Investment and Risk Management in Electrical Infrastructure

Oliver Wright

Supervisor: Dr. Daniel Kuhn
Second Marker: Professor Berç Rustem

Department of Computing,
Imperial College London,
United Kingdom,
SW7 2AZ

oliver.wright06@imperial.ac.uk

17/June/2010

Abstract

This project explores the problems faced by an independent system operator, when investing in new electrical infrastructure. The expansion of power generation and transmission capacities must be made to ensure the power system remains resilient under uncertain demand and economic conditions. Recently, higher emphasis has been placed on the environmental impact of new power plants, as energy companies are under increased political pressures. A multi-stage stochastic programming model is proposed that allows for investment in new wind and solar farms, and is suitable under an emissions trading scheme. In order to obtain a computationally tractable problem a linear decision rule approximation is applied to both the primal and dual problems which results in the solution of a tractable linear program. A case study of the electricity system in El Salvador is undertaken to assess possible investment opportunities and to evaluate the model. The results highlight a mix of renewable resources as the best option to facilitate rising demand for electricity, likely increases in fossil-fuel prices and the need for emission-free energy.

Acknowledgments

I would like to pay special thanks to my supervisor Dr. Daniel Kuhn for his continued support throughout this project. Your enthusiasm for this subject has kept me motivated and has introduced me to a topic which I have found both interesting and enjoyable.

I would also like to thank two very dedicated and enthusiastic Ph.D students, Wolfram Wiesemann and Angelos Georghiou. I do not think I could have overcome some of the difficulties I have faced in this project without your help. Thank you to Professor Berç Rustem for always having confidence in me and for his useful comments.

Without sounding too clichéd I would like to thank my incredibly supportive family for all their advice and inspiration throughout my academic and personal life. I dedicate this project to my Granddad, David Curtis, area manager of the Yorkshire electricity board.

Contents

Chapter 1

Introduction

1.1 Motivation

This project is concerned with developing an efficient and computationally tractable model for the expansion of electrical infrastructure. The task of the system operator can be described as follows. An existing electrical infrastructure consists of several different regions, each with a number of generators, connected via a number of transmission lines. At some initial point in time an investment decision is made regarding which of the candidate generators and transmission lines to build. The new facilities are then put into service with the existing system for some time period; then the demand for power, plant and transmission line failures are all revealed. The system operator then adjusts the power produced at each generator and transported between regions in order to satisfy demand, whilst also minimising operating costs. In an extended version of the system the decision maker will have the opportunity to make further investment and operational decisions sequentially at later time periods, allowing him to tailor his decisions to the realisations of the uncertain economic conditions. Hence, the aim of the system operator is to choose the optimal investments in order to minimise the expected cost whilst hedging against the worst case realisation of the failure and demand patterns that can reasonably be expected to occur.

Usually the cost of power cuts to businesses can be significantly higher than the initial investment cost of the generators and as a result the reliability of the system is crucial. Beenstock illustrates the importance that these investment decisions can have on businesses that require a continuous supply of electricity, and states that often generators are built as a means of insurance against power outages [4]. In this situation the system operator may be faced with excess supply which also undoubtedly incurs extra, and sometimes significant, costs. Therefore it is of the utmost importance that the decision maker carefully balances the trade off between lower costs, and reliability. Achieving lower costs is not as straightforward as reducing investment as there is a trade-off between initial costs and operating costs, which is detailed in Section 2.4. The complexity of the problem is added to by increasing competition in the sector due to its ongoing liberalisation and the popularity in viable environmentally friendly alternatives. It is quite clear then that the decisions on whether to invest in new facilities or not can have significant consequences on the robustness of the system and cost to the system operator. The amount of money invested in such facilities adds to the importance of this decision. In a case study issued shortly after the electric power blackout in north-eastern U.S. and Canada an estimated investment of around $50 to $100 billion was needed to modernize the electrical grid [31]. All of this has motivated research within the stochastic programming community for more accurate and viable models to solve the variety of problems faced in modern electricity markets.

1.2 Recent Works

A number of works have been published on short-term optimal portfolio management for an electricity supplier. Fleten *et al* use a stochastic programming model to coordinate the generation and dispatch of resources into a competitive energy market over a two year horizon [20]. It was shown that due to the deregulation of electricity markets and the introduction of electricity exchanges, methods previously used in other financial markets can be applied to optimise the value of an electrical portfolio [44]. The timing of when to release water in hydro-electric plants is a major consideration when trying to optimise profit. The difficulty arises as the trading periods of financial instruments are usually of a much shorter time horizon(hours) than the dispatch of electricity(months). Pritchard *et al* propose a dynamic programming solution to this problem, taking into account multiple trading periods [40]. Other models that focus on optimal power generation and trading in the short to midterm are studied by Gröwe-Kuska & Römisch, Nürnberg & Römisch and more recently Doege *et al*[25, 38, 15].

The model developed in this project is concerned with a long-term view of expansion planning, as opposed to short term scheduling and trading as described above. An influential case study on expansion of the Brazilian hydroelectric power system comes from Leopoldino *et al*. They study a centralised, regulated system of interconnected generating units with a maximum production and transmission capacity, and use Monte-Carlo simulation to minimise cost whilst catering for the expected increase in load demand over time [32]. Nogales *et al* later proposed a decomposition methodology for the optimal flow problem of a multi-area decentralised system. An advantage of their model is that they cater for the increased competition in the sector and it is applicable to an independent system operator [37]. Chaton and Doucet have recently applied linear decision rules to a case study of Hydro-Québec [10].

1.3 Contributions

A model for a long-term (decades) capacity expansion problem suitable for an individual system operator in a deregulated market is proposed. It accounts for a multi-regional system, such as a national grid, with uncertain demand and allows for multiple investment periods. A linear decision rule (LDR) approximation, proposed by Kuhn *et al*, is applied to both the primal and dual stochastic programs in order to obtain a tractable linear program. This, method allows us to estimate the error introduced through the duality gap [30]. In an attempt to obtain a more accurate solution a continuous piecewise linear decision (CPLDR) rule approximation, proposed by Georghiou *et al*, is later applied. The contributions of this project will primarily focus on the following tasks:

- In Section 2.2 the LDR techniques used to obtain tractable reformulations of the stochastic capacity expansion problem are introduced. An improvement of this technique using CPLDR is introduced in Section 2.2.2.
- Chapter 3 introduces two stochastic capacity expansion models, a two-stage and multi-stage formulation. An implementation of these models is also provided which provides simulations for the uncertainties in the system.
- The models will be evaluated based on a case study of the electricity system in El Salvador. The numerical results are detailed in Chapter 5.
- Finally, an evaluation of the capacity expansion models is undertaken and ideas for future work proposed in Chapter 6.

Chapter 2

Background

2.1 Mathematical Overview

An overview of some of the technical concepts which are required to understand the main workings of this project will now be given. It is assumed that the reader has a general knowledge of probability and linear algebra or linear programming, so basic concepts of probability will not be covered. Instead an overview of linear and stochastic programming is given so that the reader can understand the methods used to obtain a tractable problem of the capacity expansion model.

Mathematical Optimisation

In its simplest form a mathematical optimisation problem is of the form

$$\begin{aligned} &\text{minimise } f_0(x) \\ &\text{subject to } f_i(x) \leq b_i, i = 1, ..., m \end{aligned}$$

The objective of such a problem is to find a vector $x = (x_1, ..., x_n)$ that minimises the function $f_0 : \mathbb{R}^n \to \mathbb{R}$. Hence x is known as the *optimisation variable* and f_0 the *objective function*. The functions $f_i : \mathbb{R}^n \to \mathbb{R}$ are known as the *constraint functions* and the constants $b_1, ..., b_m$ the bounds of the constraints [7].

Linear Program

A linear program is a specific type of mathematical optimisation where the objective function and constraints are linear. That is, they can be written as a linear combination of the vector of variables x. They are generally of the form

$$
\begin{aligned}
\text{minimise } & c^\top x \\
\text{subject to } & Ax \leq b \\
& x \geq 0 \\
& x, c \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m
\end{aligned} \tag{2.1}
$$

As before our aim is to find an x which minimises the objective function $c^\top x$. The feasible region of the problem is defined to be those values of x which satisfy the conditions $Ax \leq b$ with $x \geq 0$. The above formulation is often referred to as the *primal* problem and provides us with an upper bound of the solution.

It is desirable to re-formulate the above problem to provide us with a lower bound to the optimal value. The *dual* of the 2.1 is defined as

$$
\begin{aligned}
\text{maximise } & b^\top y \\
\text{subject to } & A^\top y \geq c \\
& y \geq 0 \\
& b, y \in \mathbb{R}^m, A \in \mathbb{R}^{m \times n}, c \in \mathbb{R}^n
\end{aligned} \tag{2.2}
$$

The *weak duality theorem* states that the dual is bounded above by the primal. More formally it states that, if x is any feasible primal solution and y is any dual feasible solution then $\sup b^\top y \leq \inf c^\top x$ [1]. The importance of the relationship between the primal and dual will become clear later, when the *duality gap* is used to estimate the loss in optimality. These are complementary, in the sense that an optimal solution to one problem yields an optimal solution to the other. The interested reader is referred to Kall and Wallace for a formal proof of these results [28, p. 72].

Dynamic Decision Making

In real life problems a decision maker may have the option to make multiple decisions at various points in time, in an attempt to optimise the performance of a stochastic system. Thus, the decisions made at a later point in time may depend on those decisions made earlier and on the uncertain data that is revealed. Edwards defines dynamic decision making by three common features [17]:

1. A series of actions is taken over some extended time to achieve a predetermined goal.
2. The environment changes spontaneously and accounts for previous decisions.
3. Decisions are interdependent so that later decisions depend on earlier actions.

[1] Here, inf and sup defined as the infemum and supremum respectively

Stochastic programs can be used to optimise dynamic systems where later decisions are modeled as functions of the uncertain data often referred to as *decision rules*. In their most basic form stochastic programs consist of only two stages, but often this is extended to a multi-stage problem.

Probability Theory

A *probability space* is defined as a triple $(\Omega, \mathcal{F}, P)$ where Ω represents the sample space of all possible outcomes, $\mathcal{F}$ is a $\sigma-$algebra or set of events of Ω assigned a probability by the *probability measure* P, with $P(\Omega) = 1$. From hereon a probability space is assumed to be of the form $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), P)$ where $\mathcal{B}(\mathbb{R}^k)$ is the Borel $\sigma-$algebra of $\mathbb{R}^k$. Intuitively speaking the Borel $\sigma-$algebra can be thought of as the set of events which are assigned probabilities by the probability measure.

A *square integrable function* of a variable, on an interval, is defined as one whose square of its absolute value is finite on that interval. The space of all Borel measurable, square integrable functions from $\mathbb{R}^k$ to $\mathbb{R}^n$ is denoted by $\mathcal{L}^2_{k,n} = \mathcal{L}^2(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), P; \mathbb{R}^n)$. The elements of the sample space are denoted by $\xi \in \mathbb{R}^k$, with their support Ξ defined as usual, the smallest subset of $\mathbb{R}^k$ with probability one. Let $\mathbb{E}(.)$ denote the expectation with respect to all Borel measurable, square integrable functions from $\mathbb{R}^k$ to $\mathbb{R}^n$.

A *stochastic process* can be thought of as a sequence of random variables indexed by time. Denote a stochastic process by $\xi := \{\xi(t), t \in T\}$ on the probability space $(\Omega, \mathcal{F}, P)$. At a fixed time point $t \in T$, $\xi(t)$ denotes the random variable of the stochastic process ξ at that time. A sequence of evolving random observations $\xi(\omega, t)$ over time t for some particular event ω is defined as a random walk, or sample path, $\{\xi(\omega, t) : t \geq 0\}$. The simplest, and most relevant, example of a sample is the price of a stock over a period of time.

The Lebesgue integral is commonly used when dealing with a probability space, as opposed to the more common Riemann integration. Intuitively speaking, the difference between the Lebesgue and Riemann integrals are in the approach to summing the area under a curve. The Riemann integral divides the domain space into portions and uses the value mapped in the range from the two end points. In contrast, the Lebesgue integral divides the range space and ask how much of the domain is mapped by the function to the range space endpoints. Measure theory provides a method for determining how much of the domain is sent to a particular portion of the range. The Lebesgue integral is denoted by $\int_\Omega P(\mathrm{d}\omega)$ where P is a measure on the probability space.

Two-Stage Stochastic Program

The most common form of stochastic program is the two-stage problem. A two-stage linear stochastic program with right hand side uncertainty as was first proposed by Dantzig is now considered [14]. In the first stage a set of decisions is made before the uncertainty is realised and these are known as *here-and-now* decisions for obvious reasons. Then, the uncertain data is revealed and the decision maker can adjust his strategy by making a second set of decisions known as *wait-and-see* or *recourse* decisions. The recourse decisions are modeled as decision rules, that is functions of the uncertain information. The decision maker aims to choose the initial decisions and decision rules which minimise the expected cost. Formally this can be writen as follows

$$\begin{aligned} \text{minimise } & c^\top x + \mathbb{E}\left[d(\xi)^\top y(\xi)\right] \\ \text{subject to } & Ax + By(\xi) \le b(\xi) \\ & x \ge 0, y \in \mathcal{L}^2_{k,m} \end{aligned} \tag{2.3}$$

with the following notation:

c : the first-stage costs with $c \in \mathbb{R}^{n_1}$

x : the first-stage decision variables with $x \in \mathbb{R}^{n_1}$

$d(.)$: the second-stage costs with $d \in \mathcal{L}^2_{k,n_2}$

$y(.)$: the second-stage decision variables with $y \in \mathcal{L}^2_{k,n_2}$

A : the first-stage coefficient matrix with $A \in \mathbb{R}^{m,n_1}$

B : the second-stage coefficient matrix with $B \in \mathbb{R}^{m,n_2}$

b : the constraint variables with $b \in \mathcal{L}^2_{k,m}$

The equation 2.3 deals with the case of *fixed recourse* implying that the recourse matrix B is independent of ξ. If the above were a case of *random recourse* then B would be replaced by $B(\xi)$ as independence is no longer guarenteed. Sochastic programs of the form 2.3 can also be written in standard form by introducing slack variables $s(\xi) \in \mathcal{L}^2_{k,m}$, s.t. $s(\xi) \ge 0$.

Multi-Stage Stochastic Program

Notation: In this section a sequence of decisions and random variables are observed, and so it is important that a distinction between the different notational meanings. The elements of the sample space are defined in a temporal structure as $\xi = (\xi_1, ..., \xi_T)$ and each of the sub-vectors $\xi_t \in \mathbb{R}^{k_t}$ are observed at the time points $t \in \mathbb{T} = \{1, ..., T\}$ respectively. For notational simplicity the history of all samples up to the time point t will be denoted by $\xi^t = (\xi_1, ..., \xi_t) \in \mathbb{R}^{k^t}$ where $k^t = \sum_{s=1}^{t} k_s$. Naturally, $\mathbb{E}_t(.)$ is defined as the conditional expectation of the history of observations ξ^t with respect to $\mathbb{P}$. Writing $A \ge 0$ for some matrix A or $v \ge 0$ for some vector v, refers to element wise inequality.

It is usually the case in a dynamic system that a sequence of decisions must be made over some extended time period $\mathbb{T}$. Hence, the two-stage model needs to be extended to a multi-stage one. In this system the first-stage decisions x are made at t_0 before the observation ξ_1, followed by a sequence of later decisions $y_t(\xi^t) \in \mathbb{R}^{n_t}, t \in \mathbb{T}$ each made after ξ^t is observed, but before future outcomes[2]. Thus, similarly to before, the objective

[2] This represents the non-anticipative nature of the system, i.e. y_t depends solely on ξ^t

is now to choose the initial decision and a sequence of decision rules to minimise the expected cost. A mutli-stage model is defined as follows

$$
\begin{aligned}
&\text{minimise } c^\top x + \mathbb{E}\left[\sum_{t=1}^{T} d_t(\xi^t)^\top y(\xi^t)\right] \\
&\text{subject to } Ax + \mathbb{E}_t\left[\sum_{s=1}^{T} B_{ts} y_s(\xi^s)\right] \leq b_t(\xi^t) \\
&\qquad x \geq 0, y_t \in \mathcal{L}^2_{k^t, m_t}
\end{aligned}
\tag{2.4}
$$

with the constraints defined as in stochastic program 2.3.

2.2 Linear Decision Rules

Stochastic programs attempt to integrate the concepts of optimisation & uncertainty and provide a huge amount of potential in solving real life problems; applicable to a number of different disciplines. However, the stochastic programming community believe that dynamic decision problems are extremely difficult to solve, especially when multiple stages are involved and the uncertain data is revealed sequentially. It has been shown recently by Dyer and Stougie that when the stochastic parameters are independently distributed, they are #P-hard, even if the stochastic parameters have the uniform distribution on the unit square [16]. In recent years many attempts have been made to come up with tractable approximations to these problems and one such popular approach at the moment is the use of linear decision rules.

One of the difficulties with stochastic programs is that they require full knowledge of the distribution of the uncertain data. This information is rarely available and the optimal solution can be very sensitive to the assumed distribution. This has revived the interest in robust optimisation as an alternative approach to handling the uncertain data, with the development of a *minimax approach*, where the decision maker minimises over the maximal expected cost of a family of distributions. Jogannathan studied such an approach that only requires knowledge of the mean and variance of the random parameters by solving an equivalent convex programming problem [27]. Linear decision rules have only been revived recently by Ben-Tal *et al* who introduced the Affinely Adjustable Robust Counterpart (AARC), a notion which was shown to have a significant reduction in the complexity of the problem [5]. Shapiro and Nemirovski suggest that, although linear decision rules almost always incur a loss of accuracy in the optimal solution, in "actual applications it is better to pose a modest and achievable goal rather than an ambitious goal which we do not know how to achieve" [41, p. 31-32]. Their success has very recently been seen in the context of portfolio optimisation over multiple periods [8], and in this project they will be applied to the problem of capacity expansion.

2.2.1 Primal and Dual Linear Decision Rules

This section discusses a theory introduced by Kuhn *et al* that uses linear decision rules on the primal and dual problem of a stochastic program to represent them as equivalent tractable linear programs. Furthermore, this technique allows us to estimate the error introduced via the optimality gap and only requires knowledge of the first two moments of the probability distribution. This approach is applicable to both our two-stage and multi-stage programs 2.3 and 2.4. Only the main theorems are stated here and the reader is advised to see the work by Kuhn *et al* for a more in-depth understanding [30].

The concept of an *affine function* $f : \mathbb{R}^n \to \mathbb{R}^m$, which is simply one that can be written as the sum of a linear function and a constant, is needed. In this section samples are taken from Ξ as defined in Appendix A and therefore it is certainly the case that $\xi_1 = 1$ almost surely. This allows for notational simplicity as affine functions of the uncertain data can be represented as linear decision rules. Formally, the structure of the decision rules are

$$x(\xi) = x_1 + \sum_{i=2}^{k} x_i \xi_i = X\xi$$

This compact notation will be assumed throughout the rest of this project.

