution of CO_2 emissions

Resiliency results were introduced in Table 5.1 for the four events considered, that are:

- Loss of 70% of nuclear energy power
- Loss of 70% of coal generation (pulverized coal and IGCC)
- Loss of 70% of installed capacity for renewables (wind, off-shore wind, solar photovoltaic and thermal, geothermal, ocean tidal and ocean thermal)
- Loss of 50% of natural gas production and pipelines in the Gulf of Mexico

The following set of plots in Figs. 5.17–5.20 contain the nodal prices for each electric regions under the stressed conditions. As a reference the baseline prices without the contingency have been included using a thinner line. We can see that the change in prices is important for the first three events, but small changes are detected for the fourth one for the reasons mentioned before.

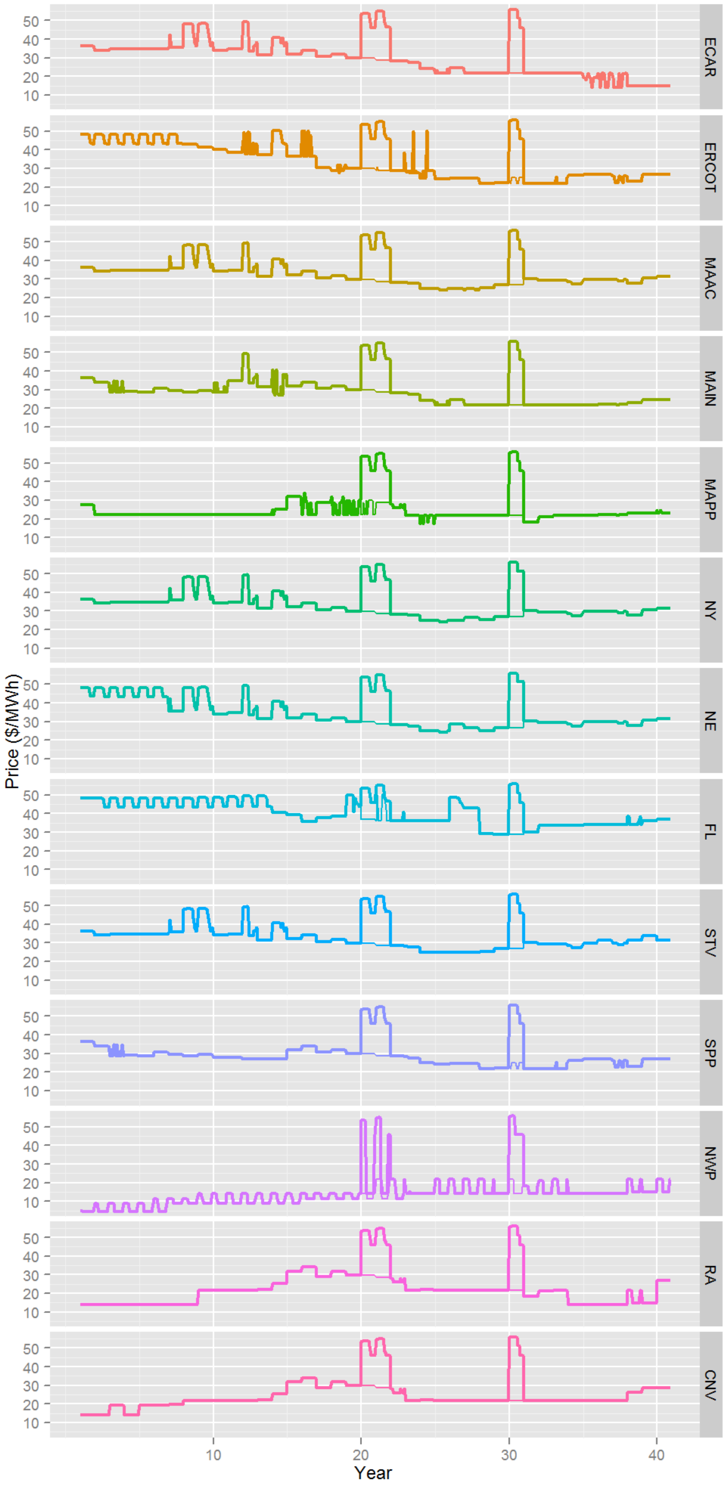


Figure 5.17: Electricity price for event 1: loss of nuclear energy

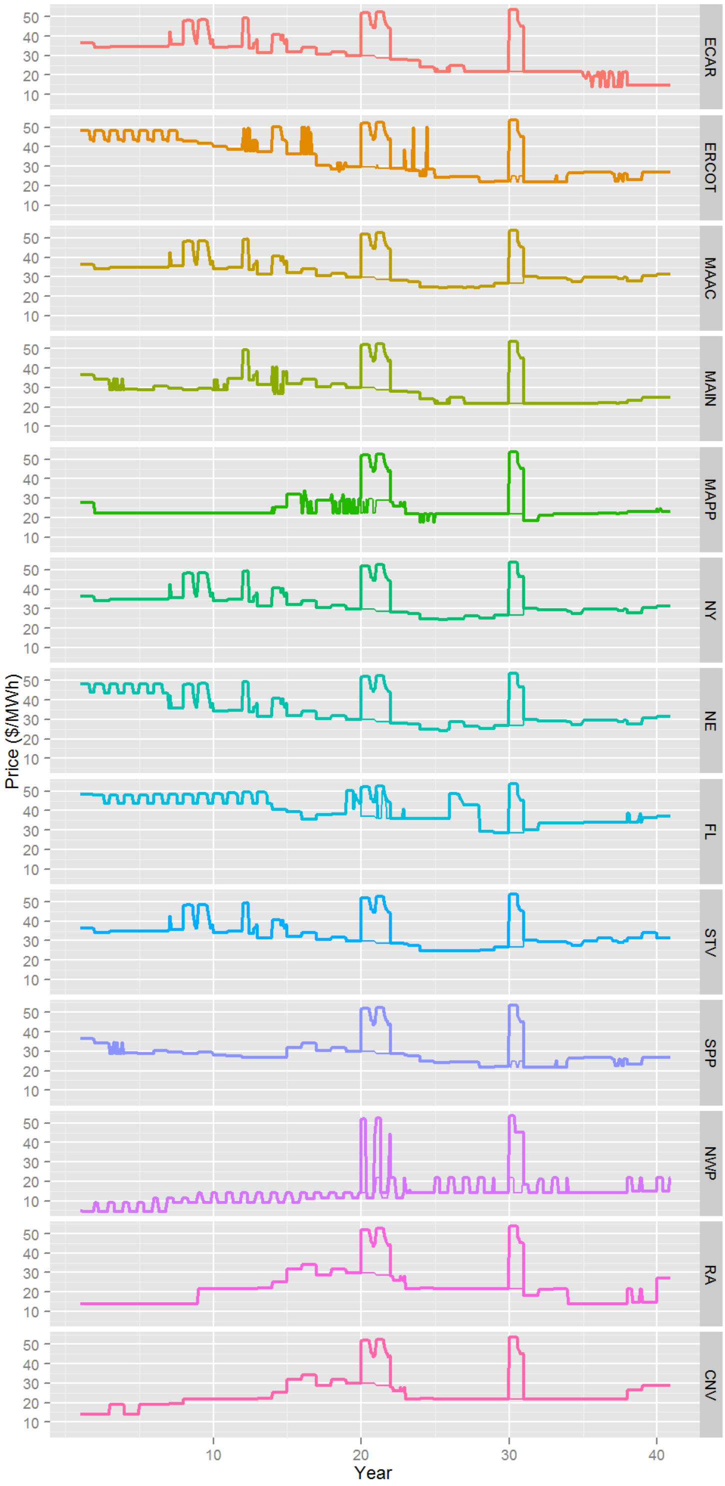


Figure 5.18: Electricity price for event 2: loss of coal generation

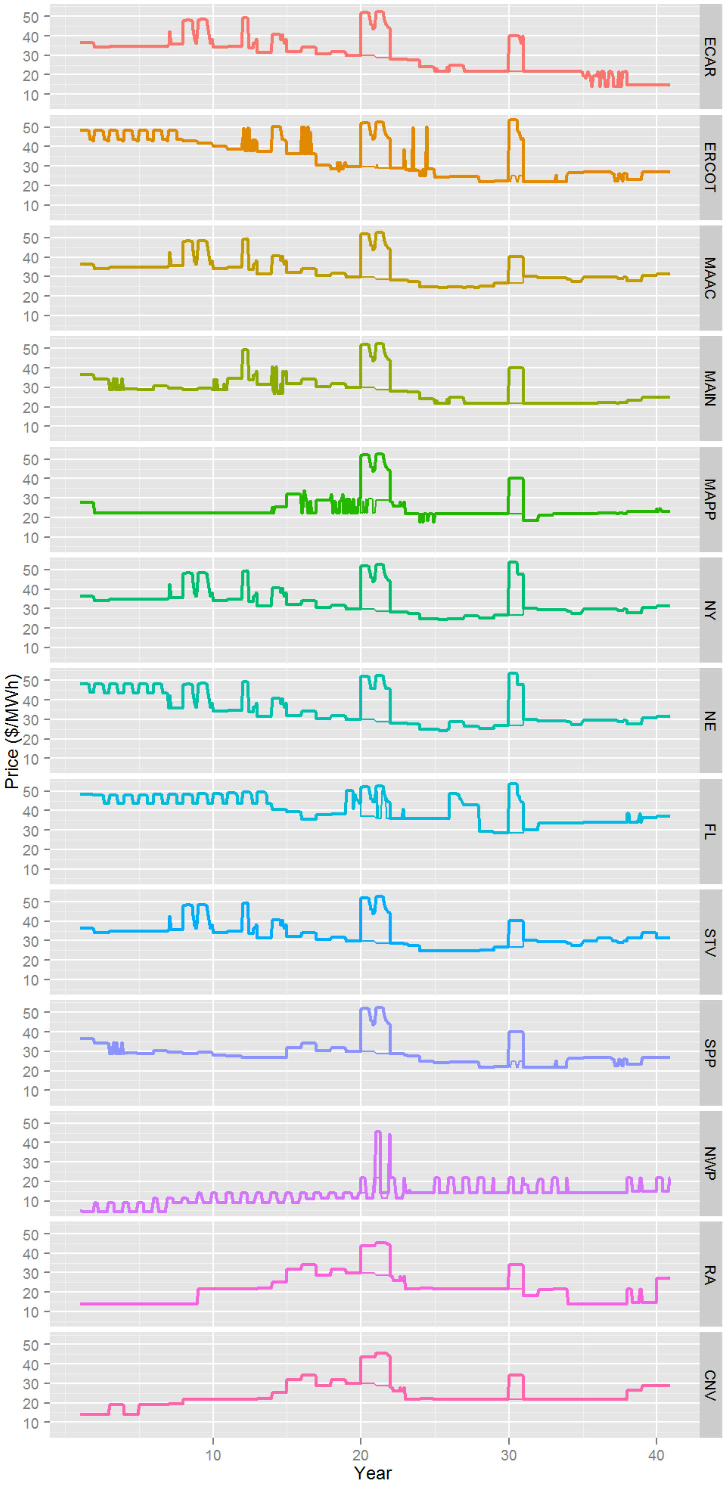


Figure 5.19: Electricity price for event 3: loss of renewables

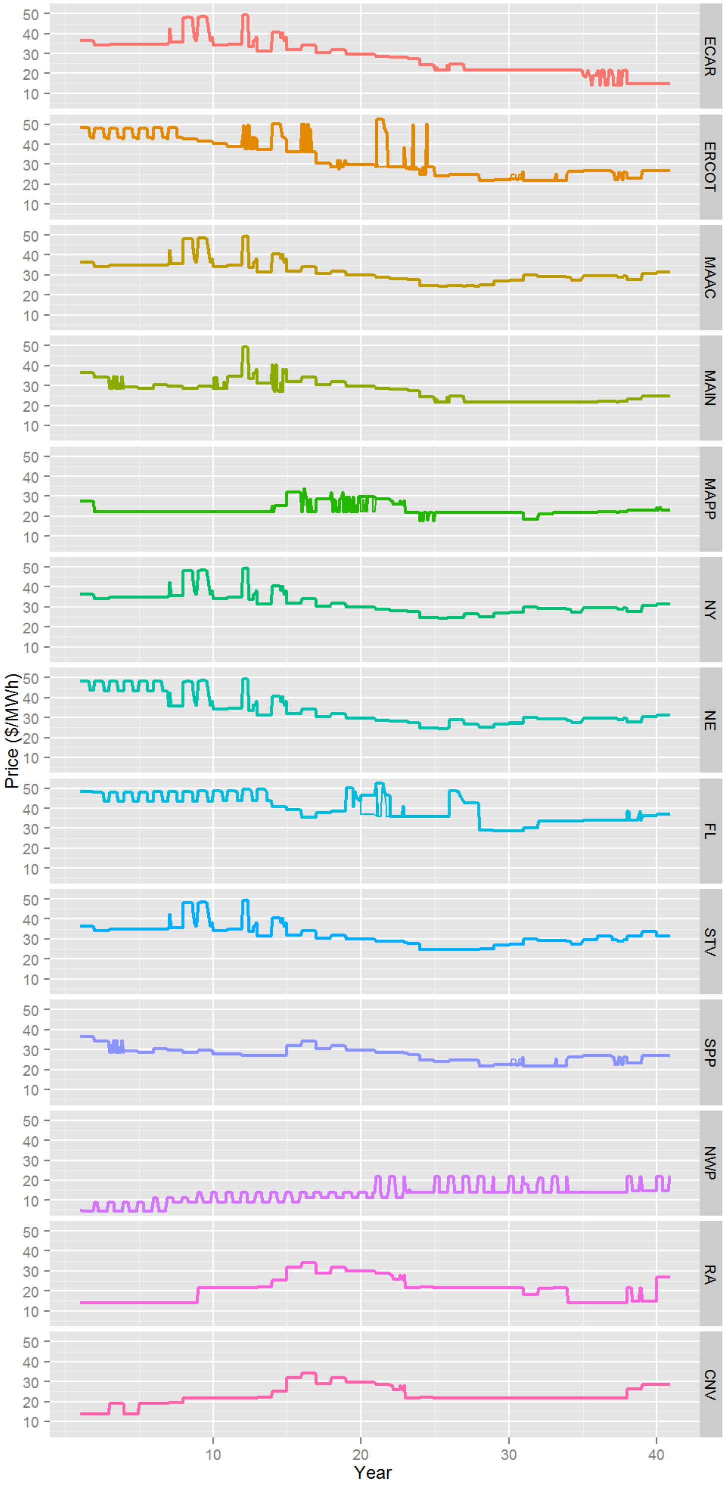


Figure 5.20: Electricity price for event 4: loss of natural gas sources

It is not a surprise that the first event is the most significant after seeing the importance of nuclear energy in the production of electricity for this particular solution. The rest of the events are also significant in terms of overall increase in cost. The impact of the last one is not as significant on the electric system since natural gas generation is not an important part of the actual production in the base case. This is due to the fact that no load duration curves are being used in this example and, thus, base load units provide the bulk of the electricity demand. For the event at year 20, ERCOT is the only electric regions with a noticeable increase in electricity prices, followed by Florida on a smaller scale. The biggest impact of this event is an average 3% increase in the natural gas price, which is small for such the type of disruptive event that it intends to model [46]. This type of event would be better modeled with an increase in the price of natural gas along with the loss of capacity.

Figures 5.17–5.20 suggest that changes in energy prices could be used to develop new metrics, in the spirit of the efforts presented in Section 2.6.3, and they could help characterize other aspects of resiliency beyond the increase in overall operational cost. The figures are compatible with the evolution of performance measures presented in Fig. 2.13, where the maximum performance deviation is here the increase in electricity prices. Recovery times are fixed in this case since the recovery of the system is predetermined by the input data. The area between the basic prices and those under the events would correspond to the consequence. Many metrics could be developed based on the consequence and, by duality theory, this would include the total operational cost if consequence was summed across all arcs using demand as weights.

5.3 Introducing DC power flow

The goal of this section is to study the effect of DC power flow constraints on the simulation. These constraints will not be used beyond this section. Use of the DC power flow model is more adequate in cases with a much more granular description of the power system, but it is implemented for this 13 node system. Average lengths are estimated and reactances (inverse of susceptance) are calculated using 0.0001 p.u. per mile. The power base for the electric system,

P_E, is set to 1 GW.

With these values the simulation is performed in NETPLAN and some significant results were found. Electric transmission is directly affected by the DC power flow constraints and the effect can be studied by comparing the average transmission without and with the constraint in Figs. 5.21 and 5.21, respectively. In general, transmission is reduced by an order of magnitude. Electricity flowing into California and Florida is greatly reduced and the flow into Texas switches direction.

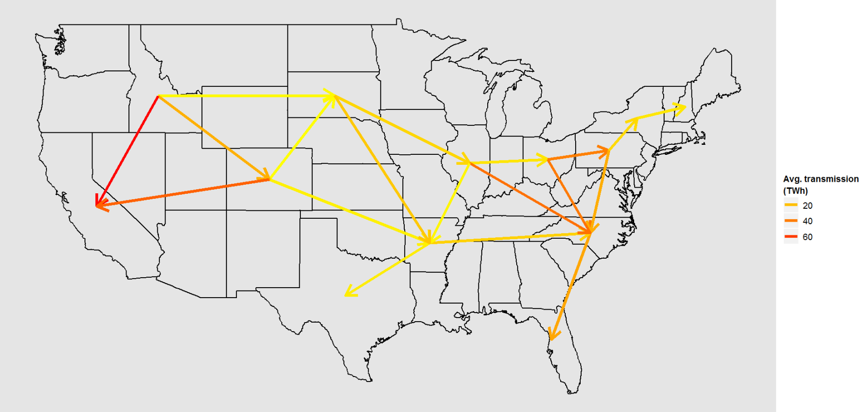


Figure 5.21: Average electricity transmission for the base case

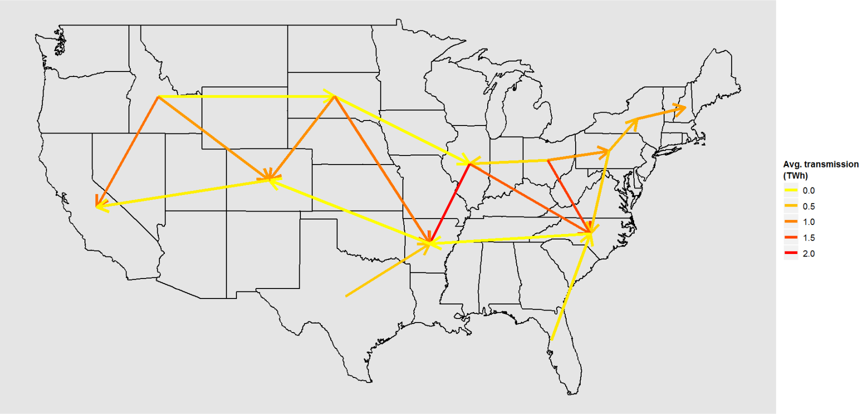


Figure 5.22: Average electricity transmission with DC power flow

These changes in transmission affects how the generation is produced in the different regions. Comparing Figs. 5.23 and 5.24 we can observe that production increases significantly in ECAR, Florida, NE, STV and CNV, while it decreases in MAIN, MAPP, and NWPP. The increases are most important in the outer regions for which it is not possible to import cheap energy. The main reductions in production happens for hydro and wind generation, which have the lowest operational cost and are dispatched first in absence of constraints.

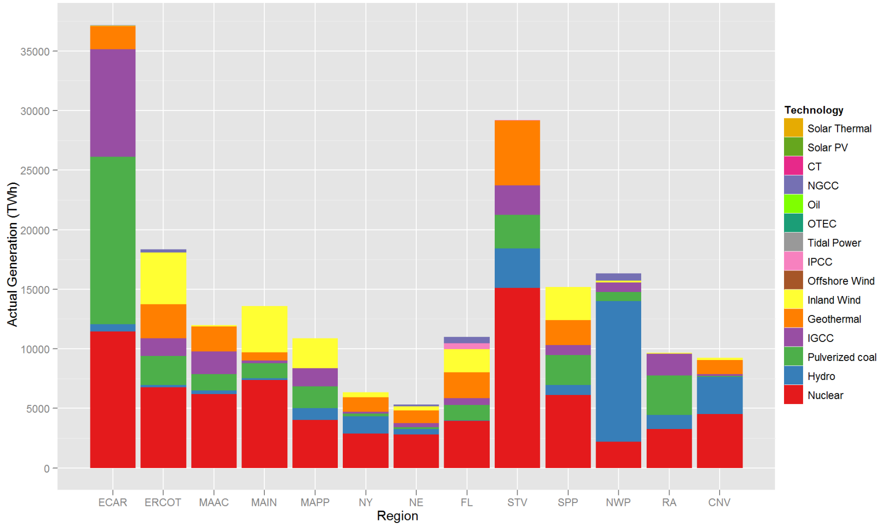


Figure 5.23: Electricity generation by area for the base case

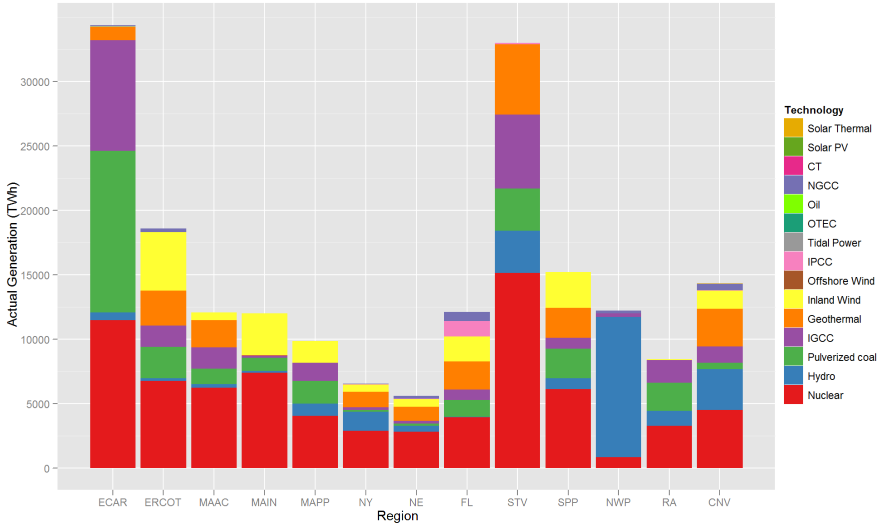
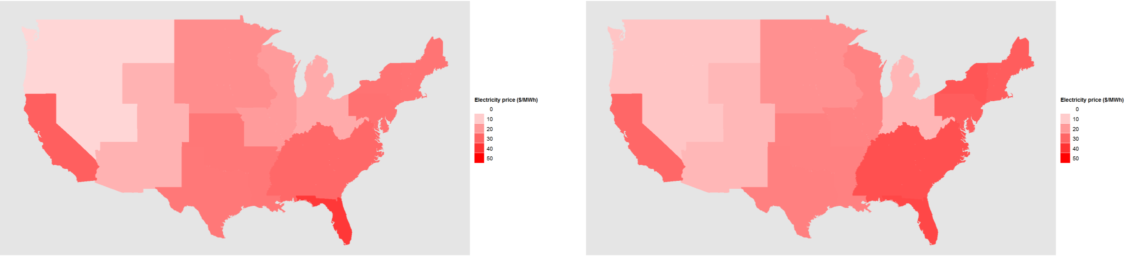


Figure 5.24: Electricity generation by area with DC power flow

The changes in generation and transmission have a significant impact on electricity prices. When comparing Fig. 5.25 with Fig. 5.11 we see that during the first years the gradient of prices is much larger. Regions that are not very interconnected (NE, ERCOT, FL, CNV) show prices that are much higher, while the cost is reduced, especially in NWPP. These price variations are directly correlated to the changes in generation across regions observed before. The differences soften as time goes by and for most cases the prices end up being similar with and without DC power flow constraints. California presents the largest increase, mainly due to the fact that electricity from the Northwest Power Pool hydro facilities cannot be dispatched as easily to satisfy its demand.

(a) Years 1–10

(b) Years 11–20



(c) Years 21–30

(d) Years 31–40

Figure 5.25: Average electricity prices with DC power flow

Resiliency calculations are also affected by the DC power flow formulation. In the base case, all regions were affected by the first three events and a few for the fourth. Figures 5.26–5.29 reveal that NWPP is not affected by any event since hydro generation is not fully utilized in normal conditions. The loss of coal-based generation does not affect ERCOT, MAIN, NY, FL, SPP, or CNV as severely. ECAR has a very small share of electricity produce from renewables and thus event 3 does not have an impact. On the other hand, the loss of natural gas has an impact on more than half of the regions (ERCOT, MAPP, NY, NE, SPP, and CNV). The magnitude of the consequence for event 4 is still the smallest, but is more spread throughout the system.

5.4 Introduction of load duration curves

We saw that the use of monthly electric system in the base case produces results that lean towards base load generation. This is due to the smaller operational cost of these units when they are constantly utilized. But actual electric load fluctuates withing the hours of the day, presenting peaks in the afternoon and evening and valleys during the night hours.

Electric demand variability is usually described using load duration curves, which are mostly used for mid- and long-term planning. A load duration curve (LDC) is obtained by ordering the demand data in descending order of magnitude, rather than chronologically. Fig. 5.30 shows an example for the New England region in 2005 [53]. An approximation to the load duration curve can be done by dividing it into several segments and taking the mean demand for said segments. In the figure 3 load steps are defined.

This method reduces the need of using uniform time steps while still capturing the different power demand levels. Following the proportions in Fig. 5.30 the electric load in the base case is divided into three load steps only for the calculations in this section. Making this change implies redefining the demand terms for PHEVs. Using the curves in Fig. 4.18 and average daily load, the terms in Table 5.3 are used to simulate the demand of a single PHEV in the electric system. With this methodology charging patterns for PHEVs are preserved.

Table 5.3: PHEV demand with load duration curves

	1st step	2nd step	3rd step
Duration	10%	40%	50%
Power (KW/vehicle)	0.282	0.167	0.056

The resulting generation of electricity in Fig. 5.31 can be compared to the base case (Fig. 5.3). In this case, natural gas combined cycle units are used more often. The differ-

Figure 5.26: Electricity price for loss of nuclear energy with DC constraints

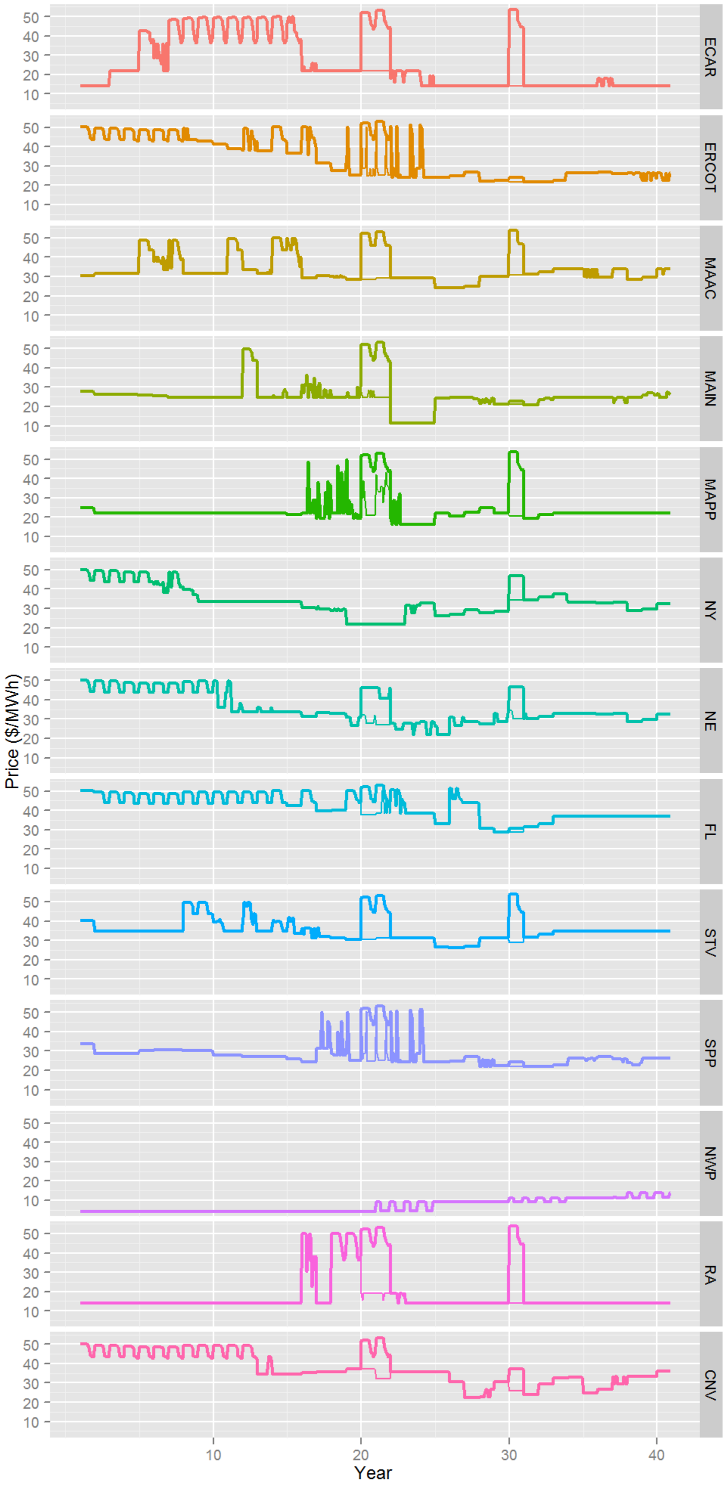


Figure 5.27: Electricity price for loss of coal generation with DC constraints

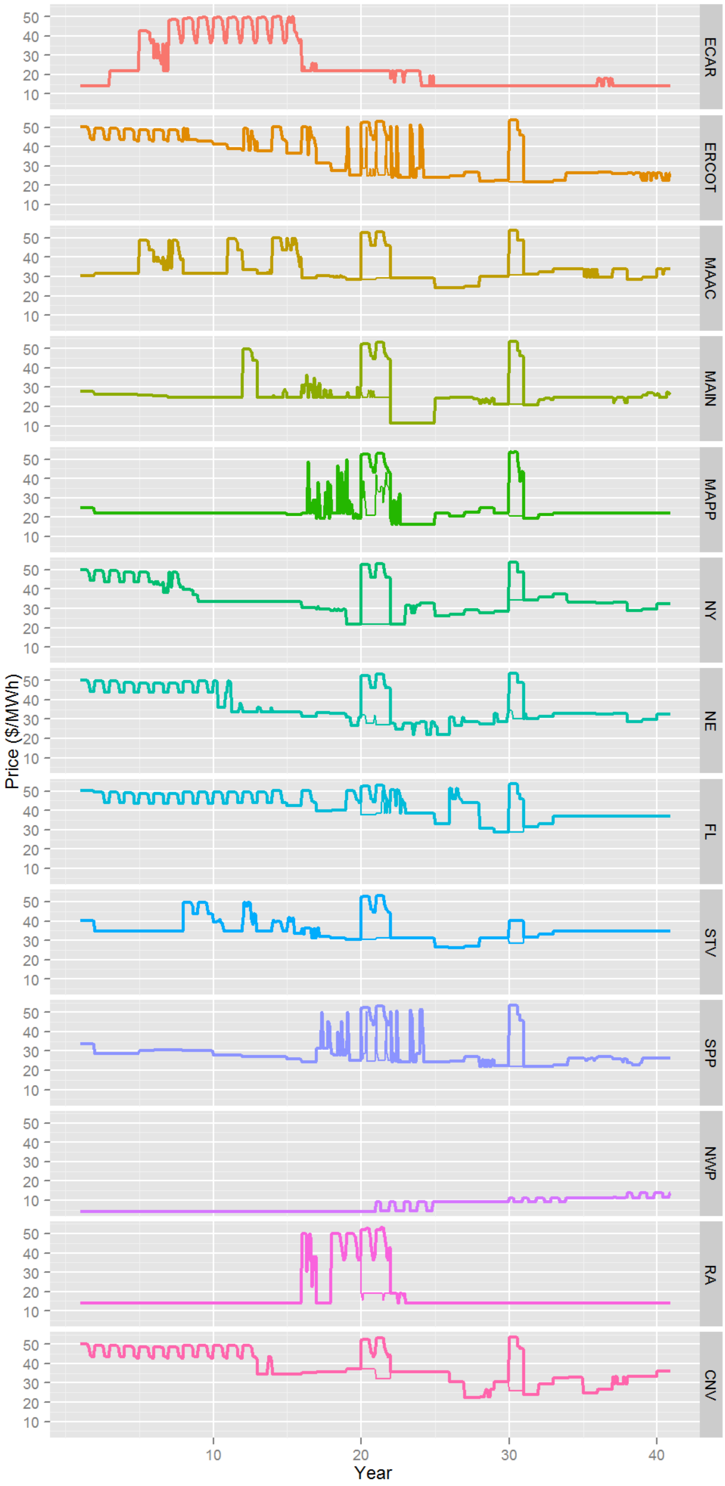


Figure 5.28: Electricity price for loss of renewables with DC constraints

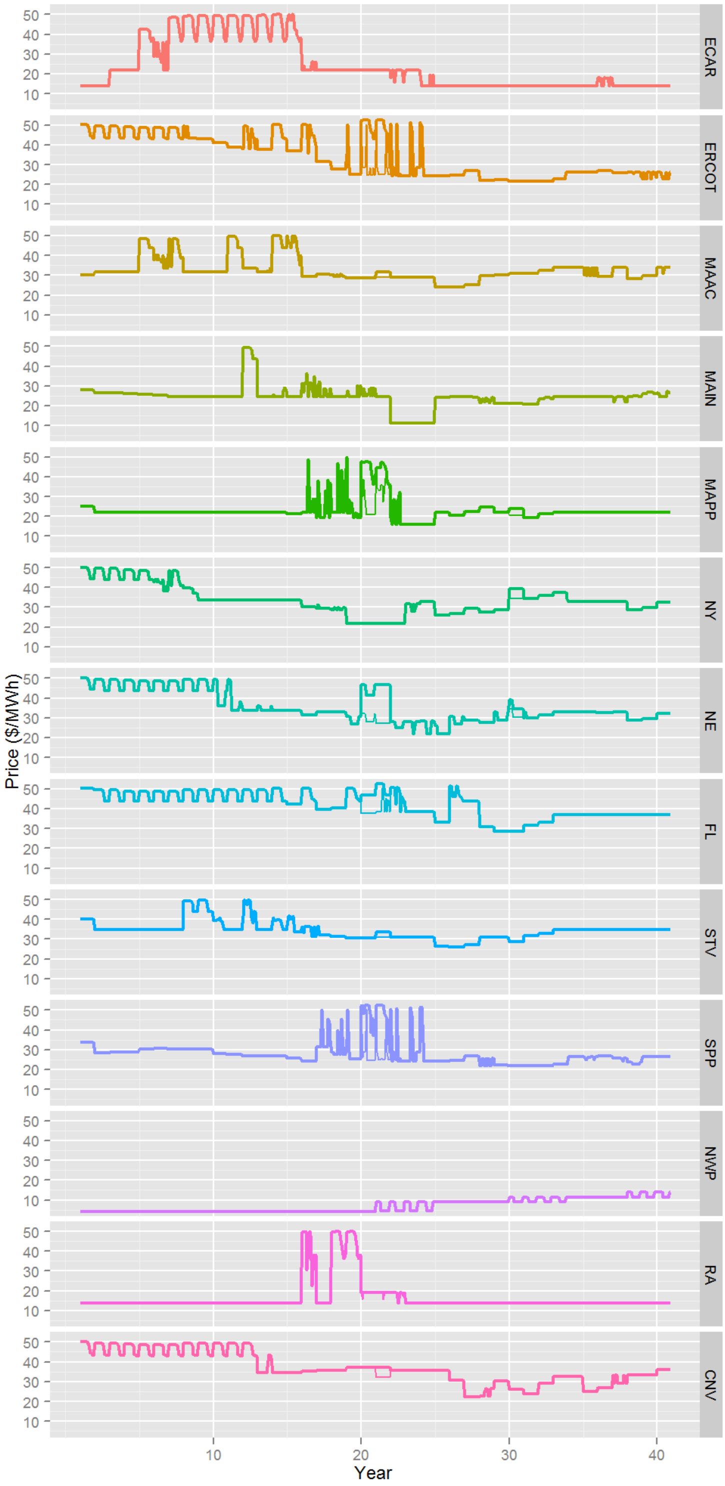


Figure 5.29: Electricity price for loss of natural gas sources with DC constraints

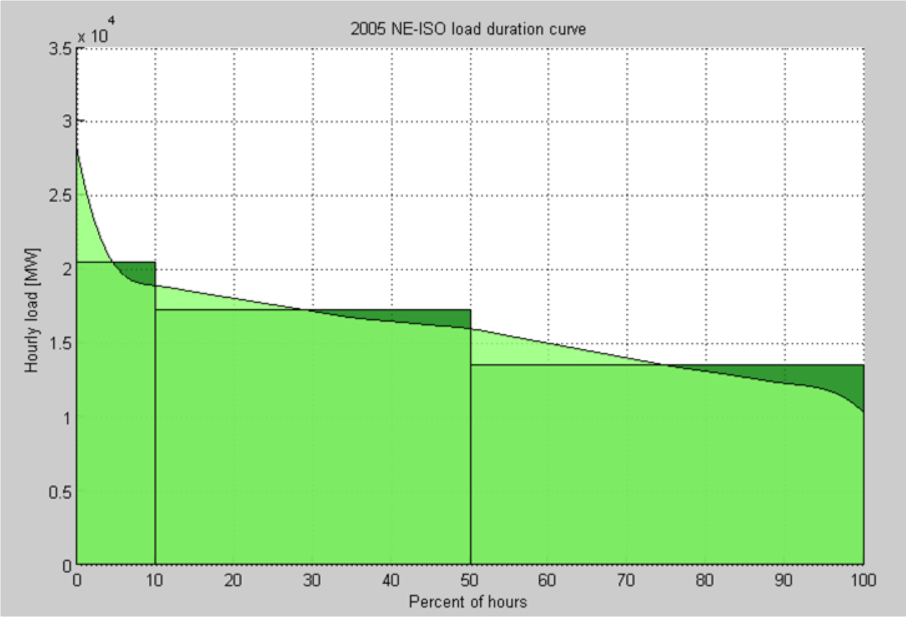


Figure 5.30: Load duration curve (Source: [53])

ence is not large but it indicates a step in the right direction. Other features in the model such as investment profiles or the transportation sector are affected very slightly or not at all. Natural gas demand increases as it is being used for as a generation fuel.

One of the drawbacks of the increase in granularity in time is the increase in model size. Introducing the three steps from the LDC generates a linear program with 1,499,114 variables and 774,360 constraints. These values were respectively 748,394 and 472,920 in the base case model. Solution time has a dramatic increase to 2 hours and 16 minutes, when the base case model was solved by NETPLAN in 17 minutes on average.