One-Stage

This first result is concerning a one-stage stochastic program with fixed recourse, which is identical to stochastic program 2.3 without any initial decisions made before the uncertain data is realised. This one stage stochastic program will be hereon referred to by $\mathcal{SP}$. Using linear decision rules the primal problem is represented by $\mathcal{SP}_u$

$$\begin{aligned} \text{minimise } & \operatorname{Tr}\left(MC^\top X\right) \\ \text{subject to } & X \in \mathbb{R}^{n\times k}, S \in \mathbb{R}^{m\times k} \\ & \left.\begin{aligned} AX\xi + S\xi &= B\xi \\ S\xi &\geq 0 \end{aligned}\right\} \mathbb{P} - a.s. \end{aligned} \tag{$\mathcal{SP}_u$}$$

The corresponding dual problem $\mathcal{SP}_l$ can be written as

$$\begin{aligned} &\text{minimise } \operatorname{Tr}\left(MC^\top X\right) \\ &\text{subject to } X \in \mathbb{R}^{n\times k}, S \in \mathbb{R}^{m\times k} \\ &\qquad AX + S = B \\ &\qquad \left.\begin{aligned} \exists x \in \mathcal{L}^2_{k,n} : XM = \mathbb{E}(x(\xi)\xi^\top) \\ \exists s \in \mathcal{L}^2_{k,m} : SM = \mathbb{E}(s(\xi)\xi^\top) \\ s(\xi) \geq 0 \end{aligned}\right\} \mathbb{P} - a.s. \end{aligned} \tag{$\mathcal{SP}_l$}$$

Here, M is the second order moment matrix defined as $M := \mathbb{E}(\xi\xi^\top)$. As stated previously, the primal $\mathcal{SP}_u$ and dual $\mathcal{SP}_l$ provide an upper and lower bound to the original approximation. This can be seen as $\mathcal{SP}_u$ is obtained by reducing the underlying feasible set, and is thus more restrictive. Conversely any feasible solution to our original problem clearly satisfies $\mathcal{SP}_l$, although not necessarily the other way round, and hence $\mathcal{SP}_l$ provides a lower bound. These problems are still generally intractable as $\mathcal{SP}_u$ has a semi-infinite quality, whilst $\mathcal{SP}_l$ requires the solution of m moment problems. This motivates us again, by some other means, to drastically reduce the complexity of these problems.

One of the useful properties of linear programs is that they are generally solvable. Thus, if $\mathcal{SP}_u$ and $\mathcal{SP}_l$ can be approximated by linear programs, then they can be efficiently solved. This is the main concept behind the primal and dual approximation, which for the one-stage case of fixed recourse can be summarised in the theorem to follow.

Theorem 2.1: *If $\mathbb{P}$ has a polyhedral support of type (Appendix A) while $\mathcal{SP}$ has fixed recourse and is strictly feasible, then $\mathcal{SP}_u$ and $\mathcal{SP}_l$ are equivalent to the linear programs $\mathcal{SP}_u^*$ and $\mathcal{SP}_l^*$ respectively. The sizes of these linear programs is polynomial in k, l, m, and n, implying that they are efficiently solvable.*

$$\begin{aligned} &\text{minimise } \operatorname{Tr}\left(MC^\top X\right) \\ &\text{subject to } X \in \mathbb{R}^{n\times k}, \lambda \in \mathbb{R}^{m\times l} \\ &\qquad AX + \Lambda W = B \\ &\qquad \Lambda h \geq 0 \\ &\qquad \Lambda \geq 0 \end{aligned} \tag{$\mathcal{SP}_u^*$}$$

$$\begin{aligned} &\text{minimise } \operatorname{Tr}\left(MC^\top X\right) \\ &\text{subject to } X \in \mathbb{R}^{n\times k}, S \in \mathbb{R}^{m\times k} \\ &\qquad AX + S = B \\ &\qquad \left(W - he_1^\top\right) MS^\top \geq 0 \end{aligned} \tag{$\mathcal{SP}_l^*$}$$

Proof. See Kuhn *et al* [30, Section 2] □

Multi-Stage

Shapiro and Nemirovski argue that multi-stage problems are generically computationally intractable, and with sampling based algorithms existing theoretical qualities deteriorate dramatically as the number of stages grows [41]. On note of this, it is useful to extend Theorem 2.1 to the multi-stage case. Again, the multi-stage problem here does not cater for initial decisions before the first uncertainty element ξ_1 is realised but, as before, can be extended to such a scenario. This multi-stage stochastic program without an initial decision will be referred to as $\mathcal{MSP}$. The primal LDR approximation $\mathcal{MSP}_u$ of $\mathcal{MSP}$ can then be formulated as

$$\begin{aligned} &\text{minimise} \sum_{t=1}^{T} \operatorname{Tr}\left(P_t M P_t^\top C_t^\top X_t\right) \\ &\text{subject to } X_t \in \mathbb{R}^{n_t \times k^t}, S_t \in \mathbb{R}^{m_t \times k^t} \ \forall t \in \mathbb{T} \\ &\qquad \left.\begin{aligned} &\sum_{s=1}^{T} A_{ts} X_s P_s M_t P_t \xi + S_t P_t \xi = B_t P_t \xi \\ &S_t P_t \xi \geq 0 \end{aligned}\right\} \mathbb{P}-a.s.\ \forall t \in \mathbb{T} \end{aligned} \tag{$\mathcal{MSP}_u$}$$

and the dual LDR counterpart, $\mathcal{MSP}_l$ as

$$\begin{aligned} &\text{minimise} \sum_{t=1}^{T} \operatorname{Tr}\left(P_t M P_t^\top C_t^\top X_t\right) \\ &\text{subject to } X_t \in \mathbb{R}^{n_t \times k^t}, S_t \in \mathbb{R}^{m_t \times k^t} \ \forall t \in \mathbb{T} \\ &\qquad \left.\begin{aligned} &\sum_{s=1}^{T} A_{ts} X_s P_s N_t P_t + S_t P_t = B_t P_t \\ &\exists x_t \in \mathcal{L}^2_{k^t, m_t} : \mathbb{E}(x_t(\xi^t)\xi^\top) = X_t P_t M_t \\ &\exists s_t \in \mathcal{L}^2_{k^t, m_t} : \mathbb{E}(s_t(\xi^t)\xi^\top) = S_t P_t M_t \\ &s_t(\xi^t) \geq 0 \end{aligned}\right\} \mathbb{P}-a.s.\ \forall t \in \mathbb{T} \end{aligned} \tag{$\mathcal{MSP}_l$}$$

where $N_t = MP_t^\top (P_t M P_t^\top)^{-1}$ and P_t acts as a truncation operator to the current time point t, that is

$$P_t : \mathbb{R}^k \to \mathbb{R}^{k^t}, \xi \mapsto \xi^t$$

As with the two-stage problem, although this greatly reduces the complexity, this problem is still exceptionally difficult to solve. Hence, for problems $\mathcal{MSP}_u$ and $\mathcal{MSP}_l$ it is still necessary to find an efficient way to solve them. It has been shown by Kuhn *et al* that these problems can also be translated into linear programs.

Theorem 2.2: *Assume that $\mathbb{P}$ has a polyhedral support of type (Appendix A), while $\mathbb{E}_t(\xi) = M_t P_t \xi$ almost surely for some $M_t \in \mathbb{R}^{k \times k^t}, t \in \mathbb{T}$. If $\mathcal{MSP}$has deterministic constraint matrices and is strictly feasible, then $\mathcal{MSP}_u$ and $\mathcal{MSP}_l$ are equivalent to the linear programs $\mathcal{MSP}_u^*$ and $\mathcal{MSP}_l^*$ respectively. The sizes of these linear programs are polynomial $k, l, m := \sum_{t=1}^T m_t$, and $n := \sum_{t=1}^T n_t$, implying that they are efficiently solvable.*

$$\begin{aligned} \text{minimise} \quad & \sum_{t=1}^{T} \operatorname{Tr}\left(P_t M P_t^\top C_t^\top X_t\right) \\ \text{subject to} \quad & \left.\begin{aligned} X_t \in \mathbb{R}^{n_t \times k^t}, \Lambda_t \in \mathbb{R}^{m_t \times l} \\ \sum_{s=1}^{T} A_{ts} X_s P_s M_t P_t + \Lambda_t W = B_t P_t \\ \Lambda_t h \geq 0 \\ \Lambda_t \geq 0 \end{aligned}\right\} \forall t \in \mathbb{T} \end{aligned} \qquad (\mathcal{MSP}_u^*)$$

$$\begin{aligned} \text{minimise} \quad & \sum_{t=1}^{T} \operatorname{Tr}\left(P_t M P_t^\top C_t^\top X_t\right) \\ \text{subject to} \quad & \left.\begin{aligned} & X_t \in \mathbb{R}^{n_t \times k^t}, S_t \in \mathbb{R}^{m_t \times k^t} \\ & \sum_{s=1}^{T} A_{ts} X_s P_s N_t P_t + S_t P_t = B_t P_t \\ & \left(W - h e_1^\top\right) M P_t^\top S_t^\top \geq 0 \end{aligned}\right\} \forall t \in \mathbb{T} \end{aligned} \qquad (\mathcal{MSP}_l^*)$$

Proof. See Kuhn *et al* [30, Section 4] □

2.2.2 Continuous Piecewise Linear Decision Rules

At this point it has been seen that restricting the decision rules to the set of linear functions results in an equivalent, computationally tractable, linear program. It has been shown that this restriction sometimes results in infeasible instances of problems and the approximate solutions may be very far from the actual solution (see the remarks from Shapiro and Nemirovski [41, p. 31-33]). To overcome this issue and obtain a more accurate solution Chen *et al* proposed the use of *segregated linear decision rules*, where the idea is to divide each element in the uncertainty vector ξ into its positive and negative components [11]. The result is a combination of linear decisions applied to the piecewise uncertainties. More recently Goh and Sim show that SLDR's obtained in a similar manner to this provide greater flexibility whilst still preserving the non-anticipative requirements [24].

In this section an approach developed by Georghiou *et al* is explored, which solves a stochastic program on a lifted probability space that is equivalent to optimizing on our original parameter space using continuous piecewise linear decision rules (CPLDR's). Furthermore, applying the primal and dual approximations discussed in the previous section to these lifted problems has been shown to provide a better approximation to the original problem than applying the primal and dual approximations to the LDR's [21]. For a more in depth understanding of this section it is highly recommended to see the paper by Georghiou *et al*.

Firstly, the lifting operator which will be used to construct the lifted parameter space $\mathbb{R}^k$ is defined. The continuous lifting operator L and a linear retraction operator R are defined as the mappings to and from the lifted space

$$\begin{aligned} L : \mathbb{R}^k &\to \mathbb{R}^{k'}, \qquad && \xi \mapsto \xi' \\ R : \mathbb{R}^{k'} &\to \mathbb{R}^k, \qquad && \xi' \mapsto \xi \end{aligned}$$

The structure of the lifted space is controlled by a series of breakpoints on each coordinate axis of $\mathbb{R}^k$ denoted by

$$z_1^i < ... < z_{r_i-1}^i \qquad\qquad i = 2, ..., k,$$

where $r_i \in \mathbb{N}$, with $r_1 = 1$ due to the degenerate nature of the first uncertain parameter ξ_1. The vectors ξ' of the lifted space $\mathbb{R}^{k'}$ with dimension $k' := \sum_{i=1}^{k} r_i$ can be written as

$$\xi' = (\xi'_{11}, \xi'_{21}, ..., \xi'_{2r_2}, \xi'_{k1}, ..., \xi'_{kr_k})^\top$$

These breakpoints define the structure of the lifting operator $L = (L_{11}, ..., L_{kr_k})$ whose L_{ij} coordinate mapping corresponding to the ξ'_{ij} axis in the lifted space is

$$L_{ij}(\xi) := \begin{cases} \xi_j & \text{if } r_i = 1, \\ \min\left\{\xi_i, z_1^i\right\} & \text{if } r_i > 1, j = 1, \\ \max\left\{\min\left\{\xi_i, z_j^i\right\} - z_{j-1}^i, 0\right\} & \text{if } r_i > 1, j = 2, ..., r_i - 1 \\ \max\left\{\xi_i - z_{r_i-1}^i, 0\right\} & \text{if } r_i > 1, j = r_i \end{cases} \tag{2.5}$$

With this lifting operator the continuous piecewise linear decision rule problems on our original space Ξ of $\mathcal{SP}$, can be solved by restricting the decision rules to those representable as $x(\xi) = X'L(\xi)$ and $y(\xi) = Y'L(\xi)$ for $X \in \mathbb{R}^{n\times k'}$ and $Y \in \mathbb{R}^{m\times k'}$. The primal for $\mathcal{SP}$ with these restrictions on x is

$$\begin{aligned} &\text{minimise } \mathbb{E}_\xi \left(c(\xi)^\top X' L(\xi)\right) \\ &\text{subject to } AX'L(\xi) \le b(\xi) \ \mathbb{P}_\xi - a.s \end{aligned} \tag{$\mathcal{SP}_u^L$}$$

and the corresponding dual problem is reformulated as

$$\begin{aligned} &\text{minimise } \mathbb{E}_\xi \left(c(\xi)^\top x(\xi)\right) \\ &\text{subject to } x \in \mathcal{L}^2_{k,n}, s \in \mathcal{L}^2_{k,m} \\ &\qquad \left.\begin{aligned} \mathbb{E}_\xi \left(\left[AX(\xi) + s(\xi) - b(\xi)\right] L(\xi)^\top\right) = 0 \\ s(\xi) \ge 0 \end{aligned}\right\} \mathbb{P}_\xi - a.s \end{aligned} \tag{$\mathcal{SP}_l^L$}$$

These problems that correspond to the CPLDR on Ξ are generally intractable for lifting operators of the form 2.5. Thus, an alternative approach to solving these problems must be taken. Consider the stochastic program $\mathcal{SP}'$ over the lifted probability space $\Xi' = L(\Xi)$

$$\begin{aligned} &\text{minimise } \mathbb{E}_{\xi'} \left(c(R\xi')^\top x(\xi')\right) \\ &\text{subject to } Ax(\xi') \le b(R\xi') \ \mathbb{P}_\xi - a.s \end{aligned} \tag{$\mathcal{SP}'$}$$

It has been shown that $\mathcal{SP}'$ is equivalent to $\mathcal{SP}$. Now consider the primal and dual LDR representations, $\mathcal{SP}'_u$ and $\mathcal{SP}'_l$, of $\mathcal{SP}'$ similar to the formulations of $\mathcal{SP}_u$ and $\mathcal{SP}_l$ as in Section 2.2.1. The following proposition is needed to achieve the goal of solving the continuous piecewise linear decisions rules formulation.

Proposition 2.1:

1. *Problems $\mathcal{SP}_u^L$ and $\mathcal{SP}'_u$ are equivalent.*
2. *Problems $\mathcal{SP}_l^L$ and $\mathcal{SP}'_l$ are equivalent.*

Proof. See Georghiou *et al* [21, Section 3] □

Now in order to reformulate $\mathcal{SP}'_u$ and $\mathcal{SP}'_l$ into tractable linear programs $\mathcal{SP}'^*_u$ and $\mathcal{SP}'^*_l$ using the techniques in the previous section the convex hull of the lifted space Ξ' must have a tractable representation of the form in Appendix A.1. For lifting operators of the form 2.5 there exists a tractable representation for *con* Ξ' if Ξ constitutes a hyperrectangle within $\left\{\xi \in \mathbb{R}^k : e_1^\top \xi = 1\right\}$. See Georghiou *et al* for a more in-depth discussion on the construction of *con* Ξ' [21, Section 4].

Thus, the problems $\mathcal{SP}'_u$ and $\mathcal{SP}'_l$ can be solved using the approximate linear program formulation and this extra approximation on *con* Ξ'. Furthermore, solving these problems is equivalent optimising the CPLDR formulations of $\mathcal{SP}_u^L$ and $\mathcal{SP}_l^L$. Finally, the following theorem establishes the benefits in solving this equivalent CPLDR formulation of $\mathcal{SP}$, as oppose to a LDR problem, by tightening the gap between the upper and lower bounds.

Theorem 2.2: *The optimal values of the approximate problems* $(\mathcal{SP}_u)$, $(\mathcal{SP}'_u)$, $(\mathcal{SP}_l)$ *and* $(\mathcal{SP}'_l)$ *satisfy the following chain of inequalities*

$$\inf\mathcal{SP}_l \leq \inf\mathcal{SP}'_l \leq \inf\mathcal{SP} = \inf\mathcal{SP}' \leq \inf\mathcal{SP}'_u \leq \inf\mathcal{SP}_u$$

Proof. See Georghiou *et al* [21, Section 3] □

2.3 Financial Mathematics

2.3.1 Cholesky Decomposition

The *Cholesky decomposition* is a decomposition of a symmetric positive-definite matrix into the product of a lower triangular matrix, and it transpose. Formally A can be decomposed as

$$A = LL^\top$$

where L is a lower triangular matrix with strictly positive diagonal entries often known as the Cholesky decomposition. The Cholesky decomposition is commonly used in Monte-Carlo simulations to generate correlated samples from a random variable or stochastic process. Given the covariance matrix $\Sigma := \text{cov}(X_i, X_j)$ one can premultiply a vector of independently sampled random variables z by the Cholesky decomposition L of Σ. The result is a vector of samples u which represent the correlated properties of the covariance matrix. So if $\Sigma = LL^\top$ and z is a vector of independent random values, then $u = Lz$ is a vector of correlated random values with respect to the covariance matrix.

2.3.2 Wiener Process

A *Wiener process* is a continuous time stochastic process, sometimes known as *Brownian motion*, which has many applications in finance and most notably in the Black-Scholes option pricing model. It is characterized by the following properties:

- $W_0 = 0$ with probability 1
- (Independent Increments) If $0 \leq t_1 < t_2 \leq t_3 < t_4$ then $W_{t_2} - W_{t_1}$ and $W_{t_4} - W_{t_3}$ are independent random variables
- If $0 \leq s < t$ then $W_t - W_s \sim N(0, t - s)$

In the Schwartz model (see Section 4.6), just as in Geometric Brownian motion, the Wiener process acts as 'shocks' on the values of successive prices from S_0. Notice how if $s = 0$ in the above then $W_t \sim N(0, t)$ and the variance is proportional to the square root of time.

2.3.3 Ito's Lemma

If a random process x is defined by the *Ito process*

$$dx(t) = a(x,t)dt + b(x,t)dz \tag{2.6}$$

where z is a standard Wiener process. Suppose also that the process $y(t)$ is defined by $y(t) = F(x,t)$. Then $y(t)$ satisfies Ito's equation

$$dy(t) = \left(\frac{dF}{dx}a + \frac{dF}{dt} + \frac{1}{2}\frac{d^2F}{dx^2}b^2\right)dt + \frac{dF}{dx}bdz \tag{2.7}$$

where z is the same Wiener process as in equation 2.6.