5.5 Sensitivity analysis

There are a number of ways that the modeling framework presented in this dissertation could be validated. The following is a list of the most relevant options:

- Comparison of forecasts to other available studies, such as EIA's "Anual Energy Outlook" [122], or DOE's "20% Wind Energy by 2030" [123];

- Repeat the numerical analysis with other tools, for instance, MARKAL/TIMES [6],

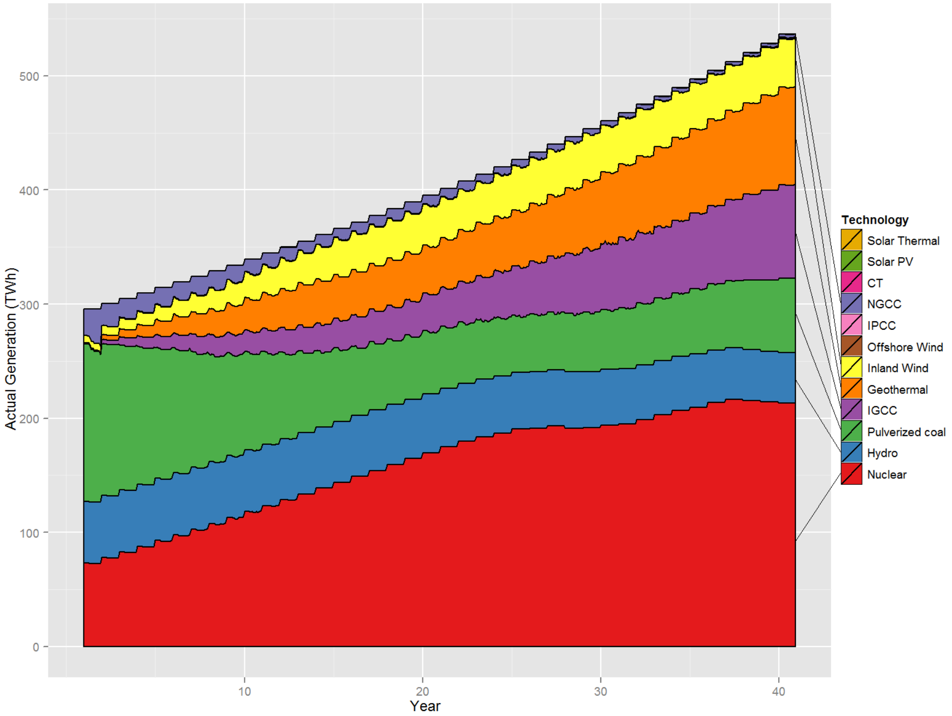


Figure 5.31: Electricity generation with LDC

NEMS [5], or ReEDS [21];

- Use the model to simulate a period of time that has already past ago and compare predicted results with actual observed investment trends;
- Perform sensitivity analysis on the modeled system and evaluate the results with respect to changes that are to be expected.

The first two options involve comparing results of tools that do not necessarily have the same modeling scope than NETPLAN does. A combination of tools or studies might be needed to validate all the features in this new framework. Comparing our results to other studies implies using results that are not based the same data. This issue can be amended by directly utilizing other planning tools, but this approach involves a considerable effort since one has to learn how to use other tools and format the data appropriately. Also, NETPLAN produces a Pareto front of solutions and the first three options provide only a limited number of cases, most often one solution.

For all those reasons, the fourth option is considered as a reasonable balance between necessary time and effort to carry out the validation and ability to examine all the features in NETPLAN. To test and validate the modeling framework for operation and investment planning of the transportation and energy system, a series of nine cases have been chosen and tested. The differences and similarities in the solutions will be explored to surface expected interactions between the systems. The cases that are tested are:

- Change inflation rate from 2% to 3%
- Increase discount rate from 2% to 3%
- Double the production cost for coal
- Increase price of gasoline and diesel by 50%, 100% and 300%
- Limit future carbon emissions to limits at first year
- Reduce future CO_2 emissions so that emissions at year 40 are reduced 20% and 40% with respect to the first year

The following table summarizes the different metrics for the minimum cost solution under the cases tested. The base case scenario has been added as a reference.

Table 5.4: Metrics under sensitivity cases

Case	Cost (Trillion $)	Emission coefficient	Emissions (billion short ton)	Resiliency (Billion $)
Base case	52.30	0.304	116.5	36.56
3% inflation	63.74	0.347	118.1	44.58
3% discount	43.39	0.300	116.5	28.52
Double coal cost	56.24	0.203	105.2	37.18
50% increase price gasoline	65.32	0.201	111.3	36.51
Double price gasoline	76.94	0.198	109.3	37.34
Quadruple price gasoline	122.67	0.255	109.9	38.74
0% increase emission	52.32	0.229	111.4	43.61
20% decrease emissions	52.43	0.124	104.4	83.42
40% decrease emissions	52.96	-0.033	52.7	88.21

The following are plots of the CO_2 emissions for each one of the cases tested. Total emissions are included in Fig. 5.32 and these are also divided in emissions from power generation and transportation in Figs. 5.33 and 5.34, respectively.

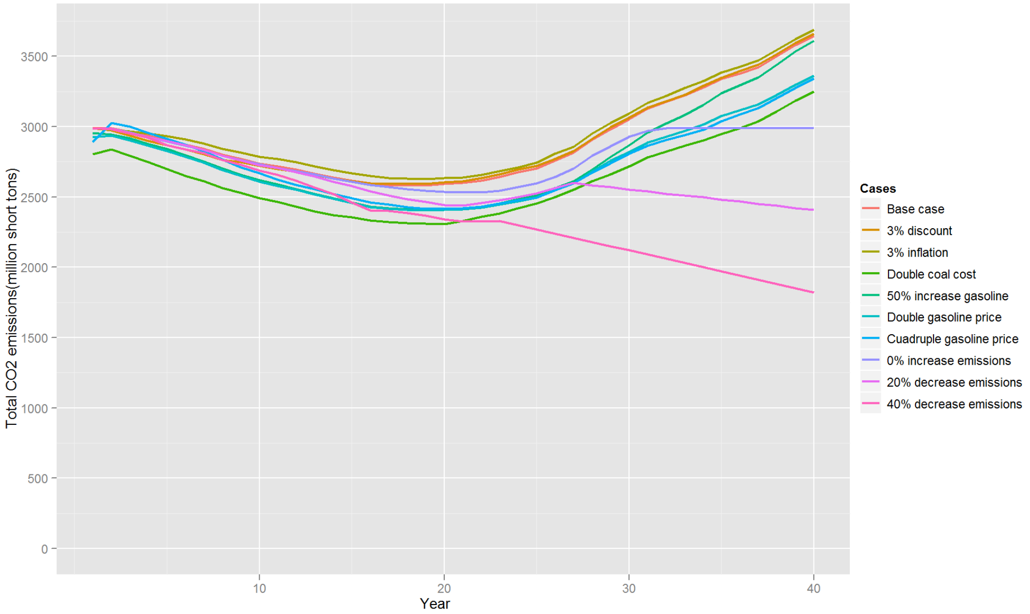


Figure 5.32: Evolution of total CO_2 emissions

The increase in inflation rate creates an expected increase in the present value of cost and resiliency. A 3% discount value has the opposite effect on both metrics but neither of them has a significant impact on emissions, just a slight increase in case of the inflation parameters. The biggest change can be observed in the transportation investment as seen in Fig. 5.35 in comparison to Fig. 5.14. The peak observed in the second year for the base case and the investment in the rest of the simulation are modified. In case of the increased inflation, more investment is done at the beginning but the situation is reversed for an increase in discount rate. This is to be expected since those parameters affect the present value of the transportation investment evolution.

This change in transportation investment has a direct effect in electricity generation. The increase in discount rate (and related decrease in early transportation investment causes a reduction in the production and transportation of coal and, thus, combined cycle units are

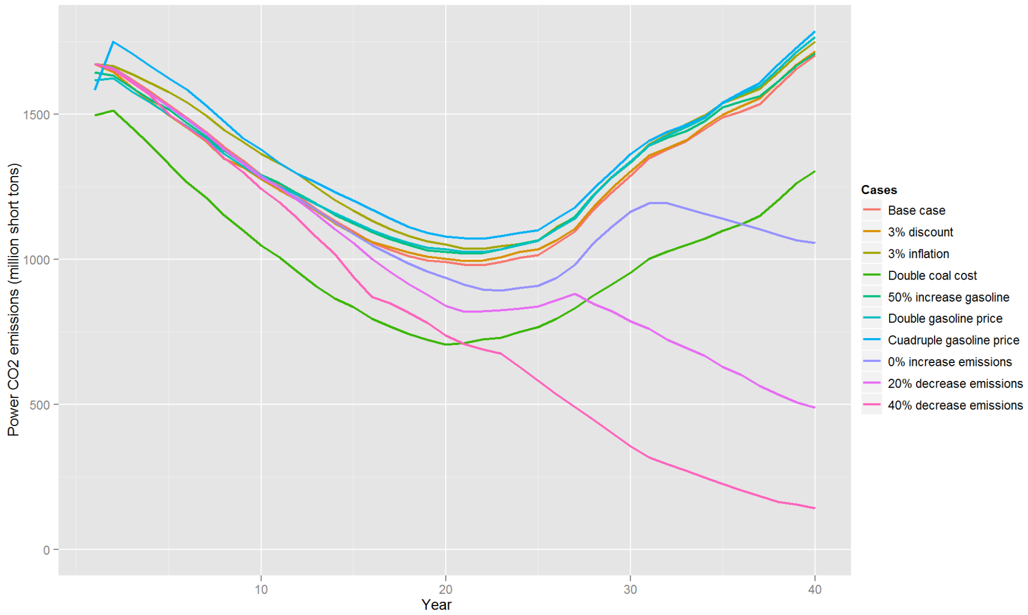


Figure 5.33: Evolution of power CO_2 emissions

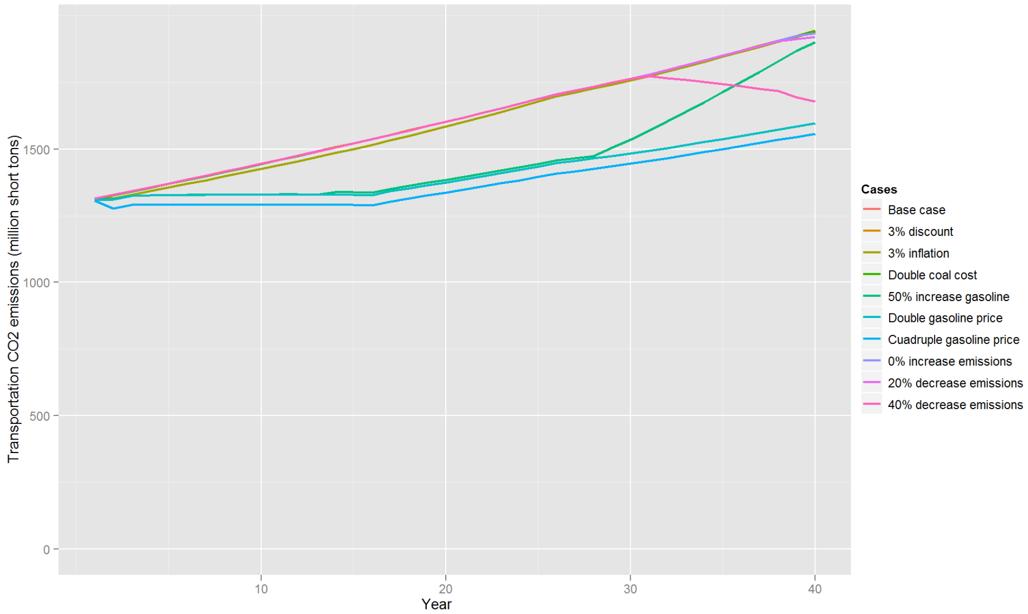
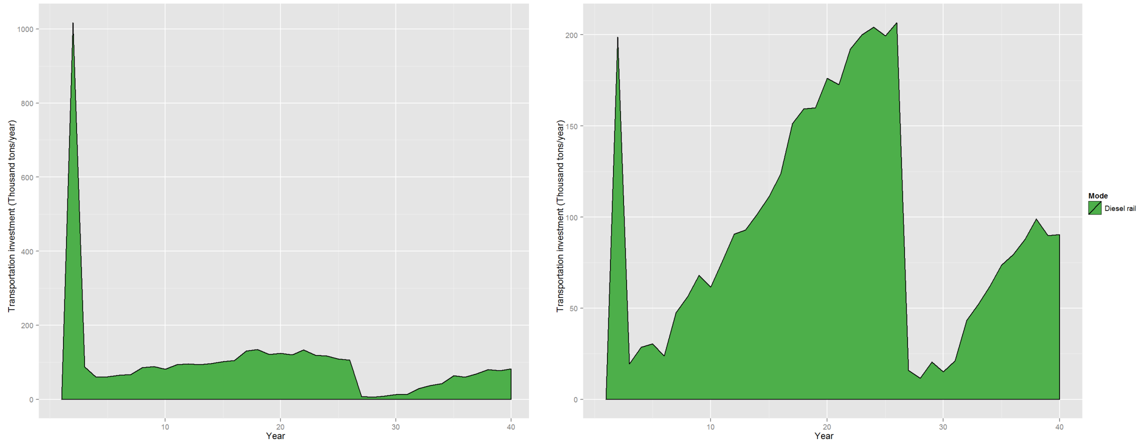


Figure 5.34: Evolution of transportation CO_2 emissions

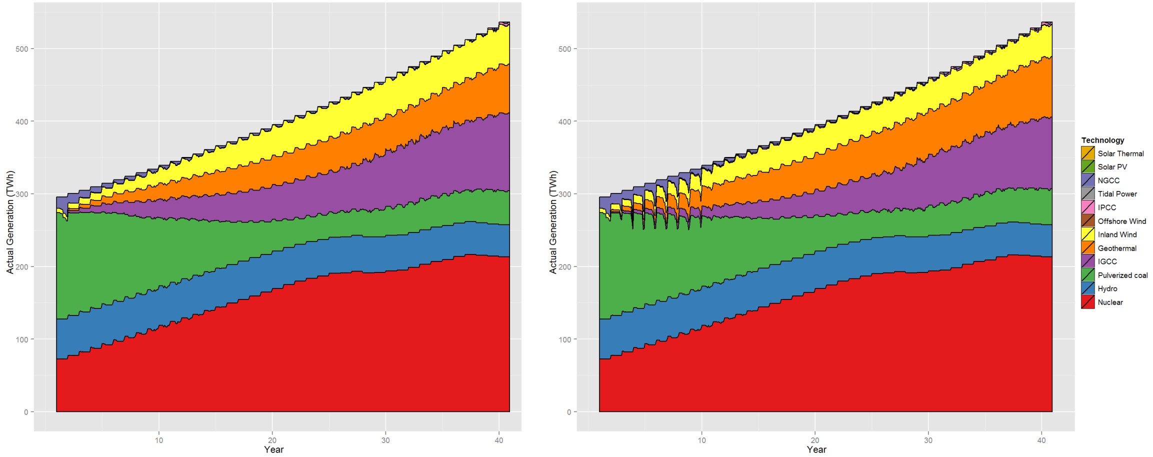


(a) 3% inflation

(b) 3% discount

Figure 5.35: Transportation investment with change in economic parameters

more heavily used in the first years. The opposite is true for the increase in inflation rate, which leads to a more pronounced use of coal-based generation and is the reason why emissions for this case increase.

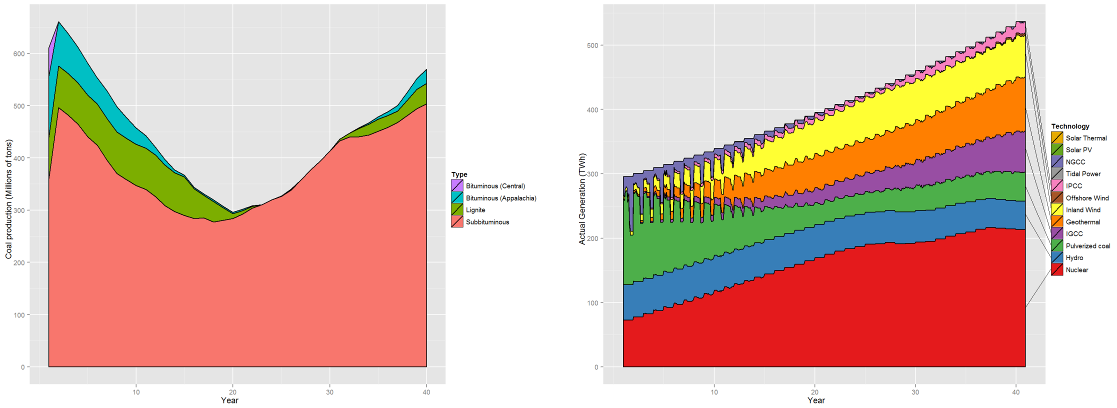


(a) 3% inflation

(b) 3% discount

Figure 5.36: Electricity generation with change in economic parameters

The increase in coal production costs has a dramatic impact in emissions from the power sector as seen in Fig. 5.33. This impact is comparable to the limitation of emissions with aggressive reductions in carbon emissions. The most significant impact is captured by the reduction of demand of coal in Fig. 5.37a. Another interesting observation is that the reduction largely affects and, in fact, it increases the demand of subbituminous coal for the second year by 50 million tons, when compared to Fig. 5.6. This increase is also reflected by a increase in transportation investment for that same year, which is the only measurable change in the transportation sector.



(a) Coal production

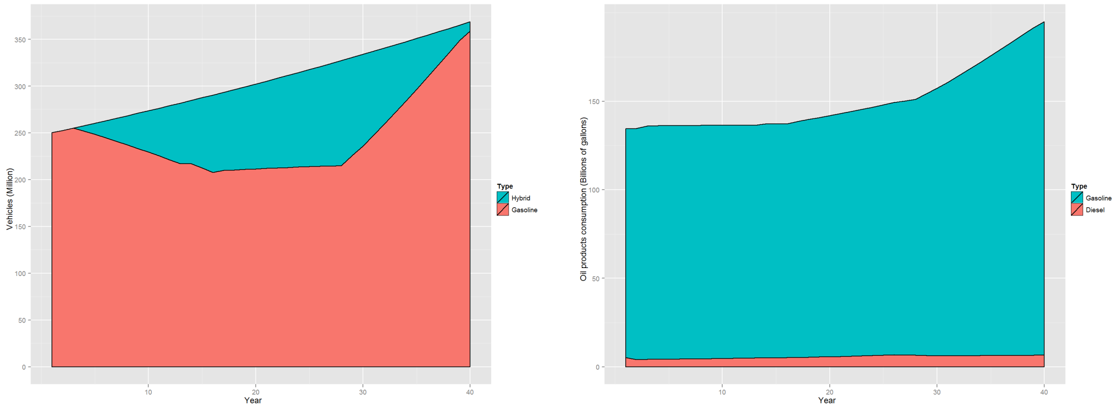
(b) Electric generation

Figure 5.37: Increase in coal production cost

In Fig. 5.37b we can see that coal-based generation is still part of the solution but with a reduced present. In the earlier years natural gas combined cycle is much more present, specially for the periods in which natural gas demand is smaller and its production price lower. The decrease investment in coal-based facilities is compensated by the inclusion of other renewable sources such as IPCC, off-shore wind and tidal power. This case is an extreme example of an introduction of a carbon tax, which can effectively switch the generation mix and investments in the system. Although not currently implemented in NETPLAN, it could prove to be an interesting tool to further explore the solution space in the multiobjective optimization.

The increase in oil-based fuels also has a significant impact on the resulting solutions. Table 5.4 shows that costs increases very significantly due mainly to the high demand of gasoline from passenger vehicles. The increase in resiliency metrics is not so radically different since the events affect mainly the electric and natural gas sectors.

Emissions in the first case decrease in the intermediate years due to the introduction of hybrid vehicles. This is captured in Fig. 5.38 along with the decrease in gasoline demand. The end in the simulation sees a return to gasoline vehicles because the savings in fuel during the final years cannot compensate the difference in investment costs, even taking into account salvage values.