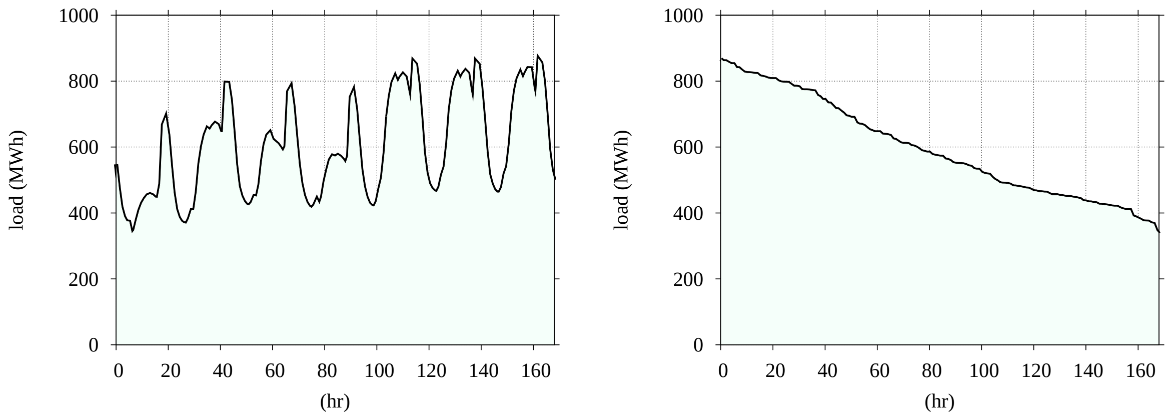
For more information on Ito processes and a formal proof of Ito's lemma see Luenberger [33].

2.4 Electricity Markets

Electricity capacity expansions is a fundamental example of a real options paradox. A real option is the right to undertake some business decision, with the hope that it will yield later rewards. In capacity expansion the option is to build new power facilities in the hope that they will reduce future operating costs, increase the resilience of the system and provide the capacity to meet future demand.

The demand for electrical energy can be seen to fluctuate throughout the day, with a minimum amount of continuous energy required throughout the day and more during certain hours. The curve that represents the demand at any hour of the day over the year (ranging in hours from $0-8760$) is often referred to as the *load curve*. It is convenient to view the demand for electricity in capacity requirements vs the utilisation, this curve is known as the *load duration curve*. The unique nature of demand has implications on the types of generators available; as a result investment decisions are made more complex. The load duration curve can be divided into regions representing the measure of capacity and the time it is required for. Two bands can be used to split the curve into three categories: the base, peak and shoulder load. An illustration of a monthly load curve is given in Figure 2.1a and the corresponding load duration curve is given in Figure 2.1b.

Plants referred to as *base load* plants will stay active constantly, rarely shutting down, unless of course they fail, to satisfy the base demand. Therefore, it makes sense that these base load plants are as economical to run as possible and hence they tend to have higher investment costs with lower operating costs in comparison to other plants. Examples of base load plants include nuclear and coal fueled plants. At peak hours for energy demand extra *peak load* plants are started up to cope with the rise in demand. These plants are only used to satisfy a fraction of the demand and it is more important that they do not cost too much to build, hence they tend to have much lower start-up costs at the expense of higher operational costs. Peak plants generally burn natural gas to fuel gas turbines such as diesel or jet fuel. In between base and peak load plants in terms or efficiency,



(a) Load curve

(b) Load duration curve

Figure 2.1: Electricity demand curves for one week [43]

investment costs and operating costs are *shoulder load* plants. Whilst it is not the aim of this project to analyse the short-term load curve and provide a model for optimal operational scheduling (see Nürnberg [38]), the trade-off the system operator must make between investment and operational costs must be taken into account.

It is also worth mentioning that not 100% of energy can be captured from production as waste is produced, and thus plants tend to have an efficiency factor associated with them. As a general rule of base load plants will tend to be more efficient and cost effective than peak load plants.

Levelised Cost

The *levelised energy cost* is the cost of generating energy at a power plant. It can be thought of as an economic assessment of the overall cost of energy production and includes investment, operating & maintenance and fuel costs. The levelised energy cost is usually used to assess the price which electricity should be sold at to ensure profit via a net present value calculation. However, it can also be used to approximate the fixed investment costs of a plant and its variable energy production and maintenance costs.

Figure 2.2 illustrates the trade off between investment and operating costs for different base and peak load plants as we described earlier.

Economies of Scale

Economies of scale refers to the cost advantages related to size of a project. They imply that as the size of a project increases there are cost advantages which result in a reduction in the cost per unit. In the electrical power industry there are economies of scale with both plant and transmission line investment. In simple terms, the larger the plant the lower the investment cost per MW. As a result these considerations should be accounted for when considering investment in new plants, clearly a full size industrial wind farm

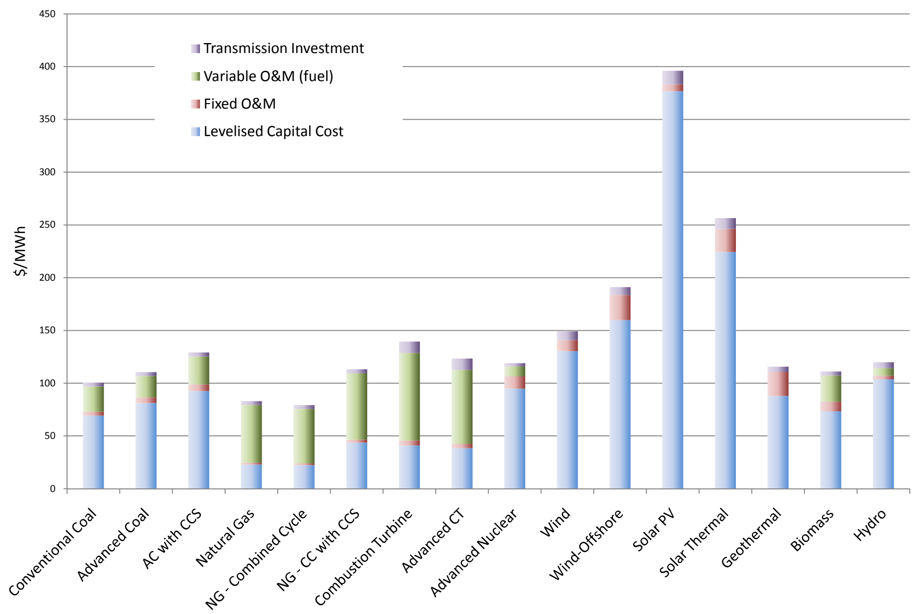


Figure 2.2: Levelised energy cost [18]

with over 50 turbines will have a lower investment cost per unit than to a private resident who wishes to purchase a small 10m turbine.

Fuel Prices

Thermal power plants burn fossil fuels to power turbines and create energy. As a result, the cost of energy generation at thermal power plants is largely dependent on the price of the fossil fuels that they burn which are traded in the commodity markets. The most common fuels used for energy production are coal, oil and gas. In recent years the commodity markets have been increasingly volatile and many companies have felt the impact of this. In the capacity expansion models of this project commodity prices need to be simulated in order to capture these fluctuating costs. The operational decision that performs best in the face of volatile prices would then be chosen.

Carbon Cap and Trade

More recently with a huge political emphasis on climate change and greenhouse gas emissions it is becoming increasingly important that the environmental impact of power generation is properly considered. It is understood that electricity generation accounts for a large portion of greenhouse gas emissions, and it accounts for 40% of emmisisons in the United States. These emissions are largely from the burning of fossil fuels such as coal, oil or gas; with coal producing more carbon dioxide than any other. To put things into perspective, the coal fired power plants in Wyoming produce more carbon dioxide in 8 hours than power generators in Vermont do in a whole year [6]. Clearly then this has become quite a significant problem and it is no wonder that emphasis has been placed on reducing emissions and creating new 'green energy'.

One proposed system for reducing greenhouse gas emissions is emissions trading, or cap and trade. The idea is to provide economic incentives to reduce greenhouse gases. A central authority provides a cap on the number of emissions, measured in carbon dioxide equivalents CO_2e, and issues an allowance to companies which give them the right to emit a number of tonnes of CO_2e. Whilst total emissions cannot exceed the cap those companies that have excess credits can sell them on the emissions market, and companies that need to pollute more can buy extra credits in the secondary markets. In 1997 a treaty was passed called the Kyoto Protocol where the most developed regions of the world agreed to legally binding targets for reducing their emissions in the six major greenhouse gases. The most common greenhouse gases are carbon dioxide, methane, nitrous oxide and sulfur. Renewable energy providers are also given certificates, or green tags, for every $1,000KWh$ that they produce, which can be sold on the market. Thus, providing a further incentive for renewable energy providers.

An example of such emissions trading system in place is the Chicago Climate Exchange (CCX) for projects in North America and Brazil. The financial instrument traded on the climate exchange is the CCX Carbon Financial Instrument (CFI), with each contract corresponding to 100 metric tons of CO_2e emissions.

Cap and Trade clearly adds to the complexity of the capacity expansion problems for the system operator. Not only must the environmental impact of emissions be considered, but there is also an added cost to producing highly pollutant energy, and a levy for green energy.

Renewable Energy

More recently emphasis has been placed on investment in renewable energy sources such as hydro, geothermal, wind and solar plants. The advantage of such plants is that they often have lower operating costs which are immune to volatile commodity prices. Their carbon emissions are also lower than traditional thermal plants. However, these advantages come at a the price of higher capital costs and the time taken to cover the initial investment can be quite long. Extra considerations must be made when choosing to invest in such resources, such as the available wind or solar radiation in a region and the amount of land that these farms will occupy. A recent phenomenon has been offshore wind farms, where larger turbines are used which can reach sizes of up to $120m$ in diameter. These turbines will often be rated at a capacity of $5MW$, in contrast to a $50m$ diameter turbine rated at approximately $700KW$.

One problem of wind and solar plants is that the energy that they produce throughout the day cannot be controlled and is unpredictable. This means that the availability of energy for use at certain times of the day cannot be guaranteed. As a result much effort is put into determining the location of wind and solar farms in regions that will allow them to perform optimally.

The power produced P in Wh from an individual wind turbine can be calculated based on its efficiency and size by

$$P = 0.5 \times \rho \times A \times V^3 \times C_p \tag{2.8}$$

where ρ represents the air density in units kg/m^3, A is equal to the rotor swept area from the turbine blades in m^2, V is equal to the wind speed in m/s. C_p is a performance coefficient or the efficiency of the wind turbine in converting the kinetic energy in the wind into mechanical energy. This factor is limited by Betz's law which states that the maximum efficiency of such a turbine cannot exceed 59%.

You can see from this equation that the power generated from a wind turbine is proportional to the wind speed cubed. Hence, proper placement of wind turbines can have a large impact on the power it produces. For this reason it is generally unlikely that industrial wind turbines are placed in regions with an average wind speed of less than roughly $6m/s$. The wind speed at a given location is clearly not constant and needs to be modeled by a probability distribution. The Weibull distribution closely mirrors the actual hourly wind speed, and is defined by the probability density function $f(x) = \eta\alpha^{-\eta}x^{\eta-1}\mathrm{e}^{-(x/\alpha)^\eta}$ on the support $x \in [0; +\infty)$; where $\alpha > 0$ is the scale parameter and $\eta > 0$ is the shape parameter. To model the hourly distribution of wind speeds the shape parameter is often chosen to be 2, which represents the Rayleigh distribution.

To the amount of energy in Wh generated from a tilted photovolatic panel can be estimated by using the global irradiance on a horizontal surface. The irradiance is a measure of power available per unit area of solar radiation and, for a period of 1 hour, is measured in Wh/m^2. The global irradiance on a tilted plane is a combination of direct, diffuse and reflected irradiance. This can be thought of as the direct sunlight on an area, sunlight from other surrounding unreflective objects the reflected sunlight from mirror like objects.

Thus, the irradiance gives the amount of energy in Wh that can be converted from the sun by every m^2 from a solar panel assuming 100% efficiency. This then needs to be multiplied by some efficiency coefficient to get the actual energy converted. The power P in Wh is then

$$P = E_g \times A \times C_p \tag{2.9}$$

where E_g is the global irradiance in in W/m^2, A is the area of the solar panel in m^2 and C_p is again the performance coefficient relating to the efficiency of the panel.

Chapter 3

Electricity Capacity Expansion

In this section the models developed for the electricity capacity expansion problem are introduced. First, a two-stage model is described which extends naturally to a multi-stage model discussed in the next section. The notation and definitions used throughout this chapter are largely influenced by a previous study of capacity expansion problems by Kuhn *et al* [29]. It may be useful to see Figure 3.1, for an illustration of a capacity expansion model.

3.1 Two-Stage

A brief introduction was given earlier to the capacity expansion problem and here the problem is clearly defined as well as the reasoning behind it. In the capacity expansion problem both the investment cost and operating cost must be accounted for in order to create a realistic long term model to minimise the overall cost. This project considers a regional or national system, and therefore the models must account for the high expense that may be incurred not only by new generators, but also by new transmission lines. As Leopoldino illustrates, in the Brazilian system only 55% of the investments go to generating facilities with a chunk of 30% on transmission lines, thus it would be unrealistic to ignore this rather significant cost [32, p. 368]. Furthermore, in a deregulated market power can be transferred between regions as a viable alternative to building new plants, and regional demand can be satisfied in this manner. Taking a long-term horizon on a regional level a number of assumptions and considerations need to be made in the model:

- Inter-period dependencies between operating periods can be disregarded
- Demand may increase or decrease between periods
- Failures are likely to occur at some point and the system must be resilient enough to handle these
- The demand of a region will be correlated with neighboring regions

- It will be assumed that energy produced from wind and solar farms cannot be stored
- Furthermore, it will be assumed that their will be no major changes in climate over our time horizon

The capacity expansion problem can then be described as follows. An existing power system has $K = \{1, ..., \overline{k}\}$ regions, $N = \{1, ..., \overline{n}\}$ generating units and $M = \{1, ..., \overline{m}\}$ transmission links. The generators are subdivided into existing $N_1 \subset N$ and candidates to be built $N_2 \subset N$. Similarly, the transmission links are divided into $M_1 \subset M$ and $M_2 \subset M$. It is sometimes useful to reference a particular class of power plants within a subset and for this we will use the superscript. For example the set of all candidate wind generators is denoted by $N_1^{\text{w}} \subset N$. Let $N(k)$ denote the number of existing and candidate generators in region k, $M_-(k)$ the transmission lines flowing into region k and $M_+(k)$ as those links flowing out of region k.

At the initial stage the decision maker chooses which plants and transmission lines to build out of the candidate choices. If a plant $n \in N_2$ is built then a cost of c_n is incurred. Similarly if a transmission line $m \in M_2$ is built then a cost d_m is incurred. The first-stage decision vectors $u = (u_n)_{n \in N_2}$ and $v = (v_m)_{m \in M_2}$ are defined by the binary variables

$$u_n = \begin{cases} 1 \text{ if plant } n \text{ is built,} \\ 0 \text{ otherwise,} \end{cases}$$

and similarly

$$v_m = \begin{cases} 1 \text{ if transmission line } m \text{ is built,} \\ 0 \text{ otherwise,} \end{cases}$$

The vector $x = (u, v)$ then represents the first stage decision variables. Next a sample $\xi(\omega, t)$ is observed from the measurable stochastic process $\xi : [0, T] \times \Omega \to \Xi$ on the probability space $(\Omega, \mathcal{F}, P)$ which represents the inherent uncertainties in the system. This process is composed of several stochastic processes, so that $\xi = (\alpha, \beta, \gamma, \delta, \pi, \zeta)$ is defined by:

1. $\alpha = (\alpha_n)_{n \in N}$, where α_n represents the stochastic failure of generator n. A value of $\alpha_n = 1$ indicates a fully functional generator and $\alpha_n = 0$ indicates a complete outage.
2. $\beta = (\beta_m)_{m \in M}$, where β_m represents the stochastic failure of transmission line m, with the values as above.
3. $\gamma = (\gamma_n)_{n \in N}$, where γ_n is the cost of producing one MWh of energy at generator n. This value will depend on both fixed O&M costs and stochastic fuel prices. It is therefore useful to think of it as a vector $\gamma_n = (\gamma_n^{\text{f}}, \gamma_n^{\text{o}}, \gamma_n^{\text{g}}, \gamma_n^{\text{c}}, \gamma_n^{CO_2})$ with cost components for fixed costs, oil, gas, coal and carbon emissions.

4. $\delta = (\delta_k)_{k \in K}$, where δ_k is the random demand at each region k defined as a factor of the total demand δ. Concretely speaking, $\delta_k = \lambda_k \delta$, with $\sum \lambda_k = 1$, so that $\sum \delta_k = \delta$. This factor model ensures demand between neighboring regions is correlated.

5. $\pi = (\pi_n)_{n \in N}$, where π_n represents the stochastic value of the carbon credits currently held for use by generator n at the beginning of the operating period. So, it represents the number of credits multiplied by the stochastic price of carbon at that time.

6. $\zeta = (\zeta)_{n \in N^{\text{w, s}}}$, where ζ_n is the random energy available in MWh for the wind or solar farm n.

After a sample is taken, the second-stage decision vector $y = (g, f)$ is made on the operational scheduling of the system conditional on the first-stage vector x in an attempt to minimise the operational cost over $[0, T]$. These decisions are written as

1. $g = (g_n)_{n \in N}$ is the energy generated at plant n in MWh. Each generator will have an upper bound $\overline{g}_n$ which represents the capacity of the plant.

2. $f = (f_m)_{m \in M}$ is the flow of energy through transmission lines between two regions. A positive flow $f_m > 0$ indicates the flow of energy from start to end, whilst a negative flow $f_m < 0$ is the opposite. Transmission lines therefore have both an upper bound $\overline{f}_m$ and lower bound $-\overline{f}_m$ representing the capacity available in both directions.