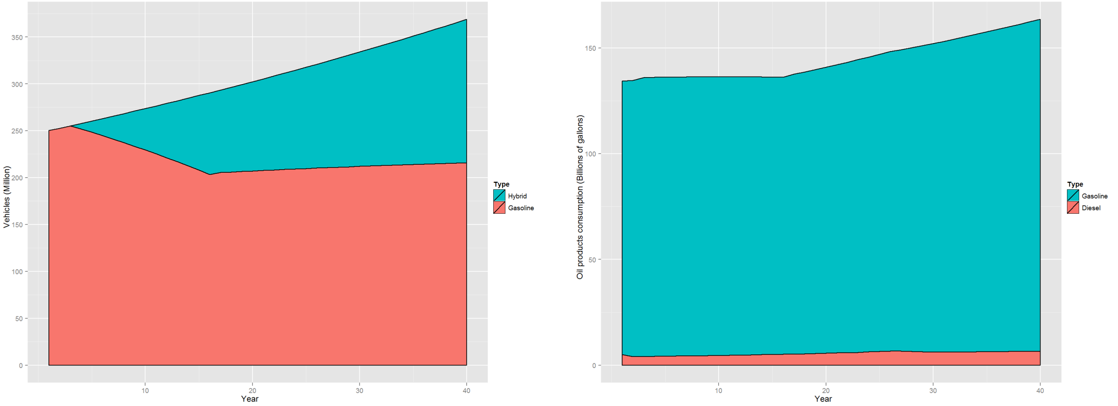
(a) Passenger transportation

(b) Gasoline and diesel demand

Figure 5.38: Transportation results with 50% increase in gasoline prices

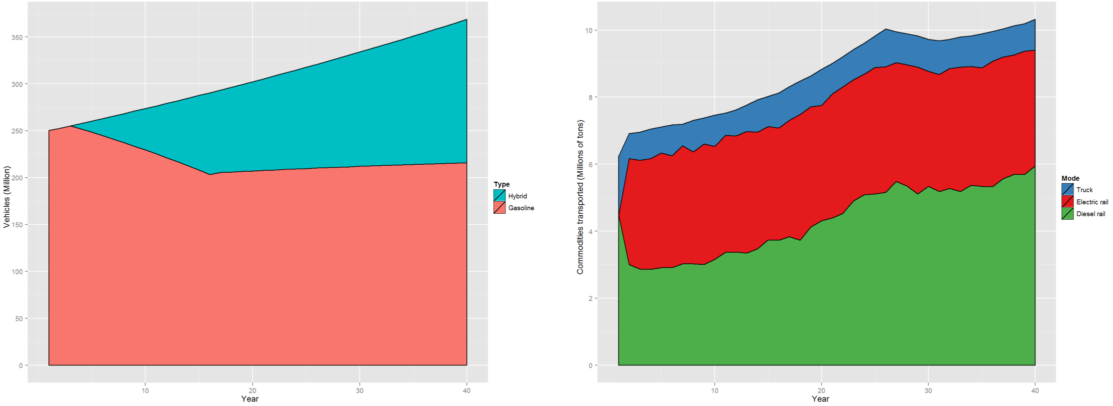
When gasoline prices are double this effect is even more evident, further reducing emissions. In Fig. 5.39 hybrids are introduced at the same pace and remain through the simulation achieving the maximum market penetration allowed. These cases begin to show the interaction between transportation investment and fuel prices. A more detailed study would allow to find the break-point for which the switch between gasoline and hybrid vehicles.

Unlike the previous examples, making oil fuels four times more expensive trigger the investment on electric rail while maintaining hybrid penetration. Figure 5.40 shows that this shift is radical and it happens primarily at year 2. As a consequence, the transportation emissions are the lowest for this case in Fig. 5.34 but triggers an increase in power emissions. The emission coefficient in Table 5.4 is larger than in the previous case even though the total emissions is practically the same. The reason is that the increase in oil fuels causes a reduction of coal transportation in the first year and hence the reference for calculating the emission index is smaller.

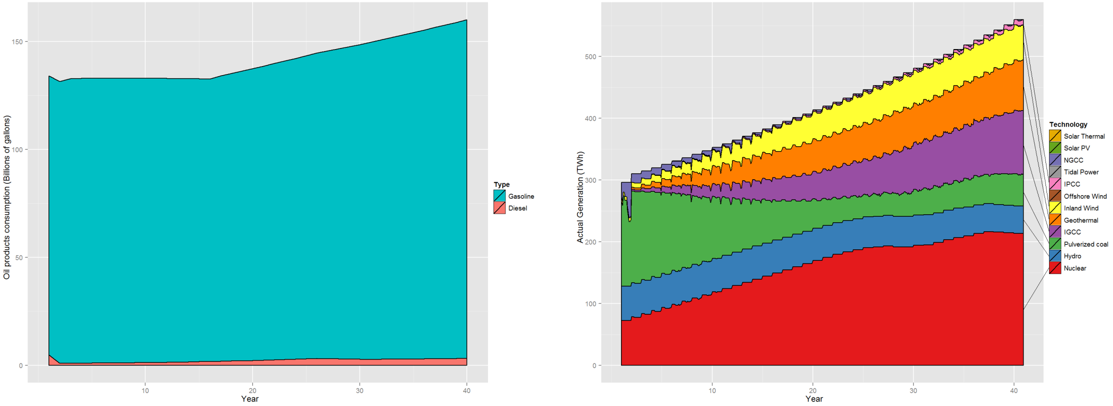


(a) Passenger transportation (b) Gasoline and diesel demand

Figure 5.39: Transportation results when gasoline prices double



(a) Passenger transportation (b) Freight transportation



(c) Gasoline and diesel demand (d) Electricity generation

Figure 5.40: Transportation results when gasoline prices quadruple

Figure 5.40d also shows a 5% increase in electricity demand with respect to Fig. 5.3, which has to be accounted for in the planning of the electric system. This increase and electricity prices are crucial in deciding the balance between diesel and electricity rail. At the same time they are related to the assigned fuel consumption for electric rail, which in this case was overestimated by assuming that tank-to-wheel efficiency was the same for electric and diesel trains. A refinement of these values would make the breaking point price for electric rail significantly smaller than in this case.

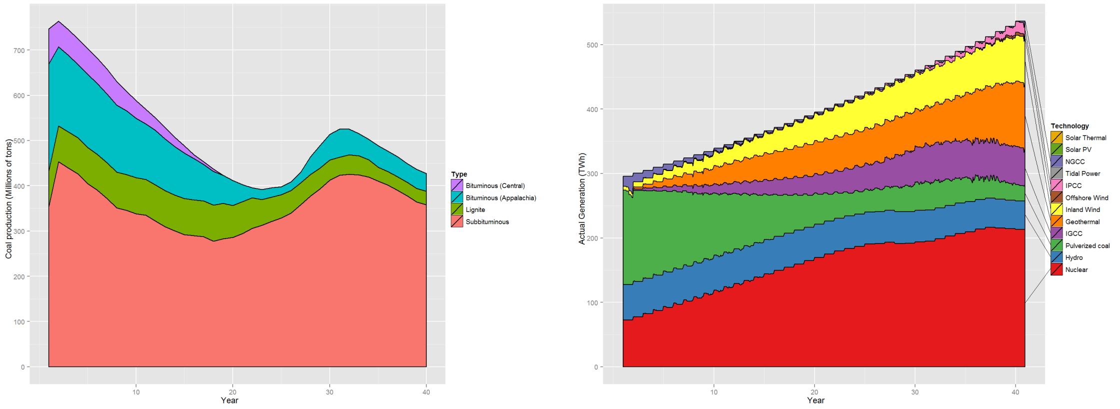
Finally, we study the cases with emission limits enforced. According to Table 5.4 the present value of cost does not increase significantly. However, the system becomes much more prone to large consequences, as the resiliency metric increases and nearly triples for the most severe case. From the emission plots (Figs. 5.33–5.34) we can determine that the biggest reduction comes primarily from the power sector. In the most severe case transportation emissions decrease in the last decade of the simulation. This corresponds to the introduction of hybrid vehicles, the only change in the transportation sector. Switching the remaining of the vehicular fleet to hybrids would allow further reduction of carbon emissions.

Figure 5.41 reproduces side by side the evolution of the electricity generation sector. As it happened in the case of coal cost increase (Fig. 5.36), electricity production switches from coal-based power plants to renewables. This is also reflected in the decrease in coal demand on the left hand side, to the point that it completely stops in the worst case.

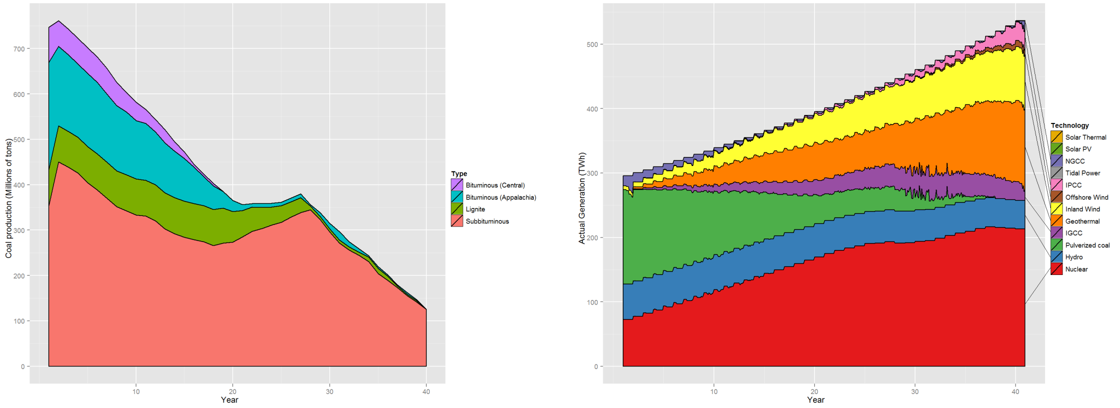
This approach is consistent with the introduction of caps or cap-and-trade mechanisms. Future development of NETPLAN could introduce these methods as decision variables, making it possible to consider new pieces of legislation as part of the solution space along with carbon taxes.

5.6 Pareto front of solutions

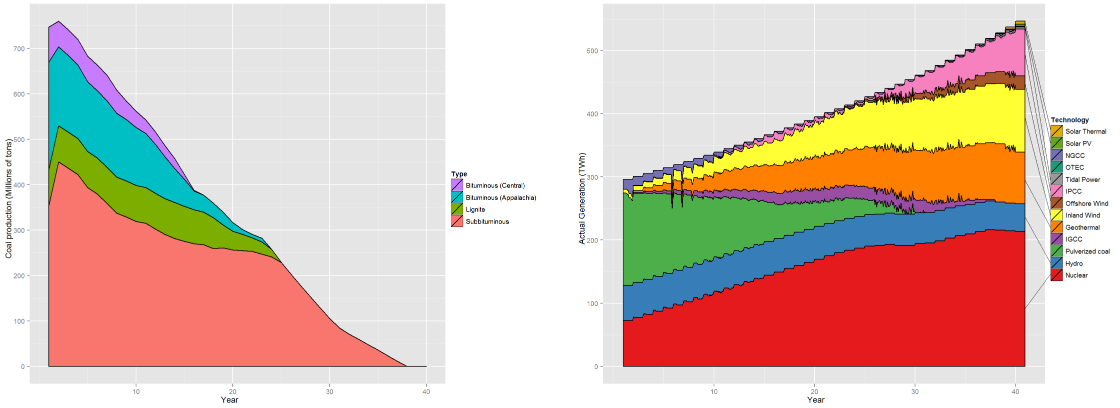
The multiobjective optimization is carried based on the system analyzed in the previous section. Eighty-two generations are evaluated with 20 individuals per generation for a total computation time of 471 hours. The solution process was stopped at this time, although the



(a) Coal production (0% increase) (b) Electric generation (0% increase)



(c) Coal production (20% decrease) (d) Electric generation (20% decrease)



(e) Coal production (40% decrease) (f) Electric generation (40% decrease)

Figure 5.41: Simulations with different carbon emission limits

process did not converge and better solutions could be found. A simple criteria to evaluate convergence consists on observing the rate at which new solutions enter the current best population for each generation. At the moment of stopping the simulation, 2 to 4 solutions were being selected as best solutions per iteration. The implementation of the parallel NSGA-II method described in Section 3.8 would greatly benefit total computation time.

The multiobjective optimizer modifies 7,384 investment variables that are enforced as minimum investments. The result is the Pareto front in Fig. 5.42, showing that they are all non-dominated. All the objectives are highly correlated and low cost values lead to high emissions and event impacts and viceversa. All three objectives are minimized.

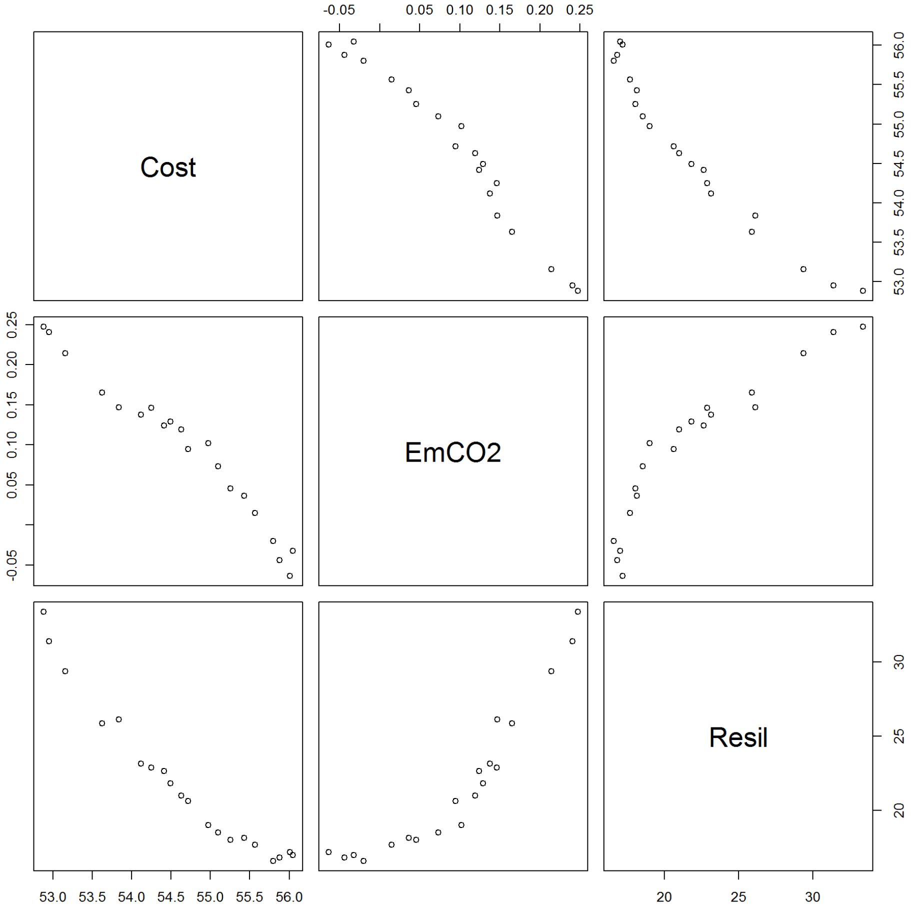


Figure 5.42: Pareto front of solutions

The different solutions are compared in the following graphs. Solutions are ordered by increasing cost and, consequently, by decreasing emission and resiliency metrics. Figure 5.43 shows that carbon emission for power and passengers also decrease as cost increases. The same is then true for total emissions.

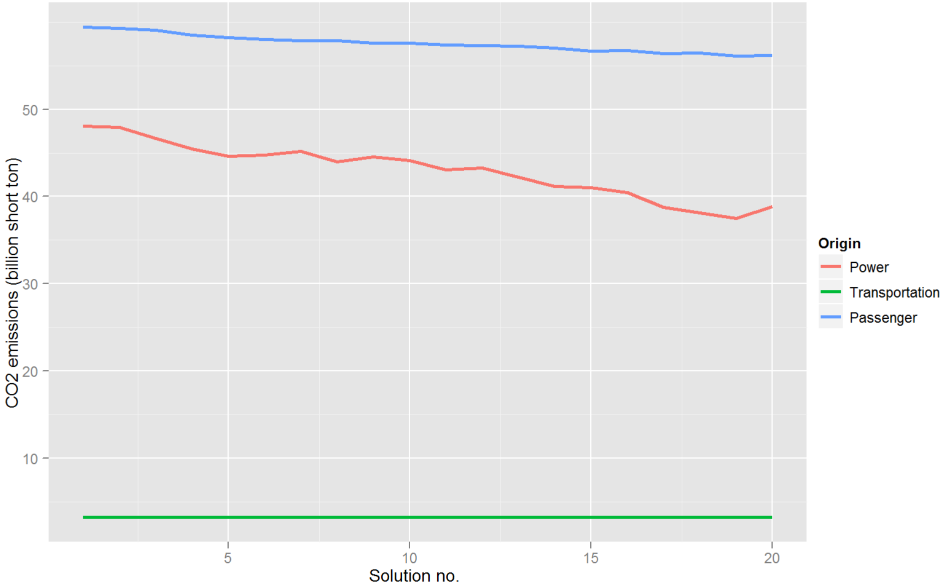


Figure 5.43: Emissions for the best solutions

Figure 5.44 contains the representation of total emissions versus the index that was developed in Section 2.6.2. The linear relationship shows that the index also does a good at tracking total emissions, while giving a higher priority to trends that decrease emissions.

The impact of events also tends to decrease as investment costs increase, except for the loss of natural gas (Fig. 5.45). Higher costs are also associated with higher levels of investments, as shown in Fig. 5.46 and, thus, the resulting system is more resilient and the impact of events on the electric system is less acute.

Figure 5.46 also shows that the investment portfolios are more diverse for solutions with higher cost and they include more renewables. This contributes to the reduction in CO_2 emissions that was identified before. The level of investment for nuclear, pulverized coal and IGCC is uniform across solutions and the higher penetration of renewables reduces the investment on combustion turbines and combined cycle units. This occurs because nuclear and coal-based technologies are less expensive than natural gas-based units when investment and

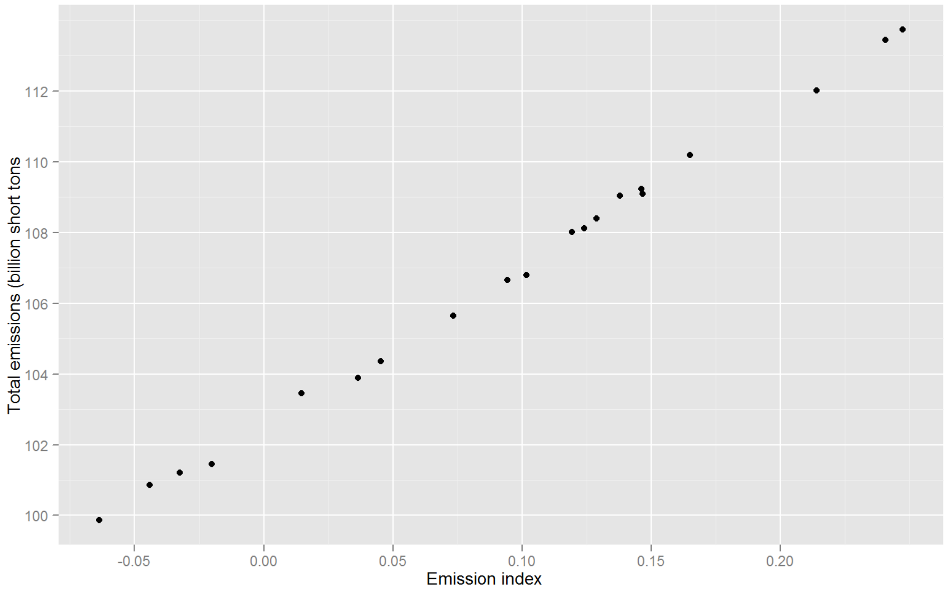


Figure 5.44: Emission index vs. total CO_2 emissions

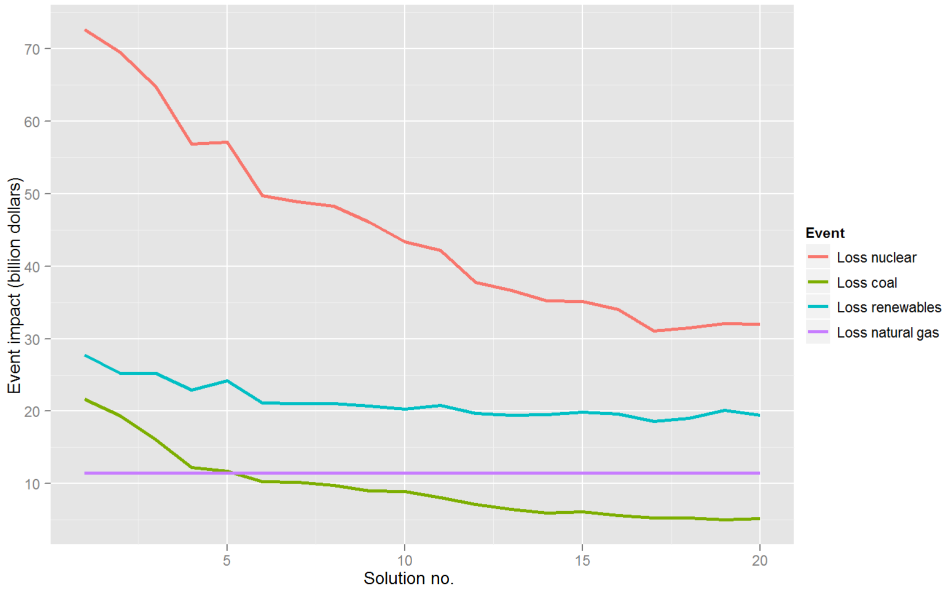


Figure 5.45: Event impacts for the best solutions

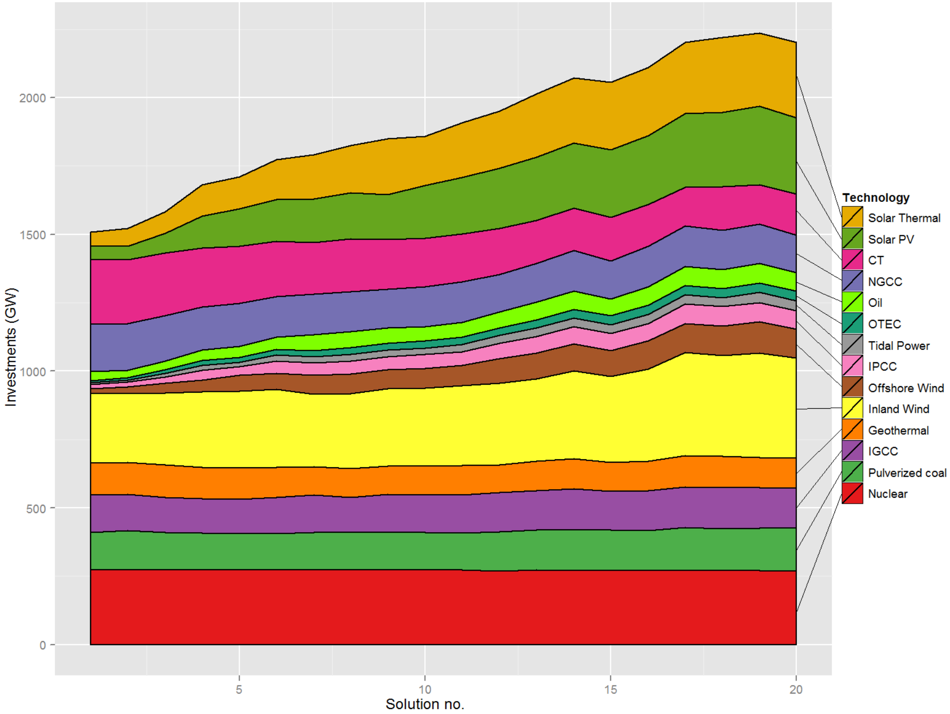


Figure 5.46: Electric generation investment for the best solutions

production costs are considered. Inclusion of the effects of renewable variability on the need for fast ramping units and on the cost of cycling thermal units would mitigate these effects.

Figure 5.47 depicts hybrid penetration with respect to the solutions ordered by increasing cost. The figure shows that investment for hybrid vehicles increases for higher levels of costs, along with renewable penetration and further reducing emissions.

To further investigate the behavior of the solutions in the Pareto front electric generation and investment, coal production and hybrid penetration graphs are included side by side in Figs. 5.48 and 5.49. The cases selected correspond to those with highest and smallest emission index and one selected in the middle of the Pareto front. The graphs on the left side represent the yearly investments for each type of generation technology. Unlike the base case in Fig. 5.4, the investment curves are not smooth and we can appreciate some high-frequency variation, which is a consequence of the optimizer not converging to the true Pareto front of solutions. If that was the case, we would see much more steady investment levels in the span of the simulation.

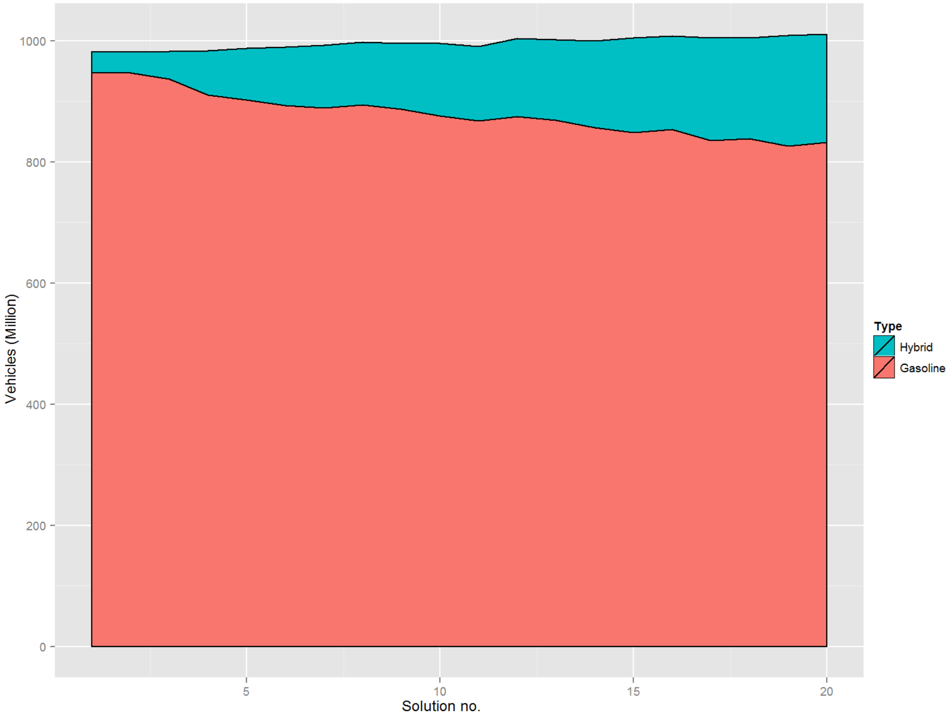
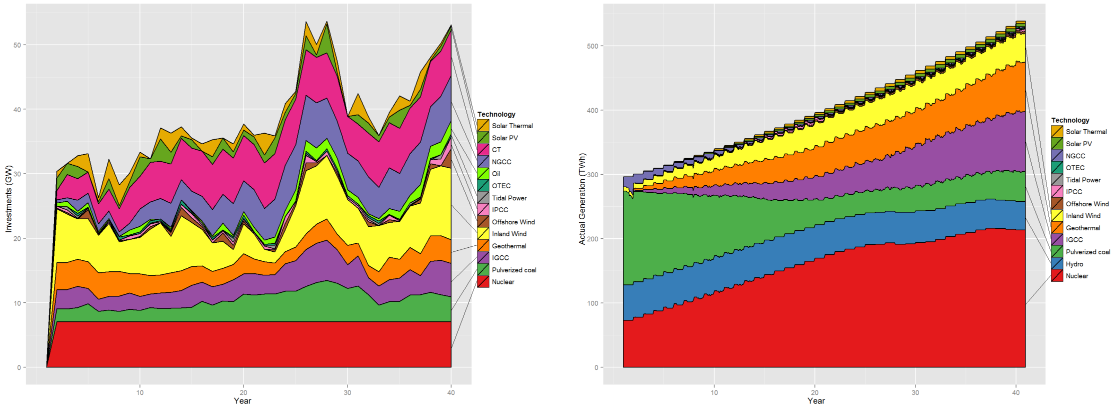
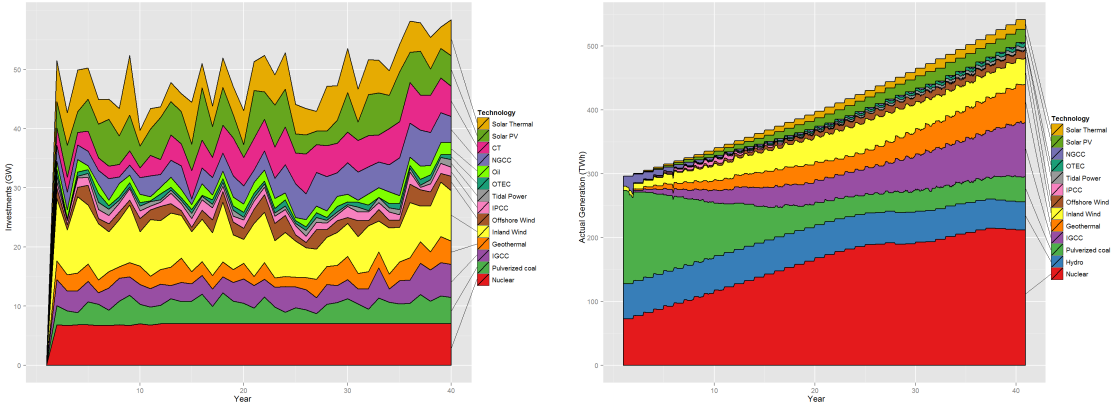


Figure 5.47: Vehicle investment for the best solutions

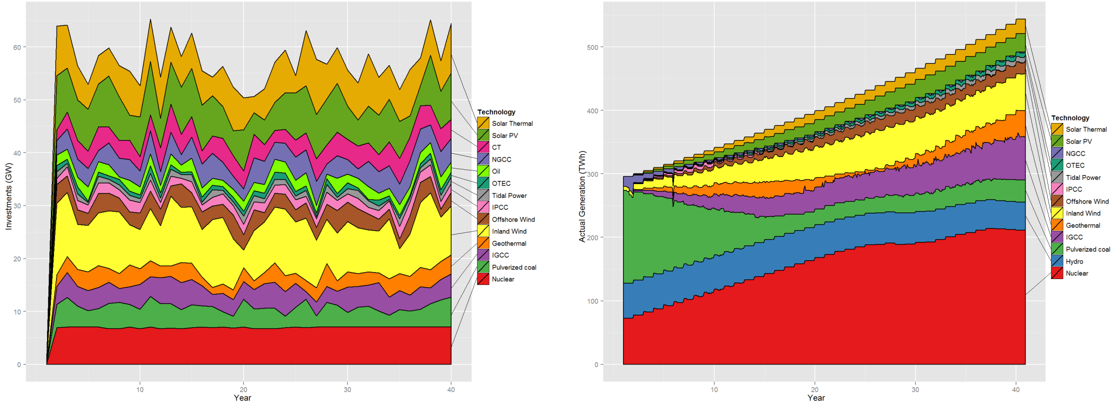
The solution with highest emissions behaves similarly to the minimum cost solution that was presented earlier. This solution includes some more renewable and hybrid investment than the base case and that is the reason for the higher cost and lower emissions. This trend is enhanced as we go towards solution with less emissions. As it was observed in earlier parallel plots, the investment on solar photovoltaic, thermal, off-shore wind and tidal power increases and so does the electricity produced from this sources. Coal production is reduced but less significantly than in the simulated cases with high coal prices or limits on emissions. Finally, hybrid penetration increases significantly across solutions. The solutions present a moderate change in energy production, according to Fig. 5.48. The differences would be more pronounced if we continued the optimization for a longer period of time, but constraints in time and resources prevented the full convergence to the real Pareto front.



(a) Electricity investment (High emissions) (b) Electric generation (High emissions)

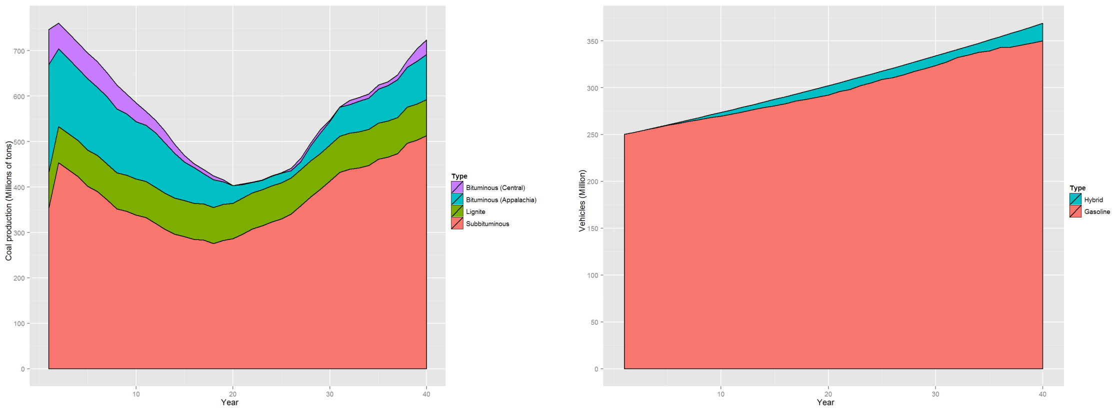


(c) Electricity investment (Medium emissions) (d) Electric generation (Medium emissions)

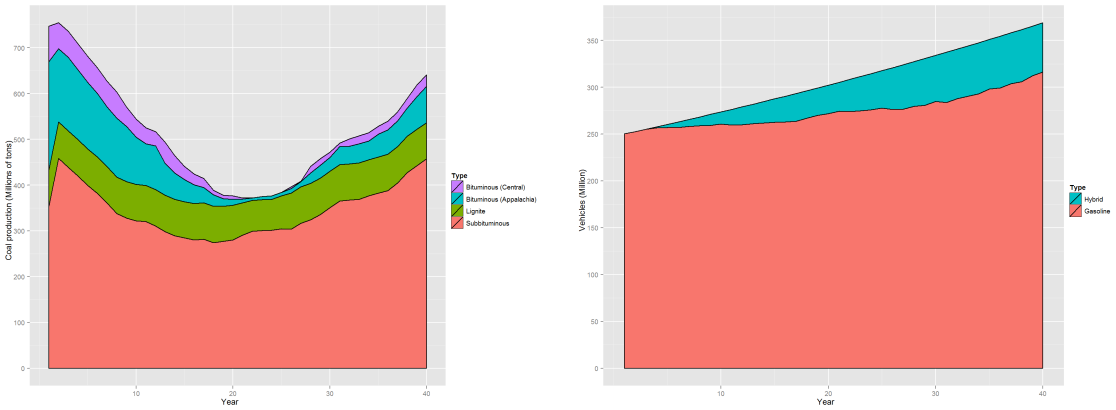


(e) Electricity investment (Low emissions) (f) Electric generation (Low emissions)

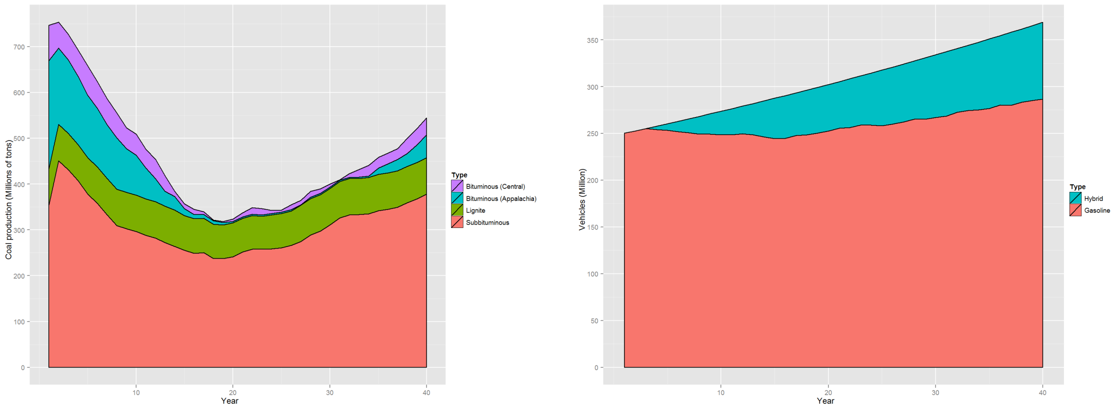
Figure 5.48: Electricity sector for different solutions



(a) Coal production (High emissions) (b) Passenger fleet (High emissions)



(c) Coal production (Medium emissions) (d) Passenger fleet (Medium emissions)



(e) Coal production (Low emissions) (f) Passenger fleet (Low emissions)

Figure 5.49: Coal production and hybrid penetration for different solutions

5.7 Sensitivity of the Pareto front

In this section the solutions in the Pareto front are tested with respect to the cases studied in Section 5.5. In all cases, the shape of the Pareto front is conserved and no solution becomes dominant.

The first plot shows the effect of increasing inflation and discount rates. Figure 5.50 shows that while emissions remain mainly unchanged, cost and resiliency are greatly affected, as they were in the minimum cost solution.

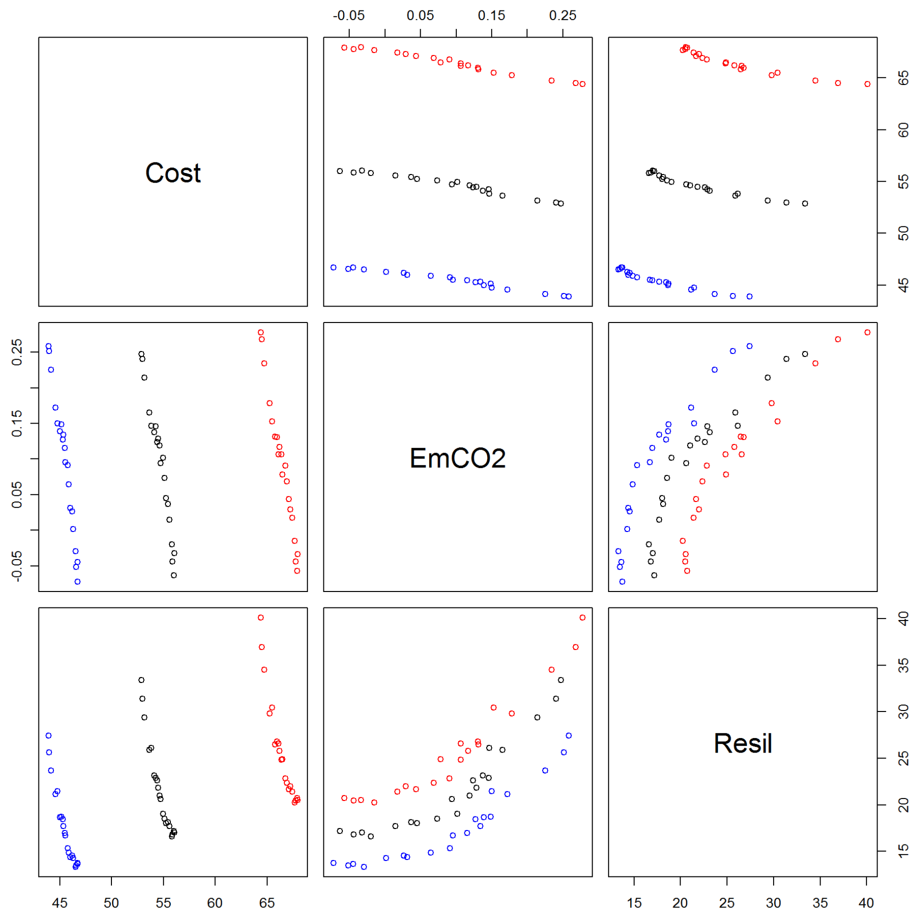


Figure 5.50: Pareto front with changes in inflation (red) and discount (blue) rates

Increase in coal production cost reduces emissions and increases the resiliency metrics. Costs increase slightly but the change is not as noticeable as with the other two metrics.