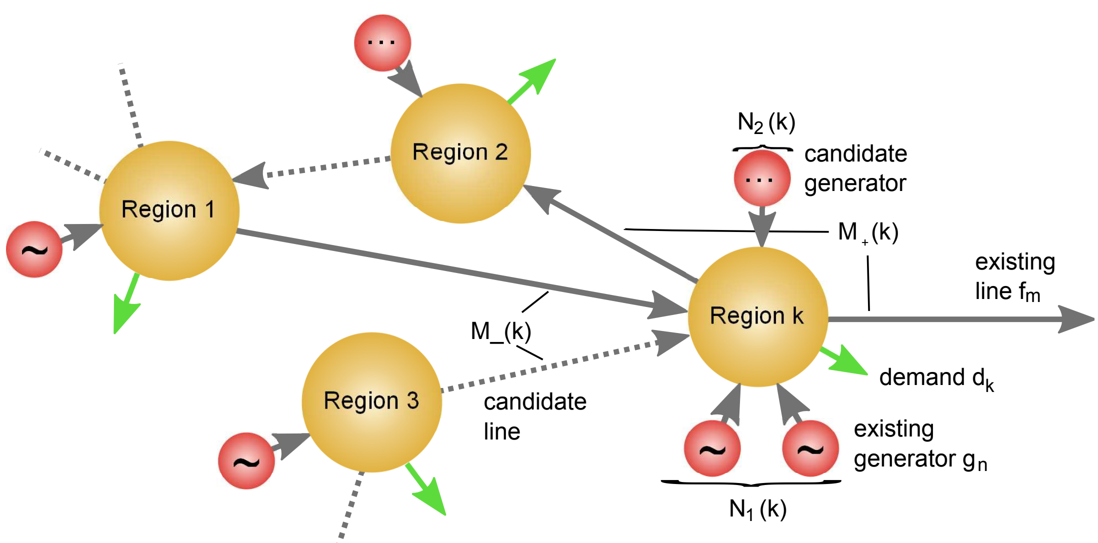


Figure 3.1: An electricity power system

In the model it is the aim to reduce operating costs over the whole time period $[0, T]$. This would result in an expectation in the objective function dependent on time, as

$$\int_{\Omega}\int_{0}^{T}\sum_{n\in N}\gamma_n(\omega,t)g_n(\xi(\omega,t))\mu(\mathrm{d}t)P(\mathrm{d}\omega) \tag{3.1}$$

where μ represents a Borel measure on the time interval $[0,T]$ which accounts for discounting and non-trivial time preferences, which will become clear in the next section. That is, μ is a measure of continuous compounding e^{-rt} where r is the interest rate. It will also be used to bias our predicted costs towards the assumptions made closer in the future than those further away. It is required that the constraints in our model represent an instantaneous moment in time and thus they need to be independent of time. In this sense it is useful to think of ξ as random vector on the augmented probability space $(\Omega',\mathcal{F}',P')$ defined as

$$\Omega'=[0,T]\times\Omega \qquad \mathcal{F}'=\mathcal{B}([0,T])\otimes\mathcal{F} \qquad P'=\frac{1}{\mu([0,T])}\mu\otimes P \tag{3.2}$$

where $\mu([0,T])$ is a normalising constant, which for continuous compounding is equal to $r/(1-\mathrm{e}^{-rT})$. Then it can be assumed without loss of generality that Ξ represents the support of ξ, and the expected cost 3.1 simplifies to

$$\mu([0,T])\int_{\Xi}\sum_{n\in N}\gamma_n g_n(\xi)P'(\mathrm{d}\xi)$$

which depends on time only through the random variable ξ, with P' a probability measure over ξ. This can then be rewritten as

$$\mu([0,T])\mathbb{E}'\left[\sum_{n\in N}\gamma_n g_n(\xi)\right] \tag{3.3}$$

where $\mathbb{E}'$ represents the expectation with respect to P'. This expectation is now independent of time as required.

Systems such as cap and trade require all emissions over a time period to be less than some threshold and thus they introduce a dependence on time. This means the constraints value h_{i+1} at time t_{i+1} depend on the value h_i at the previous time point t_i. To ensure a decoupling from time a sale of credits rule is proposed, where all the current carbon credits held by the system operator are sold at the beginning of the period. This is valid as in the model it is assumed that the independent system operator controls the energy production at all facilities, and thus is the owner of all the carbon credits. This leaves the constraints independent of time, except of course through ξ, whilst still accounting for the existing credits. In this situation all carbon credits required for operation are then bought from the spot markets as needed.

3.1.1 Two-Stage Model

A two-stage model of the capacity expansion model can now be formally defined. The objective function relates to the investment costs from first stage decisions, the expected operating costs conditional on these decisions and the final term is the sale of all current carbon credits, again conditional on the investment decisions.

$$
\begin{aligned}
\underset{u,v,f,g}{\text{minimise}} \quad & \sum_{n\in N_2} c_n u_n + \sum_{m\in M_2} d_m v_m + \mu([0,T])\mathbb{E}'\left[\sum_{n\in N}\gamma_n g_n(\xi)\right] \\
& -\left\{\sum_{n\in N_1}\pi_n + \sum_{n\in N_2} u_n\pi_n\right\} \\
\text{subject to} \quad & \\
& \sum_{n\in N} g_n^{\mathrm{CO_2e}}(\xi) \le \mathrm{CO_2e} && \forall n\in N \\
& \sum_{n\in N(k)} g_n(\xi) - \sum_{m\in M_+(k)} f_m(\xi) + \sum_{m\in M_-(k)} f_m(\xi) \ge \delta_k && \forall k\in K \\
& g_n(\xi)\ge 0 && \forall n\in N_1 \\
& g_n(\xi)\le \alpha_n\overline{g}_n && \forall n\in N_1 \\
& g_n(\xi)\ge 0 && \forall n\in N_2 \\
& g_n(\xi)\le \alpha_n\overline{g}_n u_n && \forall n\in N_2 \\
& g_n(\xi)\le \zeta_n && \forall n\in N^{\mathrm{w,\,s}} \\
& f_m(\xi)\ge -\beta_m\overline{f}_m && \forall m\in M_1 \\
& f_m(\xi)\le \beta_m\overline{f}_m && \forall m\in M_1 \\
& f_m(\xi)\ge -\beta_m\overline{f}_m v_m && \forall m\in M_2 \\
& f_m(\xi)\le \beta_m\overline{f}_m v_m && \forall m\in M_2 \\
\text{where} \quad & \\
& \Xi = \big\{\xi=(\alpha,\beta,\gamma,\delta,\pi,\zeta) : 0\le\alpha,\beta\le e, e^\top(e-\alpha)+ \\
& \quad e^\top(e-\beta)\le B, \gamma\in[\underline{\gamma},\overline{\gamma}], \delta\in[\underline{\delta},\overline{\delta}], \pi\in[\underline{\pi},\overline{\pi}], \zeta\in[\underline{\zeta},\overline{\zeta}]\big\}
\end{aligned}
\tag{3.4}
$$

In the above model notice that the term $\sum_{n\in N_1}\pi_{sn}$ corresponding to the sale carbon credits from existing plants could be removed as this is a constant does not impact the investment decisions. Any investment or operational decision must satisfy a number of constraints. The first constraint is optional and is included only if a multi-objective decision is required to minimise costs subject to a maximum rate of carbon emissions[1] over the time period $[0,T]$. The demand constraint in the second line ensures that the total power supplied to a region k at any instantaneous time point must not drop below the expected demand at that region. The remaining generation and flow constraints ensure that all second stage decisions adhere to the capacity available multiplied by the

[1]The rate of carbon emissions is in terms of million metric tonnes (mmt) of CO_2e per MWh

availability at that facility due to failures. For candidate plants the multiplication by the binary decision variables u_n from the first stage ensure $g_n \equiv 0$ if the plant is not built, that is if $u_n = 0$. The result is analogous for the transmission lines. The added constraints for wind and solar plants ensure that the power produced at wind or solar farm cannot exceed the random energy provided from the weather at that time. As renewable energy operation and fuel cost are negligible the amount of energy produced will only be limited by availability and weather, and thus will equal the minimum of the maximum of the two constraints.

3.2 Multi-Stage

The multi-stage model consists of a number of different investment and operational decisions occurring over a total time period of $[0, ..., T]$. Investment decisions may only be made at time points which have a succeeding operational period. This ensures that investments are made in the face of uncertainty with respect to the resulting operational decisions to be made in the next stage. Over the time period there may be a number of possible outcomes regarding the growth or contraction of demand and this can be represented via a scenario tree with the notation taken from [2].

A scenario tree of S nodes represents the possible outcomes of the random processes over time. Thus, at each node s there is an associated stochastic process ξ_s. Each node occurs with probability p_s, such that $\sum_{s \in \mathcal{L}} p_s = 1$. The time point associated with node s will be denoted by t_s. It is assumed that each node s in the tree, except the root node $s = 0$, has a unique parent, which is denoted by $a(s)$. Furthermore, the subtree of node s will be identified by $\mathcal{T}(s)$, so that $\mathcal{T}(0)$ represents the whole scenario tree. The scenario tree may not necessarily be balanced and a nodes children may not correspond to the same time period. It is also necessary to reference the leaf nodes, which represent the outcome scenarios, and these will be denoted by $\mathcal{L}$. Figure 3.2 provides an illustration of an example scenario tree.

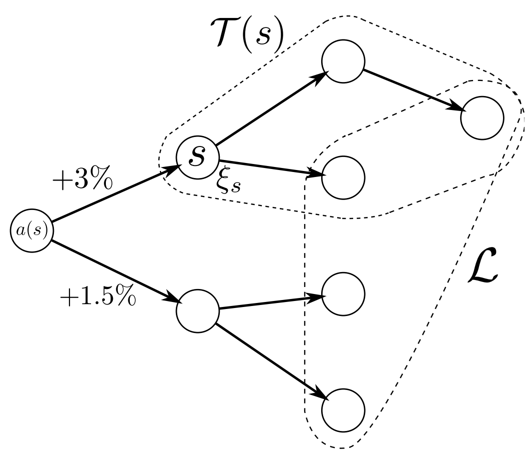


Figure 3.2: Scenario tree for the multi-stage model

This formulation will require that the objective function takes the expected investment and operational costs over all nodes, for their respective time periods. This is a reasonable goal, as the uncertainty with regards to demand growth over time will impact whether investment decisions are made today or at a later time point. Postponing some decisions to a later time point may result in a saving should this scenario not occur.

The notation defined in Section 3.1 will now be extended in the natural way to the multi-stage case. The investment decision vector $x = (x_s)_{s\in\mathcal{T}(0)/\mathcal{L}}$ is defined by $x_s = (u_s, v_s)$ where s is a non-leaf node in the scenario tree, and u_s and v_s are defined identically to before. In a similar manner the second stage decision variables can be extended to $y = (y_s)_{s\in\mathcal{T}(0)/\{0\}}$ with $y_s = (g_s, f_s)$ where s is a non-root node in the tree. Decisions are then made to ensure costs are minimised over the respective time intervals $\left[t_{a(s)}, t_s\right]$, which partition the interval $[0, T]$.

The mapping $\xi_s : \left[t_{a(s),}, t_s\right] \times \Omega_s \to \Xi_s$ can now be considered as a random vector on the probability space $(\Omega'_s, \mathcal{F}'_s, P'_s)$ defined similarly to 3.2 as

$$\Omega'_s = [t_{a(s)}, t_s] \times \Omega_s \qquad \mathcal{F}'_s = \mathcal{B}([t_{a(s)}, t_s]) \otimes \mathcal{F}_s \qquad P'_s = \frac{1}{\mu_s\left(\left[t_{a(s)}, t_s\right]\right)}\mu_s \otimes P_s$$

with μ_s representing the Borel measure on $\left[t_{a(s)}, t_s\right]$, accounting for discounting as before, with probability density function $\mathrm{e}^{-r_s t}$. In this multi-stage case it is not certain that $t_{a(s)} = 0$ and thus the normalising constant $\mu_s\left(\left[t_{a(s)}, t_s\right]\right)$ is equal to $r_s / \left(\mathrm{e}^{-r_s t_{a(s)}} - \mathrm{e}^{-r_s t_s}\right)$. Thus, the expected operating cost analogous to 3.3 is

$$\mu(\left[t_{a(s)}, t_s\right])\mathbb{E}'_s\left[\sum_{n\in N}\gamma_{sn} g_{sn}(\xi_s)\right]$$

where $\mathbb{E}'_s$ is the expectation of of ξ_s with respect to the probability measure P'_s.

3.2.1 Multi-Stage Model

With this notation the two-stage model 3.4 extends to a multi-stage scenario tree formulation. The two-stage model can be thought of as a generalisation of a two node scenario tree with $t_0 = 0$ and $t_1 = T$ where node 0 is the root node and node 1 the only child node.

$$
\begin{aligned}
\underset{u,v,f,g}{\text{minimise}} \quad & \sum_{s\in\mathcal{T}(0)/\mathcal{L}} p_s \left\{ \sum_{n\in N_2} c_{sn}(u_{sn} - u_{a(s)n}) + \sum_{m\in M_2} d_{sm}(v_{sm} - v_{a(s)m}) \right\} \\
& + \sum_{s\in\mathcal{T}(0)/\{0\}} p_s \mu\big(\big[t_{a(s)}, t_s\big]\big)\mathbb{E}'_s \left[\sum_{n\in N} \gamma_{sn} g_{sn}(\xi_s) \right] \\
& - \sum_{s\in\mathcal{T}(0)/\{0\}} p_s \left\{ \sum_{n\in N_1} \pi_{sn} + \sum_{n\in N_2} u_{a(s)n}\pi_{sn} \right\} \\
\text{subject to} \quad & \\
& \sum_{n\in N} g_{sn}^{\mathrm{CO_2e}}(\xi_s) \le \mathrm{CO_2e} && \forall n \in N \\
& \sum_{n\in N(k)} g_{sn}(\xi_s) - \sum_{m\in M_+(k)} f_{sm}(\xi_s) + \sum_{m\in M_-(k)} f_{sm}(\xi_s) \ge \delta_{sk} && \forall k \in K \\
& g_{sn}(\xi_s) \ge 0 && \forall n \in N_1 \\
& g_{sn}(\xi_s) \le \alpha_{sn}\overline{g}_n && \forall n \in N_1 \\
& g_{sn}(\xi_s) \ge 0 && \forall n \in N_2 \\
& g_{sn}(\xi_s) \le \alpha_{sn}\overline{g}_n u_{a(s)n} && \forall n \in N_2 \\
& g_{sn}(\xi_s) \le \zeta_{sn} && \forall n \in N^{\mathrm{w,\,s}} \\
& f_{sm}(\xi_s) \ge -\beta_{sm}\overline{f}_m && \forall m \in M_1 \\
& f_{sm}(\xi_s) \le \beta_{sm}\overline{f}_m && \forall m \in M_1 \\
& f_{sm}(\xi_s) \ge -\beta_{sm}\overline{f}_m v_{a(s)m} && \forall m \in M_2 \\
& f_{sm}(\xi_s) \le \beta_{sm}\overline{f}_m v_{a(s)m} && \forall m \in M_2 \\
& u_{sn} \ge u_{a(s)n} && \forall n \in N_2 \\
& v_{sm} \ge v_{a(s)m} && \forall m \in M_2 \\
\text{where} \quad & \\
& \Xi = \big\{ \xi = (\alpha, \beta, \gamma, \delta, \pi, \zeta) : 0 \le \alpha, \beta \le e, e^\top(e-\alpha) + \\
& \qquad e^\top(e-\beta) \le B, \gamma \in [\underline{\gamma}, \overline{\gamma}], \delta \in [\underline{\delta}, \overline{\delta}], \pi \in [\underline{\pi}, \overline{\pi}], \zeta \in [\underline{\zeta}, \overline{\zeta}] \big\}
\end{aligned}
\tag{3.5}
$$

In the multi-stage model the final constraints on the investment decisions u_s and its predecessors $u_a(s)$ ensure that if a plant is built at an earlier stage of a branch, then it remains active throughout that scenario. The added difference in the objective function then ensures that a cost is only incurred at the node in which the facility is built, and no other. For example, consider a 3 stage linear tree, where plant 3 is built in the second stage, that is $u_{23} = 1$ but $u_{13} = 0$. In this case the objective function takes cost of c_{23} at node 2, but no other[2].

[2]As this is a linear tree the probability term p_2 is clearly equal 1 and is disregarded

Chapter 4

Sampling

In order to implement the models 3.4 and 3.5 and apply a the primal and dual LDR and CPLDR approximations the moment matrices $M := \mathbb{E}(\xi\xi^\top)$ need to be calculated. Thinking of the stochastic process ξ as a random vector on the augmented probability space allows the Monte-Carlo simulations for the random components of ξ to be implemented.

In this section a description of how each component is simulated in such a way that correlated samples are produced where necessary is given. It will end with an outline of an algorithm which can be used to simulate the stochastic processes.

4.1 Inverse Transform Sampling

Sample random variables are generated by inverse transform sampling. This states that if X is a continuous random variable with cumulative distribution function F_X, with inverse F_X^{-1} and if Y is uniformly distributed on $[0, 1]$ then $X = F_X^{-1}(Y)$ has cumulative distribution function F_X. It is easy to calculate the inverse of the cdf for the most common distributions and to generate uniformly distributed random numbers in any modern programming language. Hence, this method provides a simple way to sample any form of distribution.

4.2 Time

In order to sample the stochastic processes, and to generate correlated samples, a sample time point $t \in [t_{a(s)}, t]$ in years is first taken and then a sample for each random vector as an outcome from Ω_s is taken. The probability measure for time is the same Borel measure μ_s that defines the augmented probability space with parameter r_s. This should now clarify why it was previously stated that μ_s not only accounts for discounting, but also non-trivial time preferences. This enforces bias towards time points closer in the future and represents that earlier projections are more reliable than those after an extended period of time. The time point t is sampled using inverse transform sampling on the distribution function corresponding to μ_s.

4.3 Wind

In order to simulate the power generated from a wind turbine or wind farm it is necessary to first determine the wind speed at a particular location, and then to calculate the power produced in MWh from the turbines at this wind speed. To simulate the wind speed the Weibull distribution is parameterised with shape $\alpha = 2$ and scale η equal to the average hourly wind speed over a year for the location of the turbine. Taking the average over the year ensures that all seasonal changes that might reasonably occur are accounted for. The energy produced from a single turbine is then determined by 2.8 and the whole farm produces the sum of the power produced by the individual wind turbines. The wind speeds are again sampled by inverse transform sampling but with a 99% confidence interval to give a finite support.

4.4 Solar

As there will be a correlation between the hourly system demand and the global irradiance throughout the day it is important that the irradiance and demand are simulated in such a way as to represent this correlation. Logically, the relationship is easily visualised as demand is highest during the day, when it is expected that there is also more sunlight, as oppose to at midnight when demand is at its lowest, as is the sunlight.

Correlated irradiance and demand can be simulated by considering the historical hourly data over at least one year for both global irradiance of a region and demand. The amount of solar energy from a plant is first calculated for each hour over the year based on the global irradiance at that hour using formula 2.9. It is important that the total area of all solar panels in the solar farm is considered and not just the individual panel. This can then be used to construct an hourly curve for the amount of energy produced by the solar farm a any hour in the year. As the years are seasonal this same energy curve can be used over all years in the time horizon. Then a sample time point $t \in [t_{a(s)}, t]$ as described in 4.2 is taken and translated to the corresponding hour of the year. The energy in MWh produced by a solar farm at this hour is then obtainable from the solar energy curve. The corresponding expected demand at this hour, accounting for future growth, can then also be estimated from the load demand curve. This will then ensure that samples are correlated based on the historical data, whilst still being representing the nature of a stochastic process as the energy produced is still random and depends on time.

An illustration of such a technique is given in Figure 4.1. Whilst studies have been undertaken to identify if there exists a correlation between the wind and irradiance, not enough reliable evidence of this topic currently exists to develop an accurate model. Thus, in this project it is further assumed that wind and solar energies are independent of one another.