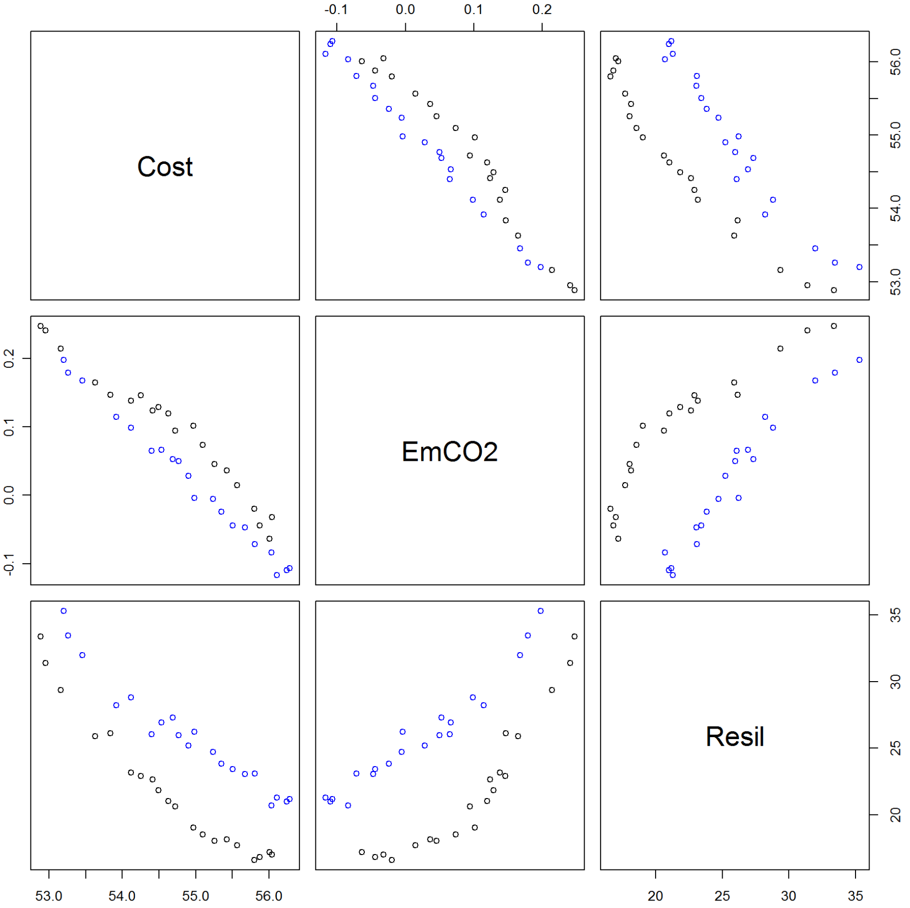


Figure 5.51: Pareto front with increased coal prices (blue)

Increase in gasoline prices (Fig. 5.52) present some unusual behaviors, like it did with the minimum cost solution. Results for the case with double increase in gasoline prices seem to have a reduction in overall cost and resiliency, while emissions decrease. The other two cases show a significant increase in cost and resiliency metrics while emissions are the highest of all cases considered.

Finally, Fig. 5.53 shows how the Pareto front solutions change when the limits for emissions are enforced. Costs increase and emissions decrease as the limits become more constraining. In the most severe case the resiliency metric becomes much larger in comparison to the rest of the cases.

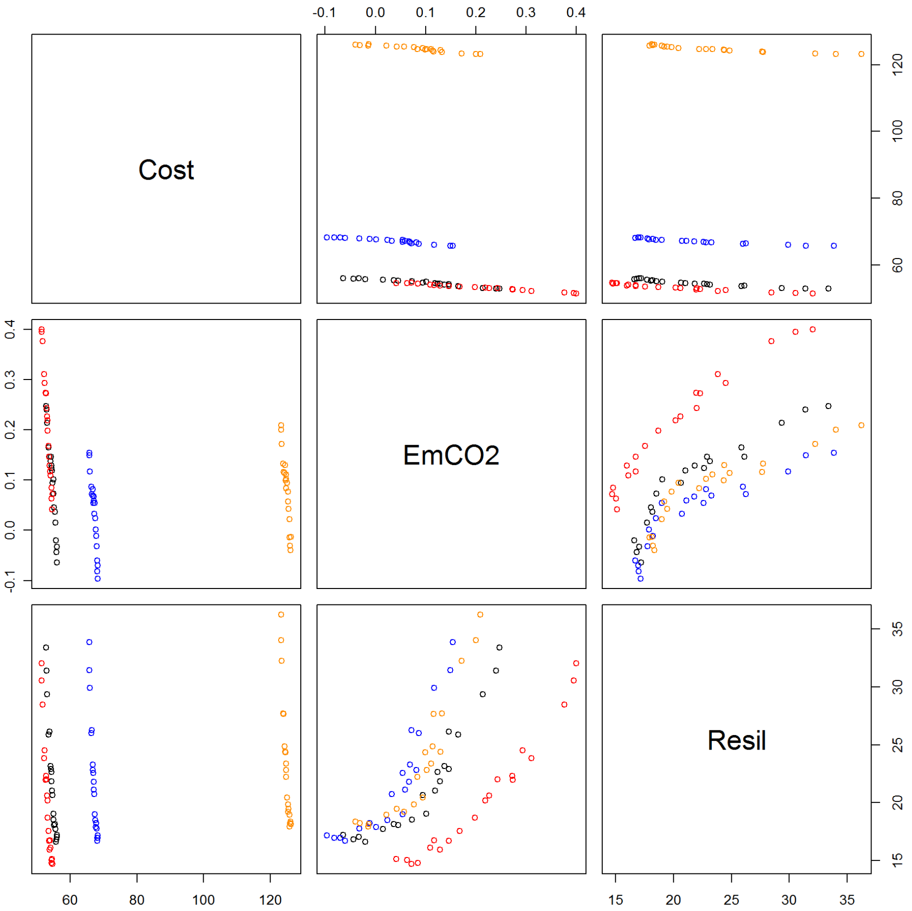


Figure 5.52: Pareto front with increases in gasoline prices (50% blue, 100% red, 300% orange)

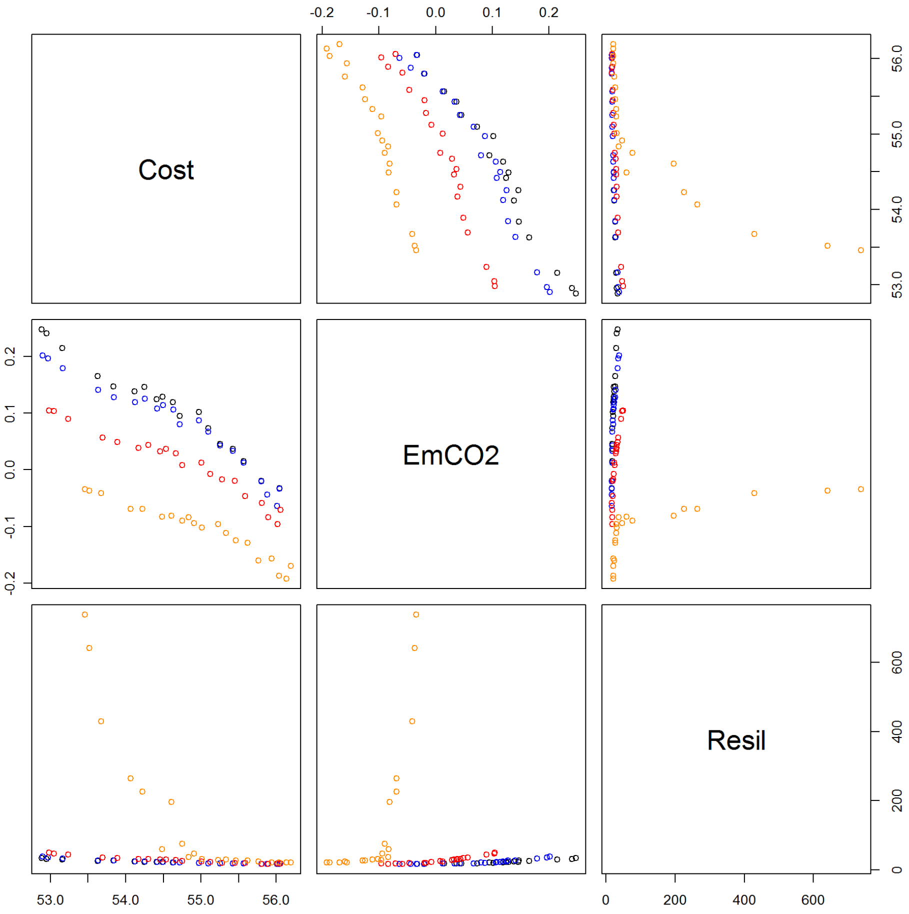


Figure 5.53: Pareto front with emissions reduction (0% blue, 20% red, 40% orange)

CHAPTER 6. CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

This work is motivated by the recognition that tools are lacking to jointly design the transportation and energy systems. Given the importance of these systems for the national economy and security, an integrating approach from the planning and operational points of view is essential in the near future. Combined optimization of energy and transportation investment portfolios would allow us to take advantage of potential interdependencies, which could significantly transform both systems in the next 40 years.

In this dissertation a modeling and software framework is proposed to address these issues, utilizing the notion of multiobjective optimization in an effective fashion. Metrics are used for cost, sustainability and resiliency (with a thorough development of the latter), and they are used to test candidate portfolios to eventually reach the Pareto front of solutions. A numerical example that models the energy, freight and passenger vehicular systems is presented. The optimization algorithms are applied to the example and the results thoroughly analyzed. Sensitivity analysis is used to validate the model and opportunities for further development and enhancement of the framework are identified.

We find ourselves in a very exciting point in time for the energy and transportation systems. Many renewable resources have reached maturity and will certainly become part of the future portfolios for energy generation. These technologies will not only reduce environmental impacts, but they will provide energy dependency and strengthen the system by introducing diversity. Nuclear is certainly a very economical option and will become part of future portfolios. Along these lines, these technologies will be most effective in the regions where they are most effective is best, e.g., Midwest for wind, or Southwest for solar and geothermal. Solid designs for a national high-voltage transmission system overlay will enable the development of the full potential of renewable technologies.

To achieve a significant reduction in greenhouse emissions it is imperative to shift away from coal-based generation, or at least to introduce new cleaner technologies combined with carbon sequestration. The evolution of the passenger vehicle fleet and the introduction of plug-in hybrids will not only further reduce emissions but it will reduce the increasing dependency on foreign oil.

Finally, NETPLAN has been shown to be an effective method to search among the space of portfolios. Once the parallelization deployment is complete, the software will be able to automatically and rapidly identify the best solutions, taking into account cost, sustainability and resiliency.

6.2 Contributions

The following is a summary of the contributions developed in this research:

Modeling transportation and energy This work presents a modeling framework to simultaneously design and optimize the energy and transportation systems. Models are developed for each system and the interactions between them are defined. At the operational level, fuel demand for transportation fleet is accounted for as well as the transportation of energy commodities through the transportation system. Investment interdependencies, such as competing right-of-way or resource scarcity, are also identified but not explicitly implemented in the framework.

Multi-objective optimization for planning A multiobjective optimizer is developed to find the best portfolios in terms of cost, resiliency and sustainability. It is based on the widely used NSGA-II algorithm, and it guides the multiobjective optimization by proposing minimum levels for investment decision variables. These are included in the minimum cost optimization problem, before selecting the full investment portfolio. Sustainability and resiliency emissions are applied to the results of this optimization and the metrics are returned to the NSGA-II selection process.

Decomposition of low-level optimization The minimum cost optimization for the energy and transportation systems is proved to be a very computational problem. Benders-like decomposition methods are successfully applied to this problem, reducing the size of the models and improving solution times for small examples. Implementation problems with external optimization libraries prevent these methods to be effective for the numerical example presented.

Parallelization framework A parallelization scheme for the evolutionary algorithm that directs the multiobjective optimization is presented. The algorithm is in the process of being deployed on a high-performance machine with 400 nodes, but it has been tested in a sequential fashion during the development of this work.

Resiliency metrics A new approach to resiliency evaluation for the energy and transportation systems has been presented. It allows the definition of events that are to be tested and the evaluation that these events have on the system operational cost. This method was applied to the numerical example with satisfactory results, showing that more diverse portfolios are more resilient.

Software design The implementation of the model is structured to allow easy development of new features thanks to a modular approach. Also, the model parameters and definition of networks is conceived separate from the source code, making it data-driven. Users can run the software without the need to understand or modify the code.

6.3 Publications

During the development of this work in the Ph.D. program I have produced the following list of publications:

1. E. Ibáñez, J. McCalley, D. Aliprantis, R. Brown, K. Gkritza, A. Somani, and L. Wang "National Energy and Transportation Systems: Interdependencies within a Long Term Planning Model," *Proc. of IEEE Energy 2030 Conference*, Atlanta, Georgia, Nov. 2008.

2. E. Ibáñez, K. Gkritza, J. McCalley, D. Aliprantis, R. Brown, A. Somani, and L. Wang "Interdependencies between Energy and Transportation Systems for National Long Term Planning," in "*Sustainable Infrastructure Systems: Simulation, Imaging, and Intelligent Engineering*," K. Gopalakrishnan, S. Peeta, editors, Springer-Verlag, Berlin, 2010, pp. 53-76.
3. J. McCalley, E. Ibáñez, Y. Gu, K. Gkritza, D. Aliprantis, L. Wang, A. Somani, and R. Brown "National Long-Term Investment Planning for Energy and Transportation Systems," *Proc. of IEEE PES General Meeting 2010*, Minneapolis, Minnesota, Jul. 2010.
4. E. Ibáñez, and J. McCalley, "Multiobjective evolutionary algorithm for long-term planning of the national energy and transportation systems," *Energy Systems Journal*, under review.
5. E. Ibáñez, "Going the extra mile: An Analysis of Gas Mileage on US vehicles," *American Statistical Association CHANCE*, under review.
6. E. Ibáñez, V. Krishnan, S. Lavrenz, D. Mejia, J. McCalley, and A. Somani, "Resiliency and robustness in long-term planning of the national energy and transportation system," *2nd RESIN Workshop*, Tucson, AZ, Jan. 2011.

The following are publications that are currently under development:

1. E. Ibanez, and J. McCalley, "A national model for integrated investment on the energy and transportation systems", to be submitted to *IEEE Trans. in Power.*
2. E. Ibanez, and L. Wang, "A Benders decomposition based branch-and-cut approach for linear programs with complementarity constraints", to be submitted to *Operations Research Letters.*
3. E. Ibanez, D. Cook, and J. McCalley, "Visualization and analysis of multiobjective solutions to the energy and transportation investment optimization problem", to be submitted to *International Journal of Critical Infrastructures.*

4. E. Ibáñez, V. Krishnan, S. Lavrenz, D. Mejia, J. McCalley, and A. Somani, "Resiliency and robustness in long-term planning of the national energy and transportation system," to be submitted to *Journal of Risk and Reliability*.

6.4 Directions of further research

There are many areas in this work that could be further expanded and would contribute to the improvement of the modeling efforts here presented. The most important ones are listed below:

Improve data quality and quantity As with any model, the quality of the results depends greatly on the quality of the data and the level of detail used. Revision of the presented data sources, use of more detailed networks, and smaller time steps will greatly benefit the model.

Include uncertainty Any long-term planning procedure faces the challenge of uncertainty in the data and projections. NETPLAN could be transformed from a deterministic to a stochastic formulation. Other options include the use of different scenarios in the resiliency evaluation.

Additional modeling Section 3.10 presented some areas of the model could be improved, including a better representation of hydro resources, the impact of high penetration of intermittent facilities (such as solar or wind), or consideration of electricity storage to dampen the reserve needs.

Further study of investment interdependencies The competition of transportation and energy investments for resources (right-of-way, materials) should be studied and relevant cases should be incorporated in the model.

Passenger formulation This work discussed the need for the passenger formulation to be further developed, and deployed in NETPLAN. Some of the issues that cannot be represented with the current model include: individual decisions, travel patterns, or passenger preferences. Air travel presents a special challenge since its routes cannot be easily linked to those for land transportation.

More sustainability metrics Using the present software, more sustainability metrics could be utilized in the national model beyond carbon emissions. These metrics would include emissions for SO_2, NO_X, particulate matter or VOCs, along with water and land demand.

Expand resiliency The resiliency methodology here presented could be further developed to include unfavorable events with changes in parameters other than capacity, such as cost or efficiency. Events could be defined for the transportation infrastructure too. Metrics based on other consequences, such as changes in nodal prices or recovery times, could be researched and implemented. Another possibility would be allowing the system to reconfigure itself after a significant event, allowing the re-optimization of investment portfolio immediately after the event.

Computational enhancements The solution of the models discussed in this work is computationally demanding. New strategies could be implemented to reduce solution times, starting with finalizing the implementation of the proposed parallel NSGA-II algorithm. The process could be modified to allow the master node to continuously fill a queue of candidates. Once enough results are returned from the worker nodes, the master would update its current best solution, thus eliminating the need of keeping separate instances of the NSGA-II algorithm running. A parallel implementation of the Benders decomposition could be performed too.

APPENDIX A. GLOSSARY

ATC Area transfer capability

Btu British Thermal Unit, English unit for energy

cf Cubic feet

CNV California-Southern Nevada Power Area

CSV Comma-separated value, file format used for NETPLAN data input

CT Combustion turbine

DC Direct current

DOE United States Department of Energy

ECAR East Central Area Reliability Coordination Agreement Region

EIA United States Energy Information Administration

ERCOT Electric Reliability Council of Texas

FL Florida subregion of SERC

GHG Green-house gases

GW Gigawatt, unit of power corresponding to one thousand Megawatts

GWh Gigawatt-hour, unit of energy corresponding to one thousand Megawatt-hours

IGCC Integrated gasification combined cycle

IPCC Integrated pyrolysis combined cycle

kW Kilowatt, unit of power corresponding to one thousand Watts

kWh Kilowatt-hour, unit of energy corresponding to one thousand Watt-hours

LMP Locational marginal price

LNG Liquefied natural gas

MAAC Mid-Atlantic Area Council

MBtu One thousand Btu

MAIN Mid-America Interconnected Network

MAPP Mid-Continent Area Power Pool

Metric ton (or tonne) One thousand kilograms

MMcf Million cubic feet

MMBtu Mega-Btu, or one million Btu

MPS Mathematical Programming System, file format for presenting and archiving linear programming

MW Megawatt, unit of power corresponding to one million Watts

MWh Megawatt-hour, unit of energy corresponding to one million Watt-hours

NE New England Power Pool

NEMS National Energy Modeling System

NERC North America Electric Reliability Corporation

NETPLAN National Energy and Transportation Planning Tool

NETSCORE-21 The 21st Century National Energy and Transportation Infrastructures Balancing Sustainability, Costs, and Resiliency, research project at Iowa State University

NGCC Natural gas combined cycle

NO_X Nitrogen oxides

NWP Northwest Pool subregion of the Western Systems Coordinating Council

NY New York Power Pool

O&M Operation and maintenance

OTEC Oceanic thermal energy conversion

PC Pulverized coal

PHEV Plug-in hybrid electric vehicle

PM Particulate matter

PV (Solar) Photovoltaic

RA Rocky Mountain and Arizona-New Mexico Power Areas

SERC Southeastern Electric Reliability Council

Short ton (or ton) English unit of weight equivalent to 2000 pounds

SO_2 Sulfur dioxide

SPP Southwest Power Pool

STV Southeastern Electric Reliability Council (excluding Florida)

Tcf Trillion cubic feet

TWh Terawatt-hour, one billion kWh

VOC Volatile organic compound

APPENDIX B. NETPLAN USER MANUAL

Introduction

This document serves as an introduction to the use of NETPLAN, including how to download, access and execute the software. It is intended to cover the basic ideas from a user point of view. For more information you may visit the NETPLAN website[1]. A file (`cplexdemo.zip`) that covers some of the basics of C++ and CPLEX has been included for future developers of the program.

Whenever possible, the connection between then model and the implementation for variables and parameters is included.

Setting up

Please download the code corresponding to latest version of NETPLAN from the NETPLAN website. Extract its content to a folder called `netplan` in your ISU "My documents". We will use these files throughout this document.