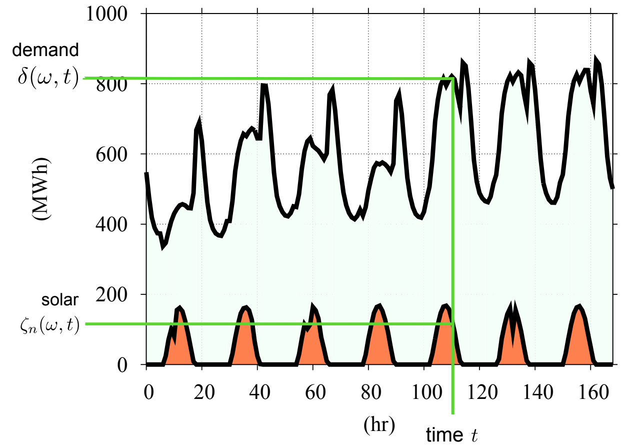


Figure 4.1: An illustration of sampling solar energy and demand

4.5 Demand

As was briefly mentioned in order to simulate demand knowledge of the load curve of the national demand for the first year of the total period $[0, T]$ is required. Then, using the same time point t that was used to sample the solar power, the hour in the year that this time belongs to can be calculated and - importantly for demand - in which year. Unlike global irradiance it is assumed that the load curve will grow over time and the given predicted growth rate and year sampled are used to predict the future hourly sample from the load curve. The regional demand is calculated as a factor of the total demand. This method will ensure demand is correlated with both solar and commodity prices.

4.6 Schwartz Type 1 Model for Commodity Prices

Commodity prices need to be simulated over a long time horizon and it is therefore best to use a model that represents the tendency of commodity prices to fluctuate around some mean value. It is also important that the simulated values are on some positive finite support, so a logarithmic price model needs to be used with a 99% confidence interval at all time points $t \in [0, T]$. The Schwartz type 1 model is used which assumes that the logarithm of the price follows a mean reverting process of the Ornstein-Uhlenbeck type. An Ornstein-Uhlenbeck process is an Ito process in S_t given by the stochastic differential equation

$$dS_t = k(\mu - S_t)dt + \sigma dW_t$$

where $k > 0, \mu$ and $\sigma > 0$ are the parameters and W_t is a Wiener process. The parameter μ represents the mean value, which the price has a tendency to follow, σ represents the volatility around the mean value caused by shocks from the Wiener process, and k is the rate at which the variable reverts towards the mean. Applying Ito's lemma 2.7 with

$y_t = x_t e^{kt}$ and solving the resulting stochastic differential equation (see Appendix B) results in the solution

$$S_t \sim S_0 e^{-kt} + \mu(1 - e^{-kt}) + \sigma\sqrt{\frac{1 - e^{-2kt}}{2k}} N_{0,1}$$

and thus S_t is normally distributed. The advantage of using such a process, as oppose to Geometric Brownian Motion[1], is that the drift term is non-constant. So when the current price is above the mean the drift will be negative, and when it is below the mean the drift will be positive.

The Schwartz type 1 model is a variation on the Ornstein-Uhlenbeck that results in a log-normal distribution of prices, ensuring positive values, defined by the Ito process

$$dS_t = k(\mu - \ln S_t)S_t dt + \sigma S_t dW_t \tag{4.1}$$

It is easy to verify that 4.1 is equivalent to an Ornstein-Uhlenbeck. As a result the solution to 4.1 is derived in a similar manner to before (see Appendix B), which results in

$$\ln S_t \sim e^{-kt} \ln S_0 + (1 - e^{-kt}) \left(\mu - \frac{\sigma^2}{2k}\right) + \sigma\sqrt{\frac{1 - e^{-2kt}}{2k}} N_{0,1} \tag{4.2}$$

and hence S_t is log-normally distributed. Correlated samples of commodity prices can then be simulated by first sampling a vector z of independent standard normally distributed random variables. Then, the Cholesky decomposition of the covariance matrix is used to transform the vector z into correlated samples $u = Lz$. The entries in this vector can then be used along with the time t and the other parameters to the Schwartz models to evaluate the solution of 4.2 for the respective commodity. This also ensures a correlation with demand indirectly through the dependence on time.

As with any stochastic model it is essential to estimate the input parameters and these approximations are commonly based on historical data. Indeed, it is not the aim of this project to develop an accurate method for estimating such parameters, as this would usually be the role of a financial analyst or economist. Thus, for simplicity the parameters k, σ and μ were estimated by a linear regression on the historical stock prices. An example of this is provided in Appendix C.

4.7 Plant and Transmission Line Failures

Finally, samples need to be provided for the failures of plants α and transmission lines β. All failures are assumed to be independent of one another and it is required that one can

[1] Geometric Brownian Motion is a logarithmic stochastic process defined by a Wiener process

specify the probability $p_i \in [0, 1]$ of failure at each facility. This was done via representing each failure as a Bernoulli trial with probability p_i of failure. After the facilities which fail are known the severity of their failures needs to be calculated in such a way that the failure constraint $e^\top(e - \alpha_s) + e^\top(e - \beta_s)$ is satisfied.

To do this a sample b representing the total failures is sampled as a uniformly distributed random variable on $[0, B]$. This is then distributed on a weighted basis by taking samples for the failed α_i, β_j as uniformly distributed random variables on $[\max(0, 1 - B), 1]$ and adjusting as follows

$$\begin{aligned}
\text{total} &= \sum_{i=1}^{n_\text{f}} \alpha_i + \sum_{j=1}^{m_\text{f}} \beta_j \\
\alpha_i &= \max(1 - \frac{\alpha_i}{\text{total}} b, 0) && \forall i = 1 \ldots n_\text{f} \\
\beta_j &= \max(1 - \frac{\beta_j}{\text{total}} b, 0) && \forall i = 1 \ldots n_\text{f}
\end{aligned}$$

An example illustration of the support for two failed plants can is given in Figure 4.2.

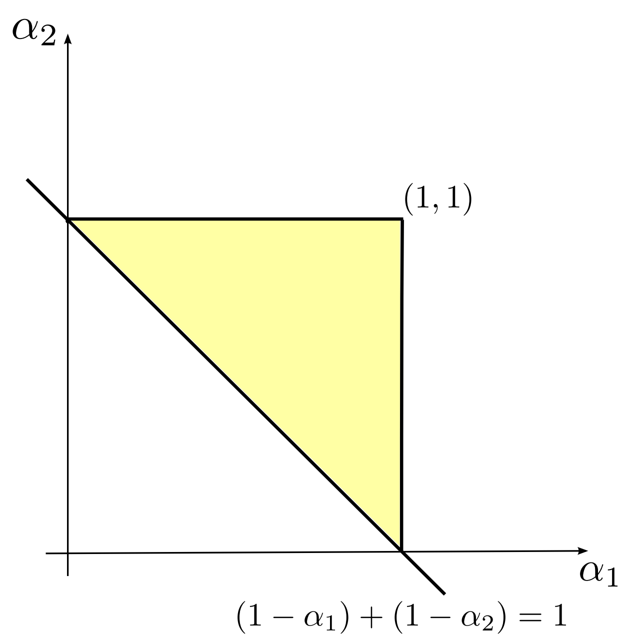


Figure 4.2: Support of two failed plants, $B = 1$

It could be argued that the use of Bernoulli trials has little impact on the actual performance of the system. The alternative here being to simply sample uniform failures for all plants and transmission lines. However, this method gives added flexibility as you can specify whether certain plants or links cannot fail, that is if $p_i = 0$.

4.8 A Sampling Algorithm

At this point a method for sampling each of the individual components of ξ has been established. All the components discussed in this chapter can be combined into Algorithm 1.

Algorithm 1 Simulation of stochastic processes

$i = 1$
while $i \leq \max$ **do**
  for all non-root nodes $s \in \mathcal{T}(0)/\{0\}$ **do**
    sample $t \in [t_{a(s)}, t_s]$
    for all $n \in N^{\text{w, s}}$ **do**
      sample (ζ_{sn}) at time t
    end for
    sample total demand δ_s for time t
    calculate δ_{sk} for each region
    sample $z \sim \mathcal{N}(0, I)$
    $u = Lz$
    $S_i = \text{Schwartz}(u_i, S_{0_i}, k_i, \mu_i, \sigma_i)$
    sample failures $(\alpha_{sn})_{n \in N}$ and $(\beta_{sn})_{m \in M}$
  end for
  $i{+}{+}$
end while

This algorithm then provides us with a complete simulation for our stochastic process ξ. The samples generated are used to calculate the moment matrices, required for the LDR and CPLDR approximations, that will be used to obtain a tractable linear program for our capacity expansion models.

Chapter 5

Numerical Results

5.1 Implementation

It was the aim of this project to design and implement the models discussed in Section 3 using the LDR and CPLDR approximations and then to evaluate the models through a series of numerical results. Here, the overall system implemented for this purpose is briefly described.

As the purpose of this project was to develop the models and evaluate them, it was decided that it was not necessary to develop a front end. Furthermore, this would not have been very user friendly as the programs sometimes take very long times to respond and it would not have been an efficient use of time. A parser in Java was developed which took as an input a text file containing the configuration of the current electricity system. As large amounts of data is required for the load demand curve and the hourly energy produced by solar farms (see Section 4.4), this data is stored in a separate text file referenced by the main input file. The input text file contains all the information on the system which is required by the model. This includes everything from the parameters for simulation to the location and capacity of generators and transmission lines. This parser then formulated the system as a model of the form 3.5 and output this to a text file. It also implemented all the simulations discussed in Section 4 and output samples for the random variables to a separate text file. Java was chosen as the structure of the electricity system and the scenario tree naturally extend to an object orientated language.

The two output files generated from the Java program were then passed as an input to a C++ parser. This C++ program reformulated the models into linear programs using the techniques detailed in Sections 2.2 and 2.2.2. This part of the system was already developed by Angelos Georghiou of the Department of Computing at Imperial College. An industrial linear programming solver, IBM ILOG CPLEX v11.2.1, was then used to solve these linear programs.

All the experiments in this section were ran on a Intel Core2 Duo 3.00GHz machine with 4GB RAM running Ubuntu v9.04.

5.2 El Salvador

The electricity sector in El Salvador has been the subject of much restructuring since the early 1990's. In this section we give an overview of the electricity markets in El Salvador and finish with a series of numerical results to evaluate our models.

5.2.1 History

Until the 1990's the electricity sector in El Salvador was dominated by the Government-owned monopoly Commission of the River Lempra (CEL), who controlled generation, transmission and distribution. In 1996 the General Law of electricity was passed which began the restructuring of the electricity sector. The law split the main activities of CEL to introduce the privatisation of electricity generation and distribution. The aim of this was to increase investment in generation capacity to meet the countries increased demand for energy and to keep energy prices low for the consumer. As a result of the restructuring a non-profit organisation the Superintendent of Electricity and Telecommunications (SIGET) was appointed as a regulatory body to enforce compliance with the laws of the energy sector. SIGET appointed a private company, the Unidad de Transacciones (UT), to oversee the operation of the transmission system and administration of the wholesale electricity market. As the UT is not allowed to participate in the buying or selling of energy its role is to ensure fair bidding within the market. As good practice they have made it part of standard operating procedure to post hourly data on the price and dispatch of energy on their website.

Since the restructuring only one government-owned company, the El Salvador transmission company (Etesal), is responsible for the maintenance and expansion of the transmission system. In more recent years a higher emphasis has also been placed on introducing more environmentally friendly initiatives and in 2006 the National Energy Council (NEC) was created. The goal of the NEC was to analyse the energy sector in El Salvador and promote the use of renewable resources and the efficient consumption of energy. El Salvador is an active participant in the SIEPAC project which coordinates the electricity systems throughout Central America and is connected by two $230Kv$ transmission lines to the neighboring countries of Guatemala and Honduras. From 2000-2006 the average annual increase in maximum demand has been 2.6% whilst average increase in installed generation capacity has been 2.9% [13, p. 25]. Whilst this sounds impressive the system is still highly vulnerable to any power outages that may occur. Demand is currently expected to grow at a rate of 5% annually [23, p. 4].

5.2.2 The El Salvador Electricity Grid

SIGET posts an annual document, the "Boletín de Estadísticas Eléctricas", on the El Salvador power system and this was used to model the system with a configuration as outlined in Figure 5.1. The current installed capacity and the transmission links was obtained from the 2008 bulletin [42, Tables 1 and 29]. In 2008 the total generation

Power Type	Location	Owner	Capacity (MW)
Geothermal			
G1	Ahuachapán	LAGEO	95
G2	Berlín	LAGEO	109.4
Thermal			
T1(Gas)	Acajutla	Duke energy	322.1
T2(Gas)	Sonsonate	CASSA	60
T3(Oil)	Santa Ana	CESSA	32.6
T4(Gas)	Ateos	GECSA	11.6
T5(Coal)	Talnique	Inversiones Energeticas	51.2
T6(Gas)	Opico	Energia Borealis	13.6
T7(Coal)	Nejapa	Nejapa power company	144
T8(Oil)	Nejapa	Hilcasa Energy	6.8
T9(Gas)	Soyapango	Duke energy	16.2
T10(Gas)	Soyapango	Textufil	44.1
T11(Gas)	Santo Tomás	Ingenio el Angel	22.5
T12(Oil)	San Rafael Cedros	Ingenio la Cabana	21
Hydro			
H1	Guajoyo	CEL	19.8
H2	Cerrón Grande	CEL	172.8
H3	5 de Noviembre	CEL	99.4
H4	15 de Septiembre	CEL	180
		Total Capacity (MW) =	1422.1

Table 5.1: Current configuration of El Salvador electricity grid [42]

capacity in the system was 1422.1MW divided between a 13 companies as detailed in Table 5.1. The current mix of electricity generation is: hydropower 33.19%, thermal 52.44% and geothermal 14.37%.

Configuration

In order to model the energy markets in El Salvador the system is viewed from the perspective of an independent system operator who coordinates the operation of all plants in order to minimise cost and emissions of the system as a whole. Whilst this might seem unrealistic this would be the perspective taken by a Government Body seeking to advance the whole system and not just independent companies. In order to configure the model information is needed on the demand, greenhouse gas emissions, investment prospects, operating costs and climate conditions. Whilst a handful of data is available on the electricity system in El Salvador, most of this information had to approximated based on data from other sources.

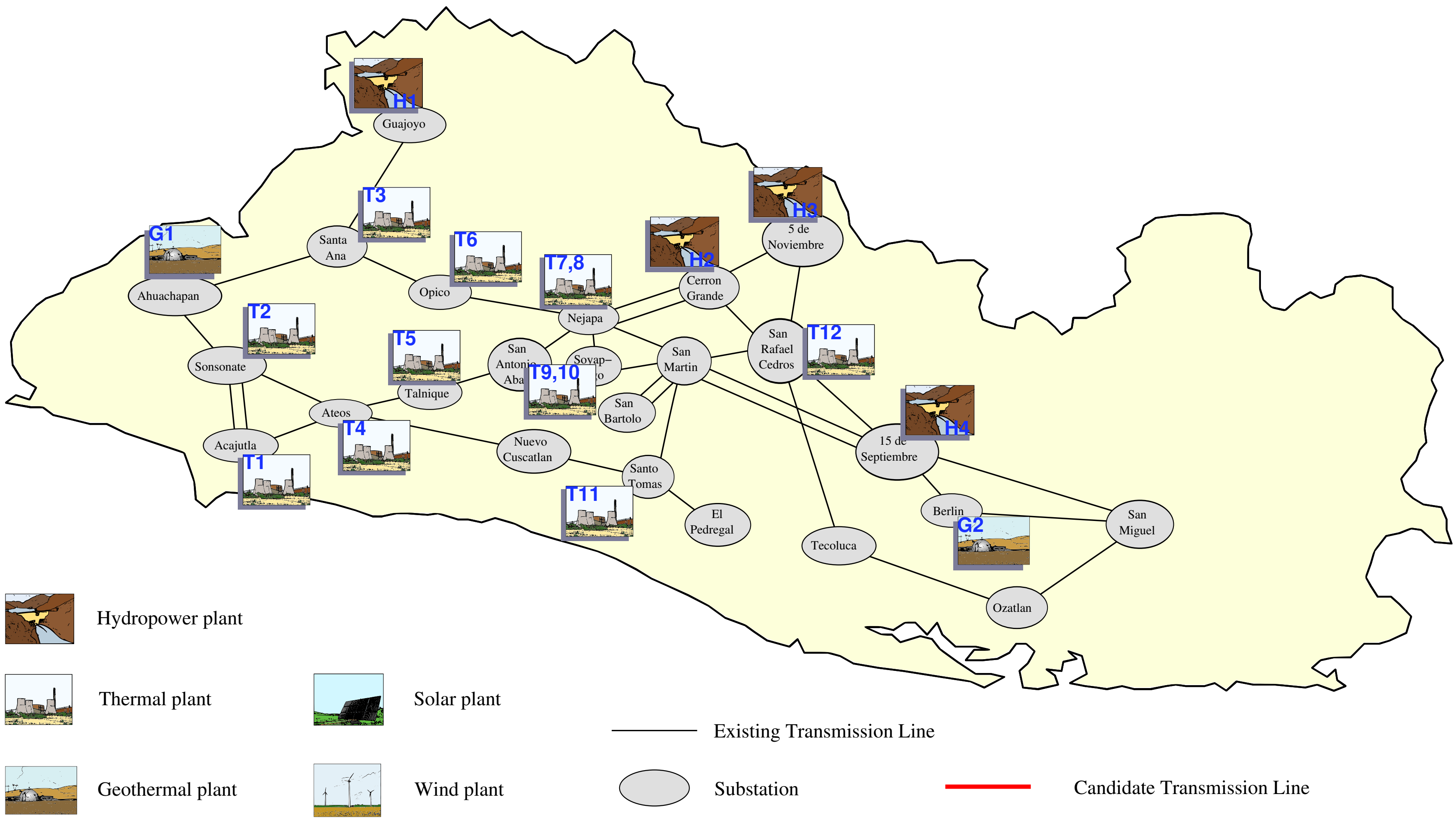


Figure 5.1: Current configuration of the electricity grid

As the UT posts hourly data on the national demand of the system a load curve for the year 2009 was constructed, and the corresponding load duration curve is given in Figure 5.2 [43]. Approximate regional factors for demand were made by considering the amount of energy consumed in each region over 2008 as a percentage of total energy consumed, which was available from the SIGET 2008 bulletin [42, Table 30]. This indicated that whilst most of the energy was generated outside of the capital, more than 50% of the demand was within the vicinity of San Salvador.

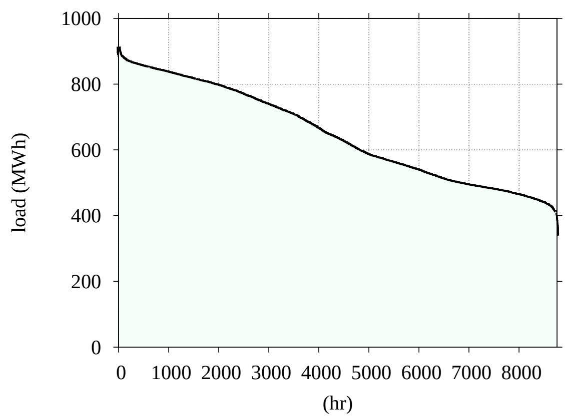


Figure 5.2: El Salvador load duration curve (2009) [43]

The rate at which thermal and geothermal plants emit carbon in CO_2e per MWh was obtainable from Carbon Monitoring for Action (CARMA). CARMA is a database containing information on the carbon emissions of over 4,000 power companies worldwide [9]. Hydro, wind and solar plants were assumed to have minimal carbon emissions as to make them negligible. Unfortunately, CARMA does not post the yearly allowance of carbon credits for the cap and trade system, and this had to be roughly estimated as a percentage of the plants total emissions for the year 2007. A rough approximation of 30% was used.

Investment proposals for new hydro, geothermal and thermal plants in El Salvador were used from a previous study undertaken in 2006 by Jun Abraham [1]. A number of investment prospects in wind and solar farms were identified based on the regional climate outside of the vicinity of San Salvador, due to the space requirements of such power sources. Data on the average wind speeds of different regions in El Salvador was obtained from the National Renewable Energy Laboratory (NREL) and an illustration of this is provided in Appendix E [34]. A number wind projects were chosen in areas where the average wind speeds exceeded $7m/s$. All the wind turbines are taken as $50m$ in diameter, at 40% efficiency. The air density was chosen to be equal to that at sea level a value of $1.2kg/m^3$. Similarly a number of selected solar projects were proposed based on the global irradiance for selected regions of El Salvador outside of the capital. The hourly data for these regions between 1998 - 2002 was obtainable from the State University of New York (SUNY) [39]. This allowed identification of several options for solar energy and simulation the of stochastic nature of solar energy as outlined in 4.4. A table of all the candidate plants is given in Table 5.2. A map of the solar intensity throughout the year is provided in Appendix D courtesy of NREL [35].

The investment costs for these plants had to be estimated from a number of sources

Power Type	Location	Capacity (MW)	Cost $
Geothermal			
GC1	Ahuachapán	50	111,485,000
GC2	Berlín	50	111,485,000
GC3	San Rafael Cedros	55	120,036,000
GC4	Berlín	44	100,186,000
GC5	Ozatlán	50	111,485,000
GC6	San Miguel	50	111,485,000
Thermal			
TC1(Coal)	Ateos	100	91,670,000
Hydro			
HC1	Guajoyo	243	322,395,000
HC2	San Miguel	59	167,345,000
HC3	15 de Septiembre	23.4	97,287,000
Wind			
WC1	Ahuachapán	11.4	24,216,000
WC2	Ahuachapán	7.5	13,446,000
WC3	San Martín	8.8	24,216,000
WC4	Guajoyo	15	24,216,000
WC5	San Martín	9	28,599,000
Solar			
SC1	Ahuachapán	7.3	6,677,000
SC2	Sonsonate	9.6	8,855,000
SC3	Cerrón Grande	7.4	6,677,000
SC4	5 de Noviembre	49.7	40,468,000
SC5	15 de Septiembre	14.6	13,140,000
SC6	San Miguel	12.7	11,009,000
	Total Capacity (MW) =	877	

Table 5.2: Candidate plants

accounting for the levelised cost and economies of scale [19, 45, 3]. A table of the average costs used for all plants is provided in Table 5.3. It should be noted that all costs in this section are rough estimates and are should not be taken as accurate quotations. These were merely used to allow calibration of the model.

Power Type	Average Cost
Thermal	\$1,108,000 per MW
Geothermal	\$ 2,182,000 per MW
Hydro	\$ 3,025,000 per MW
Solar	\$50 per m^2
Wind	\$ 700 per MW

Table 5.3: Approximate average investment costs

A number of investment opportunities in new transmission lines were also estimated. These were based on a weighted basis of the distance from the two end nodes of the transmission lines. The proposals are outline in Table 5.4.

Line	From Node	To Node	Capacity (MW)	Cost \$
LC1	Nejapa	San Antonio Abad	150.6	5,800,000
LC2	Ateos	Nuevo Cuscatlán	150.6	12,083,000
LC3	Acajutla	Ateos	203	18,850,000
LC4	Tecoluca	El Pedregal	75.3	4,350,000
LC5	Guajoyo	Santa Ana	150.6	17,400,000
LC6	Soyapango	Nejapa	150.6	7,250,000
LC7	San Bartolo	San Martín	150.6	6,766,000
LC8	Ozatlán	San Miguel	75.6	11,116,000

Table 5.4: Candidate transmission lines

All the candidate plants and transmission lines can then be identified as in Figure 5.3.

Operating and maintenance costs consist of a fixed and variable portion and it is assumed that variable costs for hydropower, wind and solar plants are negligible. The O&M costs for geothermal and thermal plants was estimated based on data provided by the Energy Information Administration (EIA) and the amount of fuel needed to produce one MWh of energy for different fossil fuels [18, 36]. These fuel costs were based on a plant that is 100% efficient, and thus they are multiplied this by an efficiency factor. A typical efficiency factor for coal and oil fired plants is around 33%, whilst gas fired plants are slightly more efficient around 50% [22].

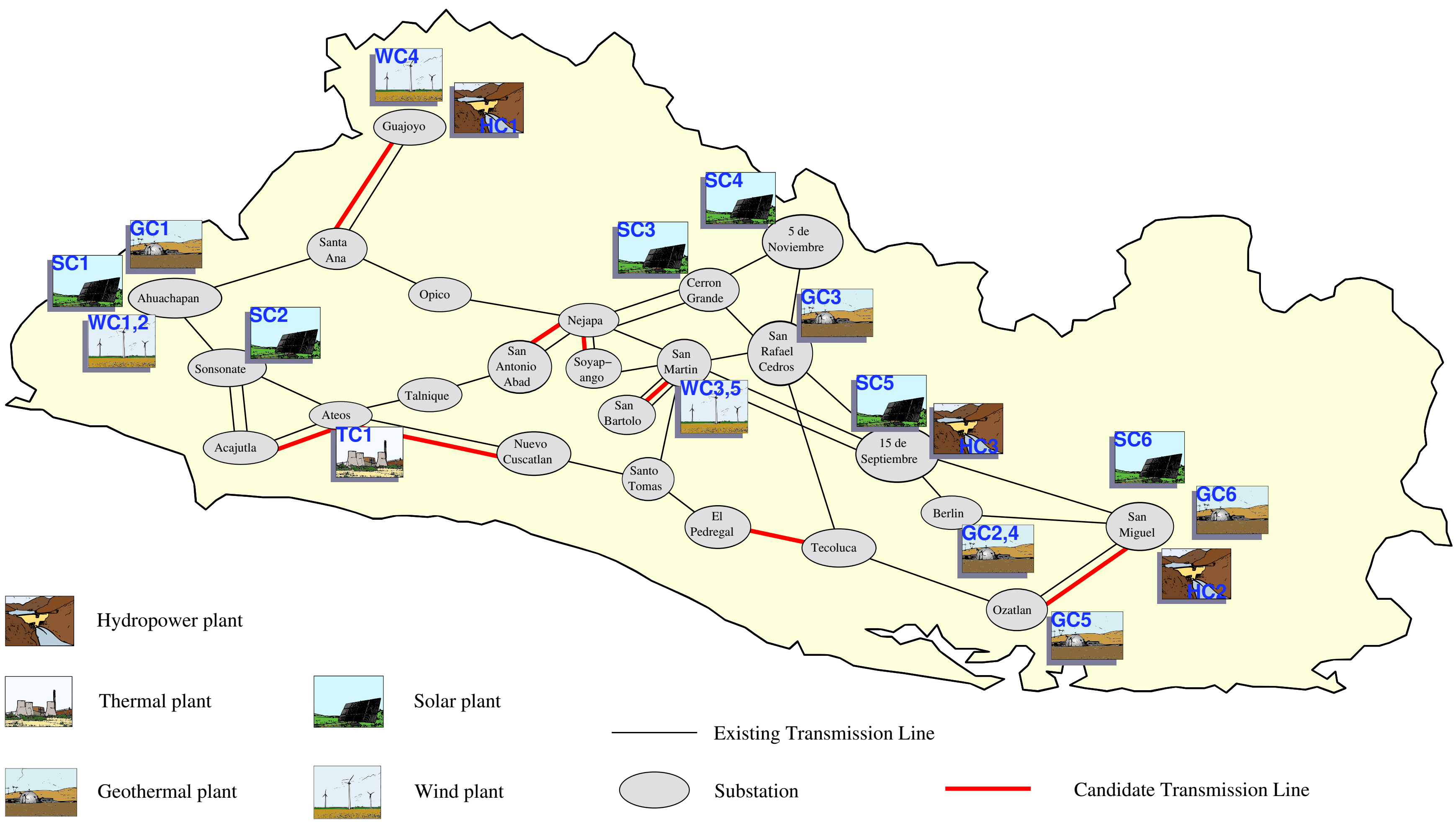


Figure 5.3: Candidate infrastructure for El Salvador

Commodity	S_0 ($)	k	σ	μ
Oil	74.49	0.13	0.32	4.9
Gas	192.51	0.41	0.53	5.74
Coal	89.04	0.08	0.28	5.03
Carbon	0.002	0.12	1.03	-3.26

Table 5.5: Schwartz parameters

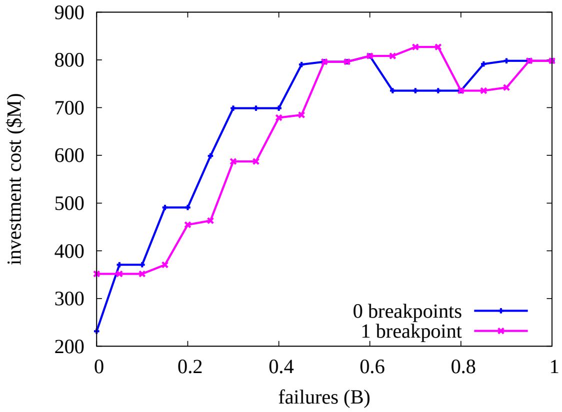
Finally, to parameterize the Schwartz models for oil, gas, coal and carbon prices a linear regression as outlined in Appendix C was performed. The historical monthly prices of WTI crude oil, Australian thermal coal & natural gas from March 2000 to February 2010 provided by the International Monetary Fund (IMF) were used [26]. Historical monthly prices of CCX Carbon Financial Instruments (CFI) were taken from January 2004 to February 2010 provided from the Chicago Climate Exchange [12]. The parameters calculated from the linear regression and used in these results - unless the results state otherwise - are given in Table 5.5.

5.3 Evaluation

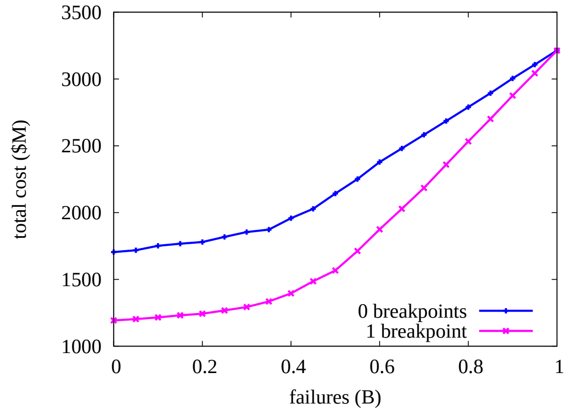
In this section a series of experiments are undertaken on the El Salvador system. The categories to consider in the evaluation are primarily the impact on total cost and investment costs. Other considerations would be the impact of investment decisions on expected carbon emissions and the congestion in the transmission links throughout the day. All results are for the upper bound of our LDR approximation unless otherwise stated. The results are based on a linear scenario tree with one investment period at time $t_0 = 0$ and a period of growth from 2009 - 2024 of 3.06% unless otherwise stated. All demand projections in this section are roughly estimated based on the projections provided by Jun [1]. This increase in system demand results in a maximum expected demand of $1402MWh$ in 15 years. Where the results do not depend on the change in variable B, a fixed failure rate of $B = 0.2$ has been chosen. A constant interest rate of 5% has been fixed throughout all experiments. Where CPLDR have been used the breakpoints are distributed uniformily over the domain of the support for the components of ξ.

The first experiment was to determine the impact of requiring a more resilient system to plant and transmission line failures. This was achieved by increasing the value of B in the support of Ξ in model 3.5. The results given in Figure 5.4b illustrate the increased benefit obtained by adding 1 breakpoint in reducing the total system cost. In Figure 5.4a for B between 0.6 and 0.8 there is a clear trade off between higher investment costs and reduced operating costs, comparing the results for 0 and 1 breakpoints. It can also be seen that the effect of the extra breakpoint is reduced as the number of failures are increased and more is said about this in section 5.3.1.

When $B = 1$ the best choice of plants to build was HC1, HC2, GC3 and GC6 and all candidate links except from LC8 were built. This highlights the vulnerability of the transmission network to failures. Under this scenario the total system capacity would be $1829.1MWh$. Conversely, when a failure of $B = 0.2$ was selected there were differing



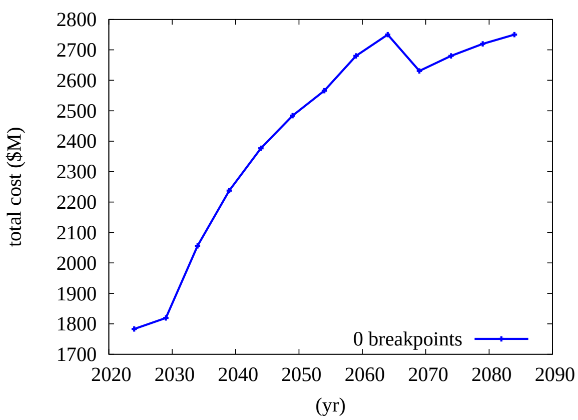
(a) Investment costs



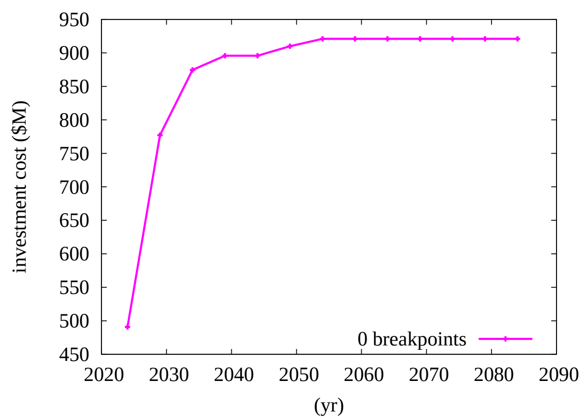
(b) Total costs

Figure 5.4: The price of reliability

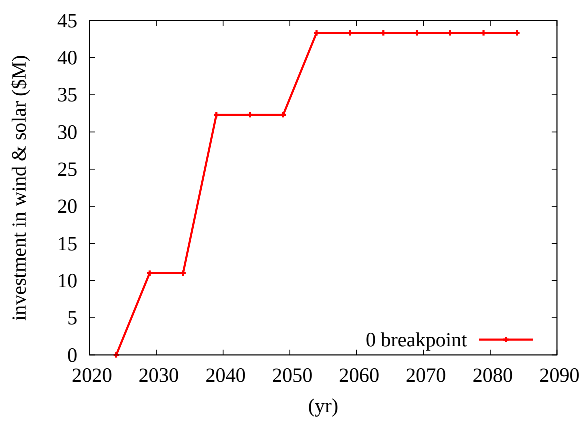
decisions from adding an extra breakpoint and resulting system capacity varied between $1622.1MWh$ and $1720.1MWh$, for 1 breakpoint and 0 breakpoints respectively.



(a) Total costs



(b) Investment costs



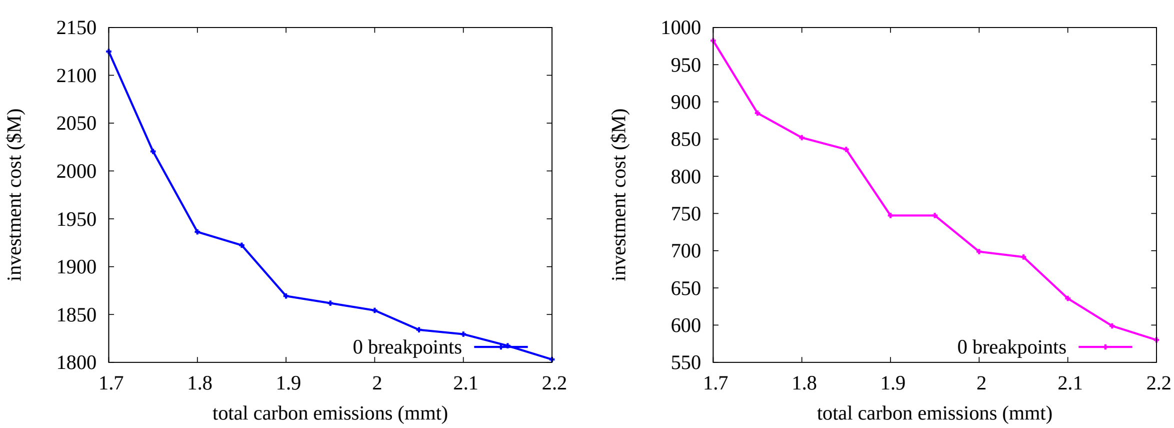
(c) Investment costs in wind and solar farms

Figure 5.5: Wind and solar time horizon

The results in Figure 5.5 are based a three-stage scenario tree with one extra node appended to our two-stage formulation, observing an operating period from 2024 until time T. The investment decisions indicate a shift away from expensive thermal plants initially by investing in plants HC1, HC2, GC3, GC5 and GC6. As the time horizon is extended further the minimal operating cost of wind and solar farms makes them viable investment options and after 45 years it is beneficial to build WC3 and SC6 initially. This results in a much cleaner energy mix than before, with a generating mix of: hydropower 40.72%, thermal 39.23%, geothermal 18.91%, wind 0.46% and solar 0.67%.

These results do support the general consensus that renewable energies are more unreliable and take longer to payback. However, part of the reason for this is also because of the naivety in the way that wind and solar farms simulated in the models. The possibility of storing the energy generated from these resources, so that energy can be supplied to the grid even when weather conditions are against us, is not considered. This is a justified assumption though as it is needed to ensure the constraints are independent of time, and without this assumption this would be violated. Further restrictions could be imposed over the support of Ξ in order to further exploit the characteristics of these resources which could make them more attractive options. These are all fairly complex tasks, and for the purposes of this project the assumptions made are reasonable and do give some insight into potential renewable investment. This topic will be discussed in more detail in Section 6.2.