You may download one of the example "data" folders or create your own with the help of this manual instructions.

Accesing CPLEX

CPLEX resides on several ISU servers, which are listed in the following URL, under the Linux categories: `http://it.eng.iastate.edu/remote`. To access any of them, you will need a telnet facility. I suggest using PuTTY, which is a free implementation of Telnet and SSH for

---

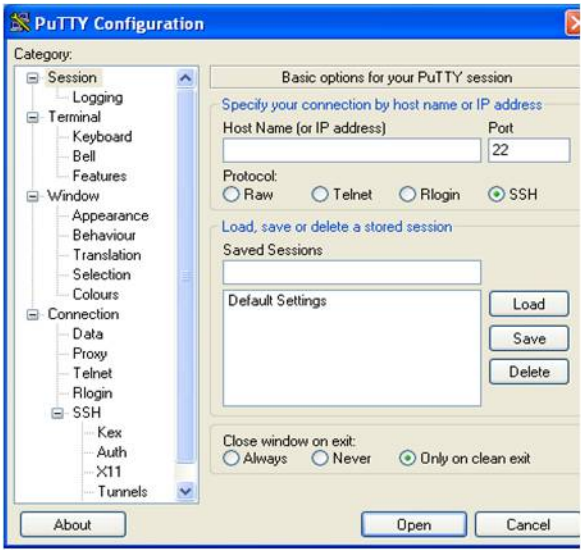
[1] `http://github.com/eibanez/NETPLAN`

Win32 and Unix platforms, along with an xterm terminal emulator.[2]

The following steps are necessary to connect to those servers:

1. Download PuTTY from `http://clue.eng.iastate.edu/downloads.shtml`. Alternatively, the official PuTTY page is `http://www.chiark.greenend.org.uk/~sgtatham/putty/`. At this page, you will find some alternatives; I used (successfully) the installer, `putty-0.60-installer.exe`.

2. Run PuTTY and get the window shown in Fig. B.1. Input the name of the server in the Host Name, e.g., `linux-8.ece.iastate.edu`.



Basic options for your PuTTY session

Specify your connection by host name or IP address

Host Name (or IP address)

Port: 22

Protocol: ○ Raw ○ Telnet ○ Rlogin ● SSH

Load, save or delete a stored session

Saved Sessions

Default Settings

Close window on exit: ○ Always ○ Never ● Only on clean exit

Figure B.1: Running PuTTy

3. Use your ISU username and password to log in. You will find yourself on a UNIX terminal emulator.

Setting the default folder

Once you are logged in the server, use the following commands to access your `netplan` folder that you previously created:

---

[2] This section has been adapted from Dr. James McCalley's class notes

```
cd YOUR_USERNAME

cd My\ Documents

cd netplan
```

NETPLAN folder

When you open the `netplan` folder you will find these folders and files:

- `data`: Folder that stores the model that is being tested
- `data/events`: Folder with the definition of events for resiliency analysis
- `src`: Folder with NETPLAN source code
- `src/nsga2`: Folder with NSGA-II implementation (includes code under development for NSGA-II parallelization)
- `CHANGELOG`: Text file to keep track of latest changes
- `Makefile`: File used to simplify the compilation procedure

During the execution, two more folders called `prepdata` and `nsgadata` will be created. The first one is used to store auxiliary files used during the optimization and most of the results. The latter will contain the results from the NSGA-II multiobjective optimizations.

The source code in the `src` folder is divided into different stages, as presented in Fig. 3.5. The main files (programs that you run) are the following:

- `preprocessor.cpp` (stage 1):Takes information from `data` folder and creates MPS and temporary files
- `postprocessor.cpp`: Takes MPS files and solves problem, writes solution in files (for checking purposes)
- `nsga2/main.cpp` (stage 2): Main file for the NSGA-II implementation, which takes the MPS and auxiliary files and solve the multiobjective problem

- `postnsga.cpp` (stage 3): Reads the individuals that form the Pareto front of solution and reports the solutions.

The rest of the files are considered libraries and contain auxiliary code for assist the programs above:

- `read.cpp`: Functions to read files (global parameters, networks, properties, etc.)
- `write.cpp`: Writing files (solutions, temporary files, etc.)
- `global.cpp`: Common global definitions
- `netscore.h`: Common tasks and misc.
- `node.cpp`: Declares a special class to store node information, along with functions to read, modify and write node data
- `arc.cpp`: Similar but for arcs
- `step.cpp`: Functions related with time and time steps
- `index.cpp`: Special variables and functions to store order of arcs and nodes (similar to the a vector index, hence the name)
- `solver.cpp`: Functions to solve a problem (includes Benders decompositions)

Basic commands

This is a list of the basic commands that are used to compile and execute the different functions in NETPLAN:

- `make`: Compiles the source code
- `./prep` (stage 1): Runs the preprocessor, taking the information in the `data` folder and creating the optimization model
- `./post`: Evaluates the minimum cost problem after it has been created

- `./nsga2` (stage 2): Runs the full multiobjective optimization
- `./postnsga` (stage 3): Evaluates the Pareto front solutions
- `make clean`: Eliminates the compiled programs (useful to clean the folder back to the original state)
- `make cleanmps`: Eliminates results and auxiliary files

Constructing a model

This will be a brief description of how to construct a functioning model. There is a small example in the website that might be helpful to visualize how it is done. References to the different formulation parameters are included.

All the data manipulation is done inside the `data` folder. The following subsections will explain the content of the different files, which use the CSV[3] format. It is recommended using plain-text editors (such as Notepad on Windows) to modify `parameters.csv`. Any other file can be edited with Notepad or Excel 2007 or newer.

parameters.csv

This is the most important file where the global parameters of the optimization are defined. It consists of two columns without a header. The first column is used for keywords and the second for values. The following is a list of the keywords, acceptable inputs or ranges and default values. The last few parameters are specific to NSGA-II.

Some values can be defined here and will be used as defaults. The philosophy is to take the most specific first and then try more relaxed definitions.

Parameters are listed as `keyword` [range] **default**

- `StepName` [letters] **None**: Definition of the key letters that will used to describe time steps. It is required and it must have one or more letters. E.g., “ym” would indicate that we are using years and months.

---

[3] Comma separated values

- `StepLength` [letters and numbers] **None**: Definition of the maximum length of the time steps (also required). E.g., “y2m12” indicates that simulation is 2 years long and each year has 12 months.
- `StepHours` [number] **1**: Defines the length in hours the last member in the time step (“m” in the example). If it used once all time steps will be the same. It can be also be used multiple time to define different time lengths, e.g., to describe different segments on the load duration curve.
- `DefStep` [letters as in StepName] **None**: Defines what is the default time step level for arcs and nodes. E.g., “ym”
- `DefDiscount` [0–1] **0**: Default discount rate (r in the formulation) that will be applied to all nodes and arcs, unless otherwise specified.
- `DefInflation` [0-1] **0**: Similar to discount but this reflects price inflation.
- `DefDemandRate` [number] **0**: Used to define the rate at which demand grows globally. E.g., “0.02” means that demands grow 2% annually by default.
- `UseDCFlow` [true/false] **false**: Use DC power flow equations.
- `CodeDC` [two letters] **None**: Define two letter code to identify nodes that use DC power flow. E.g., “EL”.
- `UseBenders` [true/false] **false**: Use Benders decomposition to solve minimum cost problem.
- `OutputLevel` [0–2] **2**: Level of output on screen (0 for most information).
- `TransStep` [letters as in StepName] **None**: Default transportation step. E.g., “y” means that all transportation is represented on an annual basis.
- `TransInfra` [letters] **—**: The first letter represents a new transportation infrastructure. The rest are the different modes that can use that infrastructure. E.g., “rt” adds infras-

tructure railroad and indicated that t (trains) can use railroad. This command should be used as many time as transportation infrastructures considered.

- `TransComm` [letters] —: The first letter defines a new commodity. The rest define what modes can transport the current commodity. E.g., "1t" indicates that commodity 1 (coal type 1) can travel by train. This command should be repeated for each commodity.
- `TransCoal` [letters] **None**: The letters correspond to the commodity codes that represent coal, i.e., the members of $\mathcal{K}_c$. The command is to be used once.
- `AddObj` [letters and numbers] —: This command adds a new sustainability objective. E.g., "emCO2" for CO_2 emissions. It can be repeated.
- `AddMetric` [letters and numbers] —: Similar to the above but it creates a metric that is not minimized, just reported.
- `NumberEvents` [integer] **0**: Declares the number of events to be used for resiliency. E.g., "2".
- `popsize` [multiple of 4] **20**: Population size for a generation (NSGA-II).
- `ngen` [number] **200**: Number of generations (NSGA-II).
- `pcross_real` [0–1] **0.75**: Crossover probability for real variables (NSGA-II).
- `pmut_real` [0–1] **0.4**: Mutation probability for real variables (NSGA-II).
- `eta_c` [number] **7**: Distribution index for real variable SBX crossover (NSGA-II).
- `eta_m` [number] **20**: Distribution index for real variable polynomial mutation (NSGA-II).
- `pcross_bin` [0–1] **0.4**: Crossover probability for binary variables (NSGA-II).
- `pmut_bin` [0–1] **0.5**: Mutation probability for binary variables (NSGA-II).
- `stages` [integer] **2**: Number of bits used to represent an investment (NSGA-II).

- `pstart` [0–1] **0.5**: Probability that each of the bits takes the value of 1 for the first randomly generation (NSGA-II).

node_List.csv

This file defines the set of energy nodes, $\mathcal{N}^E$. Nodes are named the following way: "CMIAy1m1", where the characters represent:

- First character is energy network: coal, electricity, etc.
- Second character is node type: Coal steam generator, storage, transmission
- Third and fourth are location: Iowa, Northwest Power Pool Area, etc.
- The rest refer to time step: month 1 of year 1, year 2, etc.

In this file we define the first four characters, i.e. network, type and location (e.g., "CMIA"). NETPLAN takes care of adding the time information. The first line of the file is reserved for the header. Starting on the second line, each line is occupied by the definition of a node. Of course, each node code needs to be unique. The first lines would go like the following (note that anything after the % is considered a comment and it's omitted).

```
code
% Coal production nodes
CPAL
CPAZ
CPAR
CPCO
...
```

arcs_List.csv

An arc is simple determined by the origin and destination nodes with a separation character, in the following fashion: "ETIAy1m1_ETMNy1m1". In this example we would be representing the electric transmission from Iowa to Minnesota for the first month of the first year.

Similar to the previous file, the first line is reserved for the header. After that, each line defines the four character code for origin and destination node in the first and second columns, respectively. For example:

```
from,to
% Coal production to transportation node,
CPAL,2TAL
CPAZ,4TAZ
CPAR,3TAR
...
```

trans_List.csv

This file introduces the transportation variables, $f_{(i,j,k,m)}(t)$: flow of transportation arc from node i to node j for commodity k using transportation mode m during time step t) are coded like this: "ttMONEy1_1TMONEy1":

- In the first half, the first two letters are the same and represent the transportation mode (train in this case)
- The next two characters represent the origin (here is Missouri)
- Destination code is defined in the next two characters (Nebraska in this case)
- The remaining characters before the underscore (_) represent time step (year 1)
- The second half begins with the commodity character (1 represents coal)
- The next character is always a "T" to represent transportation.
- The remaining characters are repeated from the first part (origin, destination, time).

NETPLAN uses the definition of commodities, infrastructures and modes from the global parameters to simplify the definition of the transportation network. This file starts with a line reserved for the header and then each line is composed by four columns:

- Origin two character code
- Destination two character code
- Distance in miles
- Characters to represent fleets allowed on that connection (if left blank, all fleets are used for that link)

This is an example:

```
from,to,mileage,fleet
AL,FL,591.5,
AL,GA,182,t
AL,MS,207.5,k
...
```

The transportation network is then coded into arcs and nodes, compatible with the energy network definition. In the example above a node called "1TMONEy1" would be created. It is this node that should be used to enter the transportation load for commodity "1" between "MO" and "NE".

Simplification for data storage

Before we continue, it is worth describing how data is stored in the rest of the files. This system is aimed at reducing the number of entries so that new cases can be modified quickly.

Let's say for example that we want to establish the efficiency of all transmission lines to 98%, for all locations and all time steps. In that case we could assign a value of 0.98 for efficiency using the following entry.

```
EL_EL   0.98
```

Using this simplified method we reduce the number of necessary data entries (and thus the chances of making errors) and it is much easier to modify later. The code I am preparing is capable of extending the simple above to all the arcs were it is required.

Now assume that we want to say that the wind generation units in Texas have a cost of 0.2, for all time steps. The necessary entry would be:

```
EWTX_ETTX   0.2
```

When dealing with time, I have been using the following scheme:

	const	y1	y2	...	y1m1	y1m2	...	y2m1	...
ET_ET	0.98	0.95		...		0.87	...		...

The above is a table of possible values of efficiencies for all electric transmission lines for different times. The way is setup allows minimizing the number of entries with respect to time steps. Note that there are empty cells in the table above. The steps to follow when you read the table are:

- Check if there is a value available in the smallest time step.
  - If such a value exists, take it. Example: 0.87 for y1m2.
  - If it doesn't, continue. Example: y1m1
- Check the following time step that includes the previous one:
  - If there is a value, take it. Example: 0.95 for y1
  - If there isn't, continue. Example: y2
- Continue until reaching the constant value.

In the case above, some values would be the following:

- y1m1 = 0.95
- y1m2 = 0.87
- y2m1 = 0.98

nodes_[Parameter].csv

All the remaining files that follow this nomenclature are used to define the rest of parameters in the model for nodes. This list presents these parameters, where the `keyword` is part of the file name (e.g., `nodes\_Demand.csv`):

- `Step` [characters in StepName] **DefStep**: Time step for a particular node
- `Demand` [number/X] **0**: Fixed demand at the node ($d_j^E(t)$ or $d_{(i,j,k)}^T(t)$). If "X" is used, demand equations 3.3 and 3.6 are not enforced.
- `DemandPower` [number/X] **X**: Definition of Demand in terms of power. If "X" is used, demand defaults to the previous parameter
- `DemandRate` [number] **DefDemandRate**: This parameter can be used to simulate a fixed annual increase in the demand.
- `CostUD` [number/X] **X**: If different than "X", a variable representing unserved demand is introduced and the cost is a penalty for not serving that demand.
- `DiscountRate` [number] **DefDiscountRate**: Discount rate to be applied to `CostUD`.
- `InflationRate` [number] **DefInflationRate**: Inflation rate to be applied to `CostUD`.
- `PeakPower` [number/X] **X**: If different than "X", constraints are created so that meeting peak power is enforced.
- `PeakPowerRate` [number] **0**: Annual rate to be applied to peak power.

arcs_[Parameter].csv

Similarly, the remaining parameters for arcs are defined here. This is a list, including the `keyword` that needs to be used in the file name.

- `OpCost` [number] **0**: Operational cost for an arc.

- `InvCost` [number/X] **X**: If different than "X", investment is allowed on the arcs, using this investment cost.
- `DiscountRate` [number] **DefDiscountRate**: Discount rate to be applied to operational and investment costs.
- `InflationRate` [number] **DefInflationRate**: Inflation rate to be applied to operational and investment costs.
- `Distance` [number/X] **X**: If different than "X", it becomes a multiplier of costs. Thus, the input of costs is in per mile terms.
- `Eff` [number] **1**: Arc efficiency
- `OpMin` [number] **0**: Minimum operational flow.
- `OpMax` [number/Inf] : Capacity at the arc due to existing infrastructure at time 0. No capacity is enforced if defined as "Inf".
- `InvMin` [number] **0**: Minimum investment on the arc.
- `InvMax` [number/Inf] **Inf**: Maximum investment on the arc.
- `InvStart` [time step] **First time step**: Defines when the first investment is allowed.
- `LifeSpan` [time step/X] **X**: If different than "X" investments have a predetermined life span.
- `Suscep` [number/X] **X**: If different than "X", the arc is endowed with susceptance and included in the DC power flow equations.
- `CapacityFactor` [number] **0**: If bigger than 0, the arc is taken into account to meet peak demand at the destination node.
- `OpSUST` [number] **0**: Coefficient to add the current arc to the calculation of a sustainability objective or metric. Substitute "SUST" by the appropriate code used in the parameter file.

arcs_TransEnergy.csv

This file defines the fuel consumption of fleet, $fuelCons_{(i,j,m)}(t)$, which creates extra demand on the energy side. The first row is reserved as a header. After that, one must write blocks of rows corresponding to the fuel consumption. Thus, a transportation link can create demand in multiple energy nodes.

The first column is matched to the first portion of the transportation arc. The second is used to define the four character code of the energy node to add demand. In the example below, the train (represented by the character `t`) creates a demand of 170.5 MMBtu per million ton-mile transported in nodes "DT01" and "DT02", that is diesel nodes at location "01" and "02".

```
from,to,const
tt0102,DT01,170.5 % MMBtu/MMton-mile
tt0102,DT02,170.5
kk0102,DT01,1678.5
kk0102,DT02,1678.5
```

events/CapacityLossN.csv

The files in the `events` folder are used to define the events to calculate resiliency metrics. Substitute "N" by the appropriate event number (from 1 up to the maximum defined in the global parameters) in the file name.

Within these files we define the fraction of capacity that is available due to a contingency. If not explicitly defined, NETPLAN will use the default of 1 and assume that the capacity was not lost. For example, the entry below indicates that pulverized coal generators are available at 70% during the second year.

```
from,to,y2
EC,,0.7
```

The software automatically identifies which are the operational years affected for each event in order to perform the minimum number of calculations that are necessary.

BIBLIOGRAPHY

[1] U.S. Energy Information Administration. (2008) Forecasts and analysis. [Online]. Available: http://www.eia.doe.gov/oiaf/forecasting.html

[2] U.S. Government. (2009, Feb.) United States Federal Budget for Fiscal Year 2010, A New Era of Responsibility: Renewing America's Promise. [Online]. Available: http://www.gpoaccess.gov/usbudget/fy10/

[3] J. McCalley, E. Ibanez, Y. Gu, K. Gkritza, D. Aliprantisa, L. Wang, A. Somani, and R. Brown, "National long-term investment planning for energy and transportation systems," in *Proc. of IEEE PES General Meeting 2010*, Minneapolis, MN, Jul. 2010.

[4] N. van Beeck, *Classification of Energy Models*, Tilburg University & Eindhoven University of Technology, 1999.

[5] U.S. Energy Information Administration. (2009) The National Energy Modeling System: An overview. [Online]. Available: http://www.eia.doe.gov/oiaf/aeo/overview/

[6] Energy Technology Systems Analysis Programme (ETSAP). MARKAL and TIMES model documentation. [Online]. Available: http://www.etsap.org/documentation.asp

[7] M. Lu and Z. D. andT. Saha, "A framework for transmission planning in a competitive electricity market," in *Proc. of the Transmission and Distribution Conf. and Exhibition: Asia and Pacific*, 2005, pp. 3705–3710.

[8] GE Power. MARS website. [Online]. Available: http://www.gepower.com/prod_serv/products/utility_software/software/en/ge_mars.htm

[9] Siemens PTI. TPLAN website. [Online]. Available: http://www.pti-us.com/pti/software/tplan/

[10] T. Tran, J. Kwon, J. Choi, D. Jeon, and K. Han, “Sensitivity analysis of probabilistic reliability evaluation of KEPCO system using TRELSS,” in *International Conf. on Probabilistic Methods Applied to Power Systems*, 2006.

[11] PCI Power Costs. GenTrader website. [Online]. Available: http://www.powercosts.com/products/gentrader.asp

[12] J. Bastian, J. Zhu, V. Banunarayanan, and R. Mukerji, “Forecasting energy prices in a competitive market,” in *IEEE Computer Applications in Power Magazine*, vol. 12, no. 3, Jul. 1999, pp. 40–45.

[13] V. Koritanov and T. Veselka, “Modeling the regional electricity network in Southeast Europe,” in *Power Eng. Society General Meeting*, vol. 1, Jul. 2003.