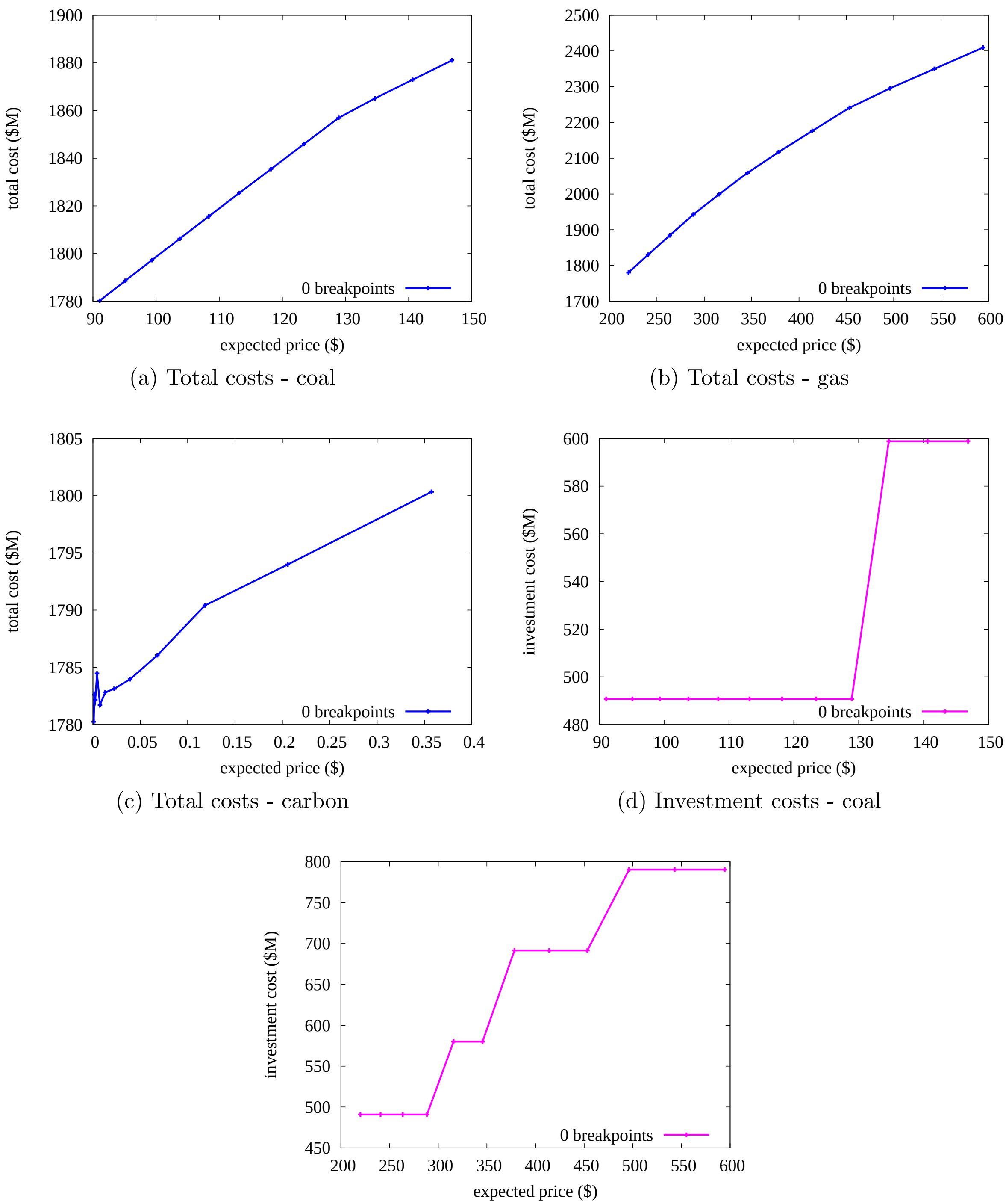
The effect of a carbon emissions cap and trade system is now analysed. Consider the rate of carbon emissions constraint in model 3.5 and the impact of limiting amount of CO_2e emitted per year in terms of million metric tonnes (mmt). When carbon is a emitted at a maximum rate of $2.2mmt/yr$ the plants built are HC1, GC3 and GC6, whilst restricting this rate right down to $1.7mmt/yr$ results in the addition of plants HC2, HC3 and GC5 as well. The resulting energy mix indicates a shift away from fossil fuel plants towards renewable resources with a lower carbon output. The reason why higher investment in both wind and solar farms is not seen are identical to the previous comments.



(a) Total cost (b) Investment cost

Figure 5.6: The price of reduced carbon emissions

Figures 5.7a, 5.7b and 5.7c illustrate the impact of fuel prices on total costs and the effect of investment costs are given in Figures 5.7d and 5.7e. These were obtained by increasing the respective parameter μ in 4.2 and taking the expectation with respect to a fixed time point. For these results the expected commodity prices were taken after 10 years. The impact of various fuels depends largely on the number of turbines of such a type.



(a) Total costs - coal

(b) Total costs - gas

(c) Total costs - carbon

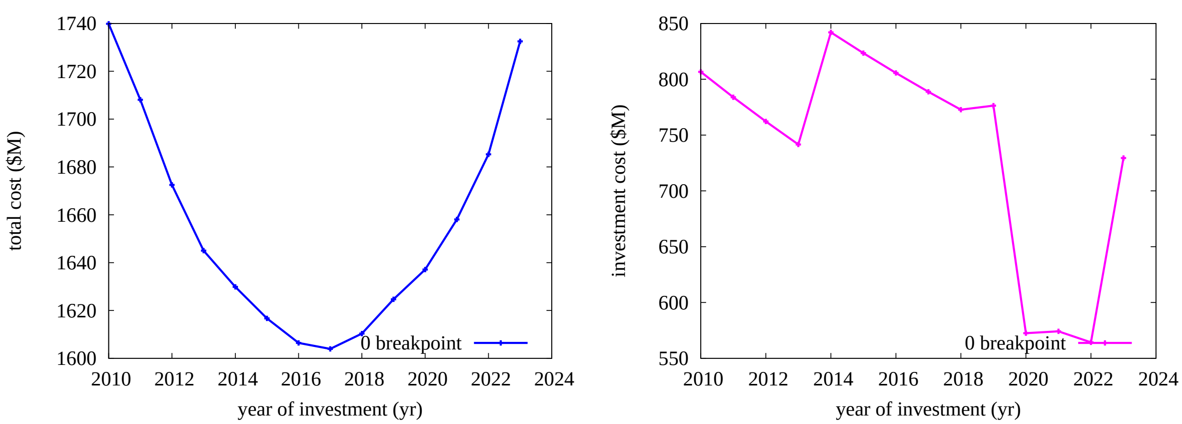
(d) Investment costs - coal

(e) Investment costs - gas

Figure 5.7: Impact of increased commodity prices

With all prices at their lowest expected cost the earlier results for $B = 0.2$ indicate that plants HC1 and GC3 were built. Even when the carbon and oil price was increased this had no impact on the investment decisions. When expected coal price exceeds $130 the additional construction of GC6 is added. The high expected gas prices exceeding $490 results in a additional investment in plants TC1, GC5 and GC6. This highlights the dependence of the system to volatile gas and coal prices and indicate that it may be beneficial to consider a more diverse energy mix to reduce this risk. The most likely choice of candidate plants to accommodate for both volatile coal and gas prices is would therefore be plants HC1, GC3, GC5 and GC6.

It is important that the time of construction of new plants is made in such a way as to still meet the expected demand, whilst benefiting from the discounted cost of delaying investment decisions to a later time. Consider a 3 node linear tree over the period 2009 - 2024 where the second investment node is placed between 2009 and 2024. The results in Figure 5.8 indicate the effect of varying this second node and indicate the best time to invest. At all variations of the three-stage scenario tree plant HC1 is always built in the 1st stage as is link LC5. The savings come by deciding on whether to build plant GC3 and links LC2 and LC3 at a later time point or in 2009 along with the others.



(a) Total cost

(b) Investment cost

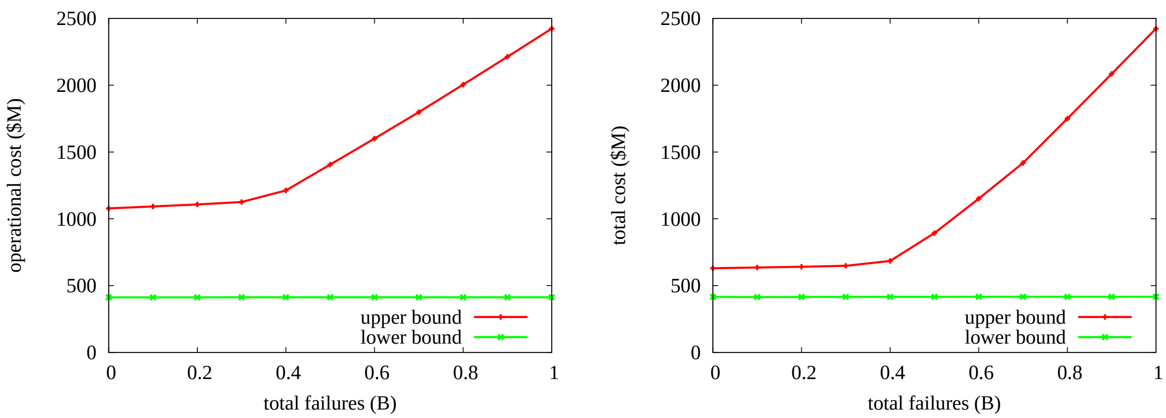
Figure 5.8: The best time to invest

Notice how the best year to invest is 2017 to reduce the overall total system cost, however this results in a total investment of $788m. Delaying the investment decisions until 2020 results in a total investment cost of $572m, at only slightly higher expected overall cost of $33m. All these factors would need to be considered in deciding whether it is worthwhile spending an extra $216m in order to benefit from a significantly smaller later saving. This also nicely illustrates the trade off between initial costs and later savings due to the nature of the economics of the different power types.

5.3.1 Accuracy of Linear Decision Rules

The investment decisions rely heavily on the accuracy of the LDR approximations to the operational schedule of the power system. As stated before this can be approximated through the gap between our upper and lower bound results. If it is assumed that plants

HC1, HC2, GC3 and GC6 are built then results for the lower and upper bound can be obtained which only account for the expected operating costs. The resulting model is one without any binary investment variables and only approximates the linear decision rules. The results in Figure 5.9 show that adding a breakpoint has very little effect on the approximation when the failure rate B is high. This could be due to the shape of the true decision rule and indeed a better approximation might be possible with a more precise placing of the breakpoint. To visualise this consider that our LDR must include the region that the true decision rules does. If the true decision rule was piecewise linear, with a global maximum at x_0, then adding a breakpoint at x_0 results in a very accurate approximation. However, if the breakpoint is not placed at x_0 our approximate LDR will not approximate the decision rule as well as the more precise breakpoint and may result in an approximation closely representing that without any breakpoints at all.



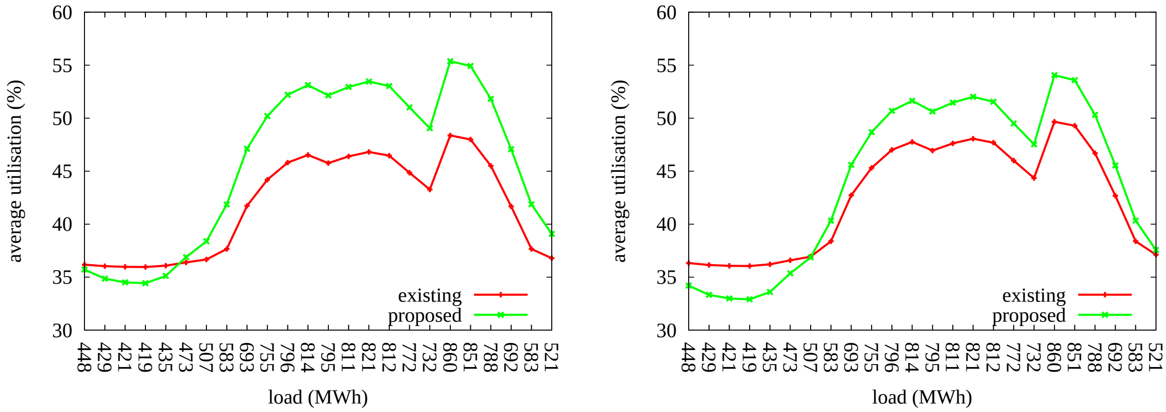
(a) 0 breakpoints

(b) 1 breakpoint

Figure 5.9: Accuracy of the LDR's

As a final consideration, and to illustrate the flexibility that our model 3.5 allows, we consider a three stage scenario tree with $B = 0.5$. We observe a period of growth from 2009 to 2019 where expected demand could either be high, at 3.77%, or low, at 2.99% each with equal probability. The results indicate that the best investment options are plants HC1 and GC3, and links LC4, LC5 and LC6 with a resulting system capacity of $1720.1MW$. Suppose now that we want to use these results to analyse basic characteristics about our proposed electrical grid to the existing. We can consider the new system in which plants HC1, GC3 and links LC4, LC5 and LC6 are existing already and use the LDR approximations given for the operating schedule to investigate the transmission line utilisation. The average utilisation of the available transmission lines in the two systems could be compared for an arbitrary projected load demand cycle or based on historical data. We used the load demand data provided from UT on for the single day of 5-Jan-2009 [43]. The results in Figures 5.10a and 5.10b illustrate the transmission line utilisation under a fully functional system and with a failed line between 15 de Septiembre and San Martín (i.e $\beta = 0.5$) as this was the highest utilised transmission line for the given date.

It might also be of interest to analyse not just the capacity mix of the new system as we have done previously, but the amount of power generated from different energy sources throughout the day. Considering the same date as with the transmission line utilisation,



(a) Fully functional system

(b) Failed transmission line

Figure 5.10: Comparison of transmission line utilisation

the total load demand throughout the day was $15,811MWh$. Figure 5.11 shows how the load demand would have been achieved from each source in the current and proposed system with plants HC1 and GC3 in place.

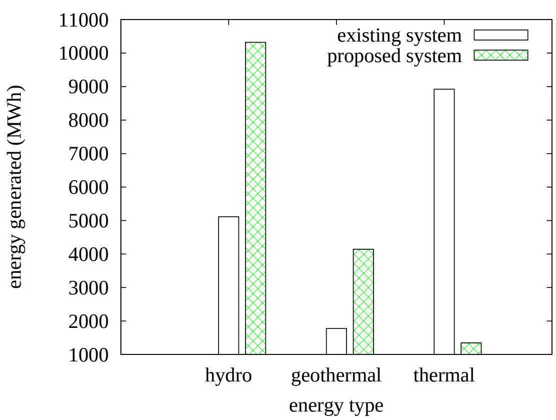


Figure 5.11: Comparison of generated energy

This set of results indicate that under higher system demand the traffic in the system actually increases under the proposed system, with the new plants and transmission lines in place. This is justified by the high demand for energy in the vicinity of San Salvador, with the majority of energy being supplied from rural areas requiring transportation. As the new plants result in a lower mix of expensive thermal energy throughout the day, and as transmission costs are ignored, it is more beneficial to transport the cheaper energy than to generate expensive thermal energy close to San Salvador. The high utilisation under all circumstances indicates the reliability of El Salvador to the transmission system. This means that satisfying load demand is highly vulnerable to transmission losses. Although not necessarily the least cost route, investigation should be undertaken into the construction of even more transmission lines to reduce the potential risk of a blackout.

It is worth taking a moment to discuss the solution time and complexity the problems within these results. This won't be discussed in depth as it is not a real concern of this project as capacity expansion proposals tend to span over many years, so anything under a week to solve the system would be viable. To solve a two stage model with no breakpoints, as in Figure 5.4, the upper bound solution time took anything from 2 minutes to more than 30 minutes. Adding a breakpoint to this the solution time increased rather significantly and times ranged from 30 minutes to anything up to 3 hours. A scenario tree with more than 2 nodes increases the solution time more significantly still, with several hours not uncommon. The lower bound for most of the problems above with binary investment variables was intractable and thus we cannot directly measure the error introduced by our LDR approximations with binary variables. It is still unsure as to the exact reason why this is however, a calculated guess would be at the large number of random variables in the model for the system failures. Whilst this is an inconvenience it could be argued that the lower bound solution would never actually be used as it is, by construction, an infeasible solution. In the LDR comparison of the upper and lower bounds in Figure 5.9 a lower bound was obtained by removing the binary decisions and the solution time was significantly longer than for the upper bound. Especially when more than 1 breakpoint is added the time ran into a number of days. The model with the highest complexity, for which a feasible upper bound solution was obtainable, was a 4 stage scenario tree which consisted of 58 binary variables, 258 decisions rules and 303 random variables. The results for this model were not included in the above analysis as they added nothing new to the existing results and were similar to those in 5.11.

Chapter 6

Conclusion

6.1 Qualitative Evaluation

In this project Linear Decision Rule and Continuous Piecewise Linear Decision Rule approaches to solving a multi-stage stochastic electricity capacity expansion problem have been explored. Both methods result in the solution of a linear program which can be solved by a variety of industrial solvers. A multi-stage capacity expansion model that allows for multi-objective optimisation with environmental constraints has been successfully developed. The inclusion of wind and solar farms as alternative investments compared to the more traditional capacity expansion models gives this model added flexibility. Whilst the simplified modeling of wind and solar resources does allow for their inclusion in the capacity expansion problem, it is clear that their are inherent inaccuracies in the assumptions made about their characteristics. An implementation has been created which transforms the electricity grid in question into a model of the form 3.5 and provides a simulation of the stochastic processes, including a Monte-Carlo simulation of commodity prices via the Schwartz model.

A detailed evaluation of the model has been undertaken through a case study of the El Salvador electricity system. The model was provides a flexible approach to various investment opportunities depending on the changing requirements of the system operator. Furthermore, the LDR approximations can be used to analyse more fundamental properties of the electricity grid as a result of the investment decisions.

The results indicate the downfall of the lower bound in its intractability for more complex problems with a high number of binary decision variables. Indeed, even when a lower bound was obtainable without binary decisions the approximate error introduced for such a complex problems was rather high. However, using CPLDR in some cases significantly reduced this gap and although precise placement of the breakpoints may reduce this gap still further.

6.2 Future Work

It is no surprise that capacity expansion problems have a vast mix of factors and complications which mean that a number of assumptions need to be made to obtain a tractable model for these problems. This project highlights a number of areas for further research which would give better understanding of the characteristics of electricity generation and improve the accuracy of the model

- It is clear that research needs to be undertaken into improving the tractability of the lower bound. This would allow better estimation of the accuracy of the primal and dual LDR approximation.
- A technique for more precise placement of the breakpoints with CPLDR could be developed and tested. This would optimize the accuracy of the CPLDR approximation and lead to more efficient use of breakpoints.
- Further work needs to be undertaken on the modeling of wind and solar resources. This could be in the form of restricting the support to ensure that, for example, it cannot happen that solar energy is at its lowest, generally during the nighttime, when demand is at its highest, which occurs generally during the daytime or early evening. Additional complexity could also be added to the model to ensure that energy from these resources can be stored, however this would introduce a dependence on time which introduces another consideration. A consideration for climate change over long time horizons could also be included.
- An extension of the model for the use in "smart grids" could be undertaken. This would allow the system operator to influence the use of energy in the system, creating a more reliable and cost efficient system.
- It could be argued that modeling transmission lines by a single link is unrealistic as in real life systems a transmission line is composed of several lines, to allow for maintenance work to be undertaken. A better understanding of the nature of the transmission system might enable incorporation of this without drastically increasing the complexity.
- The model could be extended to cope for rolling blackouts. These are intentional power outages in certain regions that avoid a complete system blackout, and are necessary to ensure that the load demand of high priority regions can still be serviced.
- Finally, an extension of the model for use in a power system consisting of multiple companies in competition which is connected to an energy spot market could be explored. This would require methods from the field of game theory and the objective would then be to maximise profit as oppose to minimising cost.