[14] Ventyx. PROMOD website. [Online]. Available: http://www1.ventyx.com/analytics/promod.asp

[15] ——. ProSym website. [Online]. Available: http://www.ventyx.com/analytics/market-analytics.asp

[16] Energy Exemplar. PLEXOS website. [Online]. Available: http://www.energyexemplar.com

[17] W. Head and H. Nguyen, “The procedure used to assess the long range generation and transmission resources in the Mid-Continent Area Power Pool,” *IEEE Trans. on Power Syst.*, vol. 5, Nov. 1990.

[18] H. Yamin, J. Finney, W. Rutz, and H. Holly, “GENCOs portfolio management using STRATEGIST in deregulated power markets,” in *Large Eng. Systems Conf. on Power Eng.*, Jul. 2001.

[19] Adica Consulting. WASP-IV model documentation. [Online]. Available: http://www.adica.com/wasp_iv.html

[20] U.S. Environmental Protection Agency. Integrated Planning Model (IPM). [Online]. Available: http://www.epa.gov/airmarkets/progsregs/epa-ipm/

[21] W. Short, N. Blair, P. Sullivan, and T. Mai. ReEDS model documentation: Base case data and model description. [Online]. Available: http://www.nrel.gov/analysis/reeds/pdfs/reeds_full_report.pdf

[22] U.S. Department of Transportation, Office of Freight Management and Operations. (2002) Freight analysis framework. Washington, D.C. [Online]. Available: http://www.ops.fhwa.dot.gov/freight/freight_analysis/faf/

[23] T. Crainic, M. Florian, J. Guelat, and H. Spiess, "Strategic Planning of Freight Transportation: STAN, an interactive-graphic system," *Transportation Research Record*, vol. 1283, pp. 97–124, 1990.

[24] Jack Faucett Associates. Highway economic requirements system - state version. [Online]. Available: http://www.fhwa.dot.gov/infrastructure/asstmgmt/hersfact.cfm

[25] HDR. RailDec, Rail Investment Decision Support Tool. [Online]. Available: http://www.hdrinc.com/15/14/1/default.aspx#raildec

[26] Transportation Research Board. (2006) Statewide travel forecasting models. [Online]. Available: http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_syn_358.pdf

[27] J. P. Sterbenz and D. Hutchison. (2006) ResiliNets: Mutlilevel Resilient and Survivable Networking Initiative. [Online]. Available: http://www.ittc.ku.edu/resilinets/

[28] U.S. Federal Communications Commission. (2010) Public Safety and Homeland Security Bureau. [Online]. Available: http://www.fcc.gov/pshs/

[29] P. V. de Carvalho, I. L. dos Santos, J. O. Gomes, M. R. da Silva Borges, and G. J. Huber, "The role of nuclear power plant operators' communications in providing resilience and

stability in system operation," in *2nd Symp. on Resilience Engineering*, Juan-les Pins, France, Nov. 2006.

[30] M. Christopher and C. Rutherford, "Creating supply chain resilience through agile Six Sigma," *CriticalEYE*, vol. 5, 2004.

[31] J. F. Castet and J. Saleh, "Survivability and resiliency of spacecraft and space-based networks: a framework for characterization and analysis," in *AIAA Space 2008 Conference and Exposition*, San Diego, California, 2008.

[32] S. Anders, D. Woods, R. Wears, S. Perry, and E. Patterson, "Limits on adaptation: Modeling resilience and brittleness in hospital emergency departments," in *2nd Symp. on Resilience Engineering*, Juan-les Pins, France, Nov. 2006.

[33] P. Kundur, J. Paserba, V. Ajjarapu, G. Anderson, A. Bose, C. Canizares, N. Hatziargyriou, D. Hill, A. Stankovic, C. Taylor, T. V. Cutsem, and V. Vittal, "Definition and classification of power system reliability," *IEEE Trans. on Power Syst.*, vol. 19, pp. 1387–1400, May 2004.

[34] L. Mili, "Taxonomy of the characteristics of power system operating states," in *2011 EFRI-RESIN Workshop*, Tucson, AZ, Jan. 2011.

[35] A. Quelhas, E. Gil, J. McCalley, and S. Ryan, "A multiperiod generalized network flow model of the U.S. integrated energy system: Part I - model description," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 829–836, May 2007.

[36] U.S. Department of Transportation, Research and Innovative Technology Administration, Bureau of Transportation Statistics. (1997) Commodity flow survey united states. [Online]. Available: http://www.bts.gov/publications/commodity_flow_survey/

[37] M. Bazaraa, J. Jarvis, and H. Sherali, *Linear Programming and Network Flows*, 2nd ed. New York: Willey, 1990, ch. 12, pp. 587–597.

[38] Environmental Protection Agency. 2002 national emissions inventory booklet. [Online]. Available: http://www.epa.gov/ttn/chief/net/2002neibooklet.pdf

[39] Energy Information Administration. (2007) Emissions of greenhouse gases in the United States 2006.

[40] E. Ibanez and J. McCalley, "Multiobjective evolutionary algorithm for long-term planning of the national energy and transportation systems," *Energy Syst. Journal*, under review.

[41] S. Fankhauser, "The social costs of greenhouse gas emissions: An expected value approach," *The Energy Journal*, vol. 15, no. 2, pp. 157–184, 1994.

[42] L. Marshall. (2009) Biofuels and the time value of carbon: Recommendations for GHG accounting protocols. [Online]. Available: http://www.wri.org/publications

[43] K. Richards, "The social costs of greenhouse gas emissions: An expected value approach," *Critical Reviews in Environmental Science and Technology*, vol. 27, no. 1, p. 279292, nov 1997.

[44] Resources for the Future. (2009) Comparison of major market-based climate change bills in the 111th congress. [Online]. Available: http://www.rff.org/wv/Documents/RFF_Major_domestic_bill_comparison_091002.pdf

[45] E. Ibanez, S. Lavrenz, D. Meja, J. D. McCalley, and A. K. Somani, "Resiliency and robustness in long-term planning of the national energy and transportation system," in *2011 EFRI-RESIN Workshop*, Tucson, AZ, Jan. 2011.

[46] E. Gil and J. McCalley, "A US energy system model for disruption analysis: Evaluating the effects of 2005 hurricanes," *IEEE Trans. on Power Syst.*, vol. PP, no. 99, Jan. 2011.

[47] North American Reliability Corporation Planning Standards. (2010) Standard TPL-002-0a – System performance following loss of a single BES element. [Online]. Available: http://www.nerc.com/files/TPL-002-0a.pdf

[48] S. Conseil. (2010) Impact of rising cost of uranium on electricity. [Online]. Available: http://energy.sia-partners.com/?p=172

[49] St. Charles County. (2010) About the new madrid fault. [Online]. Available: http://www.scchealth.org/docs/ems/docs/prepare/newmadrid.html

[50] U.S. Department of Energy. (2009) National electric transmission congestion study. [Online]. Available: http://www.congestion09.anl.gov/documents/docs/Congestion_Study_2009.pdf

[51] Bureau of Transportation Statistics. (2005) Freight facts and figures. [Online]. Available: http://ops.fhwa.dot.gov/freight/freight_analysis/nat_freight_stats/docs/05factsfigures/index.htm

[52] Asociation of American Railroads. (2007) National rail freight infrastructure capacity and investment study. [Online]. Available: http://www.aar.org/~/media/aar/Files/natl_freight_capacity_study.ashx

[53] E. Gil, "Integrated network flow model for a reliability assessment of the national electric energy system," Doctoral thesis, Iowa State University, Ames, Iowa, 2007.

[54] J. Salmeron, K. Wood, and R. Baldick, "Analysis of electric grid security under terrorist threat," *IEEE Trans. Power Syst.*, vol. 19, no. 2, pp. 905–912, May 2004.

[55] E. Ibanez, K. Gkritza, J. McCalley, D. Aliprantis, R. Brown, A. Somani, and L. Wang, "Interdependencies between energy and transportation systems for national long term planning," in *Sustainable Infrastructure Systems: Simulation, Imaging, and Intelligent Engineering*. Berlin: Springer-Verlag, 2010, pp. 53–76.

[56] A. M. Geoffrion, "Generalized Benders' decomposition," *J. Optim. Theory Appl.*, vol. 10, pp. 237–260, 1972.

[57] J. Bloom, "Solving an electricity generating capacity expansion planning problem by

generalized Benders' decomposition," *Operations Research*, vol. 31, no. 1, pp. 84–100, Jan./Feb. 1983.

[58] C. Coello and G. Lamont, *Applications of Multi-Objective Evolutionary Algorithms*, 1st ed. New Jersey: World Scientific, 2004.

[59] L. Chambers, *The Practical Handbook of Genetic Algorithms Applications*, 2nd ed. Boca Raton, CA: Chapman & Hall, 2001.

[60] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, 2002.

[61] IBM ILOG, *CPLEX Optimization Studio*, 2010, Version 12.2. [Online]. Available: http://publib.boulder.ibm.com/infocenter/cosinfoc/v12r2/

[62] American Institute of Aeronautics and Astronautics. (1998) Guide for the verification and validation of computational fluid dynamics simulations. [Online]. Available: http://www.aiaa.org/content.cfm?pageid=363&id=629

[63] M. Milligan and K. Porter. (2008) Determining the capacity value of wind: An updated survey of methods and implementation. [Online]. Available: http://www.nrel.gov/wind/pdfs/43433.pdf

[64] Y. Wang and J. Liao. (2008) Analyses of wind energy impact on wfec system operations. [Online]. Available: http://www.nrel.gov/wind/pdfs/39477.pdf

[65] X. Wang and J. McDonald, *Modern power system planning*, 3rd ed. London: McGraw-Hill, 1994.

[66] A. Quelhas, E. Gil, and J. McCalley, "Nodal prices in an integrated energy system: A generalized network flow model," *Int. J. of Critical Inf.*, vol. 2, no. 1, pp. 50–69, May 2006.

[67] A. Quelhas and J. McCalley, “A multiperiod generalized network flow model of the U.S. integrated energy system: Part II - simulation results,” *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 837–844, May 2007.

[68] U.S. Energy Information Administration. (2007) Coal transportation: Rates and trends. [Online]. Available: http://www.eia.doe.gov/cneaf/coal/page/trans/ratesntrends.html

[69] ——. (2007) Form FERC-423 database: Monthly cost and quality of fuels for electric plants data. [Online]. Available: http://www.eia.doe.gov/cneaf/electricity/page/ferc423.html

[70] ——. (2008) Annual coal report. [Online]. Available: http://www.eia.doe.gov/cneaf/coal/page/acr/acr_sum.html

[71] ——. (2008) Productive capacity of coal mines by state. [Online]. Available: http://www.eia.doe.gov/cneaf/coal/page/acr/table11.html

[72] ——. (2008) Average sales price of U.S. coal by state and disposition. [Online]. Available: http://www.eia.doe.gov/cneaf/coal/page/acr/table33.html

[73] ——. (2008) U.S. coal consumption by end-use sector. [Online]. Available: http://www.eia.doe.gov/cneaf/coal/quarterly/html/t32p01p1.html

[74] ——. (2008) U.S. coal imports. [Online]. Available: http://www.eia.doe.gov/cneaf/coal/quarterly/html/t18p01p1.html

[75] ——. (2008) U.S. coal exports. [Online]. Available: http://www.eia.doe.gov/cneaf/coal/quarterly/html/t7p01p1.html

[76] ——. (2009) Distribution and production of oil and gas wells by state. [Online]. Available: http://www.eia.doe.gov/pub/oil_gas/petrosystem/petrosysog.html

[77] ——. (2009) Natural gas wellhead value and marketed production. [Online]. Available: http://www.eia.doe.gov/dnav/ng/ng_prod_whv_a_EPG0_FWA_dpmcf_a.htm

[78] ——. (2009) Natural gas summary. [Online]. Available: http://www.eia.gov/dnav/ng/ng_sum_lsum_dcu_nus_m.htm

[79] ——. (2009) About U.S. natural gas pipelines. [Online]. Available: http://www.eia.doe.gov/pub/oil_gas/natural_gas/analysis_publications/ngpipeline/index.html

[80] ——. (2009) Impact of the 2008 hurricanes on the natural gas industry. [Online]. Available: http://www.eia.doe.gov/pub/oil_gas/natural_gas/feature_articles/2009/nghurricanes08/nghurricanes08.pdf

[81] ——. (2009) U.S. natural gas imports by country. [Online]. Available: http://www.eia.gov/dnav/ng/ng_move_impc_s1_m.htm

[82] ——. (2009) U.S. natural gas imports by point of entry. [Online]. Available: http://www.eia.gov/dnav/ng/ng_move_poe1_a_EPG0_PRP_DpMcf_a.htm

[83] ——. (2009) State to state natural gas pipeline capacities and flows. [Online]. Available: http://www.eia.gov/pub/oil_gas/natural_gas/analysis_publications/ngpipeline/StatetoState.xls

[84] ——. (200) The basics of underground natural gas storage. [Online]. Available: http://www.eia.doe.gov/pub/oil_gas/natural_gas/analysis_publications/storagebasics/storagebasics.pdf

[85] ——. (2009) EIA-191A field level underground natural gas storage data. [Online]. Available: http://www.eia.gov/pub/oil_gas/natural_gas/data_publications/natural_gas_annual/current/data/eia191a.xls

[86] ——. (2009) Underground natural gas storage capacity. [Online]. Available: http://www.eia.gov/dnav/ng/ng_stor_cap_a_EPG0_SAC_Mmcf_m.htm

[87] ——. (2009) Natural gas consumption by end use. [Online]. Available: http://www.eia.gov/dnav/ng/ng_cons_sum_a_EPG0_vgt_mmcf_a.htm

[88] ——. (2009) Natural gas frequently asked questions. [Online]. Available: http://www.eia.doe.gov/ask/ng_faqs.asp

[89] ——. (2009) U.S. natural gas exports by country. [Online]. Available: http://www.eia.gov/dnav/ng/ng_move_expc_s1_m.htm

[90] ——. (2009) State data for consumption - retail sales of electricity by state by sector by provider, 1990-2008. [Online]. Available: http://www.eia.doe.gov/cneaf/electricity/epa/sales_state.xls

[91] ——. (2010) International energy statistics: Total electricity net imports. [Online]. Available: http://tonto.eia.gov/cfapps/ipdbproject/IEDIndex3.cfm?tid=2&pid=2&aid=23

[92] J. Gifford, R. C. Brown, and J. Voss, "Characterization of power production technologies," *Handbook of Data Mining for Power Systems*, under review.

[93] National Renewable Energy Laboratory. (2010) Wind resource assessment. [Online]. Available: http://www.nrel.gov/wind/resource_assessment.html

[94] U.S. Energy Information Administration. (2010) Form EIA-860: Annual electric generator report. [Online]. Available: http://www.eia.doe.gov/cneaf/electricity/page/eia860.html

[95] U.S. Environmental Protection Agency. (2010) U.S. imports by country of origin. [Online]. Available: http://www.eia.gov/dnav/pet/pet_move_impcus_a2_nus_ep00_im0_mbbl_a.htm

[96] ——. (2010) Oil supply and disposition. [Online]. Available: http://www.eia.gov/dnav/pet/PET_SUM_SND_D_NUS_MBBL_A_CUR.htm

[97] ——. (2010) World crude oil prices. [Online]. Available: http://www.eia.gov/dnav/pet/pet_pri_wco_k_w.htm

[98] ——. (2010) Weekly retail gasoline and diesel prices. [Online]. Available: http://www.eia.gov/dnav/pet/pet_pri_gnd_dcus_nus_w.htm

[99] ——. (2010) Factors affecting gasoline prices. [Online]. Available: http://www.eia.doe.gov/energyexplained/index.cfm?page=gasoline_factors_affecting_price

[100] ——. (2010) NEMS: Petroleum market module. [Online]. Available: http://www.eia.doe.gov/oiaf/aeo/assumption/pdf/petroleum.pdf

[101] ——. (2010) Emission facts: Average carbon dioxide emissions resulting from gasoline and diesel fuel. [Online]. Available: http://www.epa.gov/oms/climate/420f05001.htm

[102] D. T. Lang, *XML: Tools for parsing and generating XML within R and S-Plus.*, 2010, R package version 3.2-0. [Online]. Available: http://CRAN.R-project.org/package=XML

[103] Google, Inc. (2010) Google Maps API. [Online]. Available: http://code.google.com/apis/maps/

[104] Rand McNally and Company, *Rand McNally Handy Railroad Atlas of the United States: Railroad Model Craftsman Edition*, 1973.

[105] MapQuest. (2010) MapQuest Maps. [Online]. Available: http://www.mapquest.com/

[106] E. Schwarm and A. Little, *Capital and Maintenance Costs for Fixed Railroad Electrification Facilities.* Transportation Research Board National Academy of Sciences, 1977, pp. 42–47.

[107] Association of American Railroads. (2010) Railroad facts. [Online]. Available: http://www.aar.org/NewsAndEvents/Press-Releases/2010/12/22-railfacts.aspx

[108] U.S. Bureau of Transportation Statistics. (2010) Energy consumption by mode of transportation. [Online]. Available: http://www.bts.gov/publications/national_transportation_statistics/html/table_04_06.html

[109] U.S. Department of Energy. (2010) Transportation energy databook. [Online]. Available: http://cta.ornl.gov/data/

[110] U.S. Bureau of Transportation Statistics. (2010) Average fuel efficiency of U.S. passenger cars and light trucks. [Online]. Available: http://www.bts.gov/publications/national_transportation_statistics/html/table_04_23.html

[111] U.S. Department of Energy. (2008) Average price of a new car, 1970-2006. [Online]. Available: http://www1.eere.energy.gov/vehiclesandfuels/facts/2008_fotw520.html

[112] A. Simpson, "Cost-benefit analysis of plug-in hybrid electric vehicle technology," in *22nd International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium and Exhibition*, Yokohama, Japan, Oct.23–28 2006. [Online]. Available: http://www.nrel.gov/vehiclesandfuels/vsa/pdfs/40485.pdf

[113] D. Wu, D. Aliprantis, and K. Gkritza, "Electric energy and power consumption by light-duty plug-in electric vehicles," *IEEE Trans. on Power Syst.*, vol. PP, no. 99, Jul. 2010.

[114] IBM developerWorks forum. (2010) "env.end() does not free up all memory in use". [Online]. Available: http://www.ibm.com/developerworks/forums/thread.jspa?threadID=351213

[115] U.S. Energy Information Administration. (2009) U.S. coal consumption by end-use sector. [Online]. Available: http://www.eia.doe.gov/cneaf/coal/quarterly/html/t32p01p1.html

[116] U.S. Bureau of Transportation Statistics. (2009) Domestic demand for gasoline by mode. [Online]. Available: http://www.bts.gov/publications/national_transportation_statistics/html/table_04_07.html

[117] ——. (2009) Fuel consumption by mode of transportation in physical units. [Online]. Available: http://www.bts.gov/publications/national_transportation_statistics/html/table_04_05.html

[118] U.S. Environmental Protection Agency. (2010) U.S. greenhouse gas inventory report. [Online]. Available: http://www.epa.gov/climatechange/emissions/usinventoryreport.html

[119] U.S. Energy Information Administration. (2009) Net generation by energy source. [Online]. Available: http://www.eia.doe.gov/cneaf/electricity/epm/table1_1.html

[120] A. Quelhas, "Economic efficiencies of the energy flows from the primary resource suppliers to the electric load centers," Doctoral thesis, Iowa State University, Ames, Iowa, 2006.

[121] U.S. Energy Information Administration. (2009) Average retail price of electricity to ultimate customers: Total by end-use sector. [Online]. Available: http://www.eia.doe.gov/cneaf/electricity/epm/table5_3.html

[122] ——. (2011) Annual energy outlook. [Online]. Available: http://www.eia.doe.gov/oiaf/aeo/aeoref_tab.html

[123] U.S. Department of Energy. (2008, may) 20% wind energy by 2030 - increasing wind energy's contribution to U.S. electricity supply. [Online]. Available: http://www.20percentwind.org/20p.aspx?page=Report