Bibliography

[1] Abraham, J. Investment prospects for geothermal power in El Salvador's electricity market. *Energy Policy 34* (2006), 3877–3886.

[2] Ahmed, S., King, A. J., and Parija, G. A multi-stage stochastic integer programming approach for capacity expansion under uncertainty. *Journal of Global Optimization 26* (2004), 2–24.

[3] Azom.com. Photovoltaic thin film manufacturing cost reduction. http://www.azom.com/details.asp?ArticleID=1167, January 2002. [online; accessed June-2010].

[4] Beenstock, M. Generators and the cost of electricity outages. *Energy Economics 13*, 4 (October 1991), 283–289.

[5] Ben-Tal, A., Goryashko, A., Guslitzer, E., and Nemirovski, A. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming 99* (2004), 351–376.

[6] Borenstein, S. Carbon-emissions culprit? coal. http://seattletimes.nwsource.com/html/nationworld/2003732690_carbon03.html, June 2007. [online; accessed June-2010].

[7] Boyd, S., and Vandenberghe, L. *Convex Optimization.* Cambridge University Press, 2004.

[8] Calafiore, G. Multi-period portfolio optimization with linear control policies. *Automatica 44 10* (2008), 2463–2473.

[9] Carbon Monitoring for Action (CARMA). International database of carbon emissions. http://carma.org/. [online; accessed June-2010].

[10] Chaton, C., and Doucet, J. A. Uncertainty and investment in electricity generation with an application to the case of hydro-québec. *Annals of Operations Research 120* (2004), 59–80.

[11] Chen, X., Sim, M., Sun, P., and Zhang, J. A linear-decision based approximation to stochastic programming. *Operations Research 56 2* (2008), 344–357.

[12] Chicago Climate Exchange (CCX). CCX Carbon Financial Instrument (CFI) contracts daily report. http://www.chicagoclimatex.com/market/data/summary.jsf. [online; accessed June-2010].

[13] Comisión Económica Para América Latina y el Caribe (CEPAL). Istmo centroamericano: Estadísticas del subsector eléctrico. Tech. rep., United Nations, October 2007. http://www.eclac.org/publicaciones/xml/1/30101/L809.pdf. [online; accessed June-2010].

[14] Dantzig, G. B. Linear programming under uncertainty. *Management Science 50*, 12 Supplement (2004), 1764–1769.

[15] Doege, J., Schiltknecht, P., and Lüthi, H. J. Risk management of power portfolios and valuation flexibility. *Operations Research Spectrum 28 2* (2006), 267–287.

[16] Dyer, M., and Stougie, L. Computational complexity of stochastic programming problems. *Mathematical Programming 106 3* (2006), 423–432.

[17] Edwards, W. Dynamic decision theory and probabilistic information processing. *Human Factors 4* (1962), 59–73.

[18] Energy Information Administration (EIA). Estimated levelized cost of new generation resources, 2016. Tech. rep., December 2009. http://www.eia.doe.gov/oiaf/aeo/electricity_generation.html. [online; accessed June-2010].

[19] Energy Information Administration (EIA). Annual energy outlook 2010. Tech. rep., May 2010. http://www.eia.doe.gov/oiaf/aeo/. [online; accessed June-2010].

[20] Fleten, S. E., Wallace, S. W., and Ziemba, W. T. Hedging electricity portfolios via stochastic programming. *The IMA volumes in mathematics and its applications 128* (1999).

[21] Georghiou, A., Wiesemann, W., and Kuhn, D. Generalised decision rule approximations for stochastic programming via liftings (working paper).

[22] Global Oil & Gas Study. Electrical generation efficiency (working document). Tech. rep., National Petrolium Council (NPC), July 2007. www.npc.org/Study_Topic_Papers/4-DTG-ElectricEfficiency.pdf. [online; accessed June-2010].

[23] Gobierno De El Salvador. Política energética. Tech. rep., May 2007. ewh.ieee.org/r9/el_salvador/biblioteca/Politica%20energetica.pdf. [online; accessed June-2010].

[24] Goh, J., and Sim, M. Distributionally robust optimization and its tractable approximations. *Operations Research* (2009).

[25] Gröwe-Kuska, N., and Römisch, W. Stochastic unit commitment in hydrothermal power production planning. In *Applications of Stochastic Programming*, S. W. Wallace and W. T. Ziemba, Eds. SIAM, 2005, pp. 663–653.

[26] International Monetary Fund (IMF). Historical database of commodity prices. http://www.imf.org/external/data.htm. [online; accessed June-2010].

[27] Jogannathan, R. Minimax procedure for a class of linear programs under uncertainty. *Operations Research 25* (1977), 173–177.

[28] Kall, P., and Wallace, S. W. *Stochastic Programming*. John Wiley and Sons, Chichester, 1994.

[29] Kuhn, D., Parpas, P., and Rustem, B. Stochastic optimization of investment planning in the electric power industry. In *Process Systems Engineering: Volume 5: Energy Systems Engineering*, E. N. Pistikopoulos, M. C. Georgiadis, and V. Dua, Eds. Wiley, 2007, pp. 215–227.

[30] Kuhn, D., Wiesemann, W., and Georghiou, A. Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming* (2009).

[31] LaCommare, K. H., and Etoa, J. H. Cost of power interruptions to electricity consumers in the united states. *Energy 12 31* (2006), 1845–1855.

[32] Leopoldino, C. M. A., Pereira, M. V. F., Pinto, L. M. V., and Ribeiro, C. C. A constraint generation scheme to probabilistic linear programs with an application to power system expansion planning. *Annals of Operations Research 50* (2005), 367–385.

[33] Luenberger, D. G. *Investment Science*. Oxford University Press, 1998.

[34] National Renewable Energy Laboratory (NREL). Wind resource map for el salvador - 50m wind power. `www.nrel.gov/wind/pdfs/el_salvador.pdf`. [online; accessed June-2010].

[35] National Renewable Energy Laboratory (NREL). Solar: annual average latitude tilt map at 40km resolution for central america from nrel. `http://swera.unep.net/index.php?id=35&idx=152`, December 2003. [online; accessed June-2010].

[36] Natural Gas. Common properties of commercial fuels. `http://www.natural-gas.com.au/about/references.html`. [online; accessed June-2010].

[37] Nogales, F. J., Prieto, F. J., and Conejo, A. J. A decomposition methodology applied to the multi-area optimal power flow problem. *Annals of Operations Research 120* (2004), 99–116.

[38] Nürnberg, R., and Römisch, W. A two-stage planning model for power scheduling in a hydro-thermal system under uncertainty. *Optimization and Engineering 3* (2004), 355–378.

[39] Perez, R. Solar: hourly solar (direct normal (DNI), global horizontal (GHI), and diffuse) data for select El Salvador sites. Tech. rep., State Univeristy of New York (SUNY), June 2004. `http://swera.unep.net/index.php?id=35&idx=195`. [online; accessed June-2010].

[40] Pritchard, G., Philpott, A. B., and Neame, P. J. Hydroelectric reservoir optimization in a pool market. *Mathematical Programming 103 3* (2005), 445–461.

[41] Shapiro, A., and Nemirovski, A. On complexity of stochastic programming problems. In *Continuous Optimization: Current Trends and Applications*, V. Jeyakumar and A. Rubinov, Eds., vol. 99. Springer, 2005, pp. 111–146.

[42] Superintendencia General De Electricidad y Telecomunicaciones (SIGET). Boletín de estadísticas elétricas 2008. Tech. rep., May 2009. `www.siget.gob.sv`. [online; accessed June-2010].

[43] Unidad de Transacciones (UT). Historical database on hourly prices and electricity dispatch. `http://216.184.107.60:8080/web/guest/home`, 2004. [online; accessed June-2010].

[44] Vehvilainen, I., and Keppo, J. Managing electricity market price risk. *Euoropean Journal of Operational Research 16 1* (2003), 136–147.

[45] Windustry. How much do wind turbines cost? `http://www.windustry.org/how-much-do-wind-turbines-cost`. [online; accessed June-2010].

Appendix A

Polyhedral Support

In both Theorems 2.1 and 2.2 proposed by Kuhn *et al* certain assumptions are made about the support of the probability measure $\mathbb{P}$. Here they are restated to keep this project self contained. The assumptions are that the support is a nonempty compact polyhedron of form

$$\Xi = \left\{\xi \in \mathbb{R}^k : W\xi \geq h\right\} \tag{A.1}$$

where $W \in \mathbb{R}^{l \times k}$ and $h \in \mathbb{R}^l$. It is further assumed, without loss of generality, that

$$W = (e_1, -e_1, \hat{W}^\top)^\top \text{ and } h = (1, -1, 0...0)^\top$$

with $\hat{W} \in \mathbb{R}^{l-2 \times k}$, and $e_1 \in \mathbb{R}^k$ is the defined as a standard unit basis vector. This restriction will enforce every element ξ to have a first element equal to 1, allowing us to represent affine functions in a compact form. It is also required for Ξ to span the whole sample space $\mathbb{R}^k$, which holds if the system $\hat{W}\xi \geq 0$ is strictly feasible.

Appendix B

Stochastic Processes

Ornstein-Uhlenbeck Solution

Proposition B.1: *The solution to an Ornstein-Uhlenbeck process*

$$dS_t = k(\mu - S_t)dt + \sigma dW_t \tag{B.1}$$

where $k > 0, \mu$ and $\sigma > 0$ are the parameters and W_t is a Wiener process is

$$S_t \sim S_0 e^{-kt} + \mu(1 - e^{-kt}) + \sigma\sqrt{\frac{1 - e^{-2kt}}{2k}} N_{0,1}$$

and hence S_t is normally distributed.

Proof. Defining $F(s,t) = S_t \mathrm{e}^{kt}$ the components of the Ito's lemma are calculated as

$$\frac{dF}{dS} = \mathrm{e}^{kt} \qquad \frac{dF}{dt} = kS_t\mathrm{e}^{kt} \qquad \frac{d^2F}{dS^2} = 0$$
$$a = k(\mu - S_t) \qquad b = \sigma$$

substituting these into Ito's Lemma 2.7 we have

$$dy_t = \left(k\mathrm{e}^{kt}(\mu - S_t) + kS_t\mathrm{e}^{kt}\right) dt + \sigma\mathrm{e}^{kt} dW_t$$
$$dy_t = \left(k\mu\mathrm{e}^{kt}\right) dt + \sigma\mathrm{e}^{kt} dW_t$$

which on solution of the differential equation gives

$$
\begin{aligned}
\mathrm{e}^{kt} S_t - S_0 &\sim \mu(e^{kt} - 1) + \sigma\sqrt{\frac{e^{2kt} - 1}{2k}} N_{0,1} \\
S_t &\sim S_0 e^{-kt} + \mu(1 - e^{-kt}) + \sigma\sqrt{\frac{1 - e^{-2kt}}{2k}} N_{0,1}
\end{aligned}
$$

and hence S_t is normally distributed as required. □

Schwartz Type 1 Solution

Proposition B.2: *The solution to the Schwartz Type 1 stochastic process*

$$
dS_t = k(\mu - \ln S_t) S_t dt + \sigma S_t dW_t
$$

where $k > 0, \mu$ and $\sigma > 0$ are the parameters and W_t is a Wiener process is

$$
\ln S_t \sim e^{-kt} \ln S_0 + (1 - e^{-kt}) \left(\mu - \frac{\sigma^2}{2k} \right) + \sigma\sqrt{\frac{1 - e^{-2kt}}{2k}} N_{0,1}
$$

Proof. Substituting $X_t = \ln S_t$ in Ito's Lemma gives

$$
\begin{aligned}
dX_t &= \left(k(\mu - X_t) - \frac{\sigma^2}{2} \right) dt + \sigma dW_t \\
dX_t &= k \left(\alpha - X_t \right) dt + \sigma dW_t
\end{aligned}
$$

with $\alpha = \mu - \sigma^2/2k$. This as an Ornstein-Uhlenbeck process of the form B.1 and hence, from the previous proposition, has solution

$$
\begin{aligned}
X_t &\sim X_0 \mathrm{e}^{-kt} + \alpha(1 - \mathrm{e}^{-kt}) + \sigma\sqrt{\frac{1 - \mathrm{e}^{-2kt}}{2k}} N_{0,1} \\
\ln S_t &\sim \ln S_0 \mathrm{e}^{-kt} + \alpha(1 - \mathrm{e}^{-kt}) + \sigma\sqrt{\frac{1 - \mathrm{e}^{-2kt}}{2k}} N_{0,1}
\end{aligned}
$$

□

Appendix C

Linear Regression

In this Appendix a it is shown how to perform a linear regression on historical futures prices in order to estimate the parameters k, μ and σ for the Schwartz type 1. Firstly, note that obviously S_0 is taken to be the current price of the futures. The other parameters can then be found by considering the relationship between consecutive observations to be linear with respect to a standard normal term ϵ as

$$\ln S_{i+1} = a \ln S_i + b + \epsilon$$

where a,b and ϵ depend on k, σ and μ through

$$\begin{aligned} a &= e^{-k\delta t} \\ b &= (1 - e^{-k\delta t}) \left(\mu - \frac{\sigma^2}{2k} \right) \\ \text{s.d.}(\epsilon) &= \sigma \sqrt{\frac{1 - e^{-2k\delta t}}{2k}} \end{aligned}$$

where s.d. represents the standard deviation. Upon rewriting these equations obtain our model parameters, corresponding to a movement of one unit of time δt

$$\begin{aligned} k &= -\frac{\ln a}{\delta t} \\ \sigma &= \text{s.d.}(\epsilon) \sqrt{\frac{2k}{1 - e^{-2k\delta t}}} \\ \mu &= \frac{b}{1 - e^{-k\delta t}} + \frac{\sigma^2}{2k} \end{aligned}$$

The parameters a, b and ϵ can be obtained by linear regression for $n+1$ realisations of the price S_i, $i \in 0, ..., n$ by calculating

$$S_x = \sum_{i=1}^{n} S_{i-1} \qquad S_y = \sum_{i=1}^{n} S_i$$

$$S_{xx} = \sum_{i=1}^{n} S_{i-1}^2 \qquad S_{xy} = \sum_{i=1}^{n} S_{i-1} S_i \qquad S_{yy} = \sum_{i=1}^{n} S_i^2$$

the parameters of the least squares fit are then given by

$$a = \frac{nS_{xy} - S_x S_y}{nS_{xx} - S_x^2}$$

$$b = \frac{S_y - aS_x}{n}$$

$$\text{s.d.}(\epsilon) = \sqrt{\frac{nS_{yy} - S_y^2 - a(nS_{xy} - S_x S_y)}{n(n-2)}}$$

It is important that these parameters are annualized as the time points $t \in [0, T]$ correspond to years. So if, for example, prices are given per month multiply k by 12 and σ by $\sqrt{12}$.

Appendix D

A Map of El Salvador Irradiation

Special thanks to NREL and UNEP SWERA for allowing the publication of Figures D.1 and E.1.

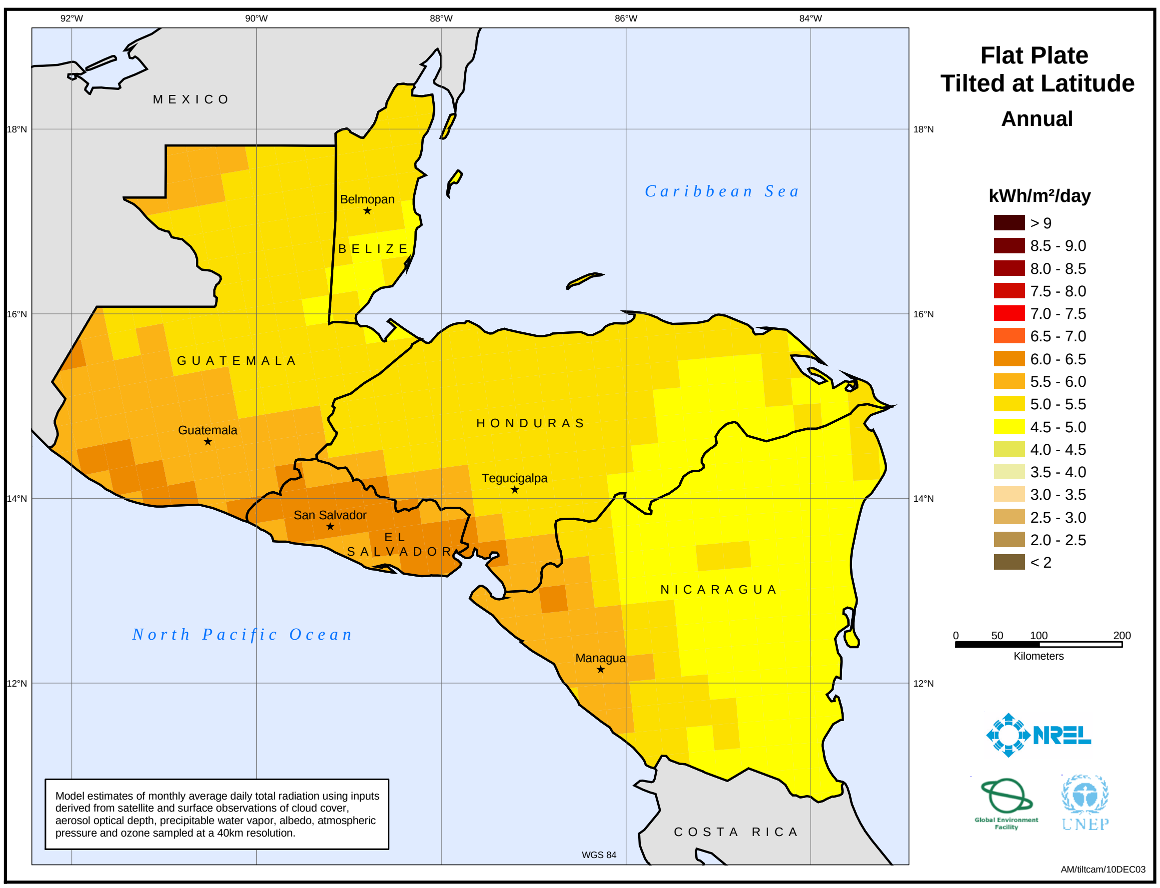


Figure D.1: Annual average irradiance of Central America [35]

Appendix E

A Map of El Salvador Wind Speeds

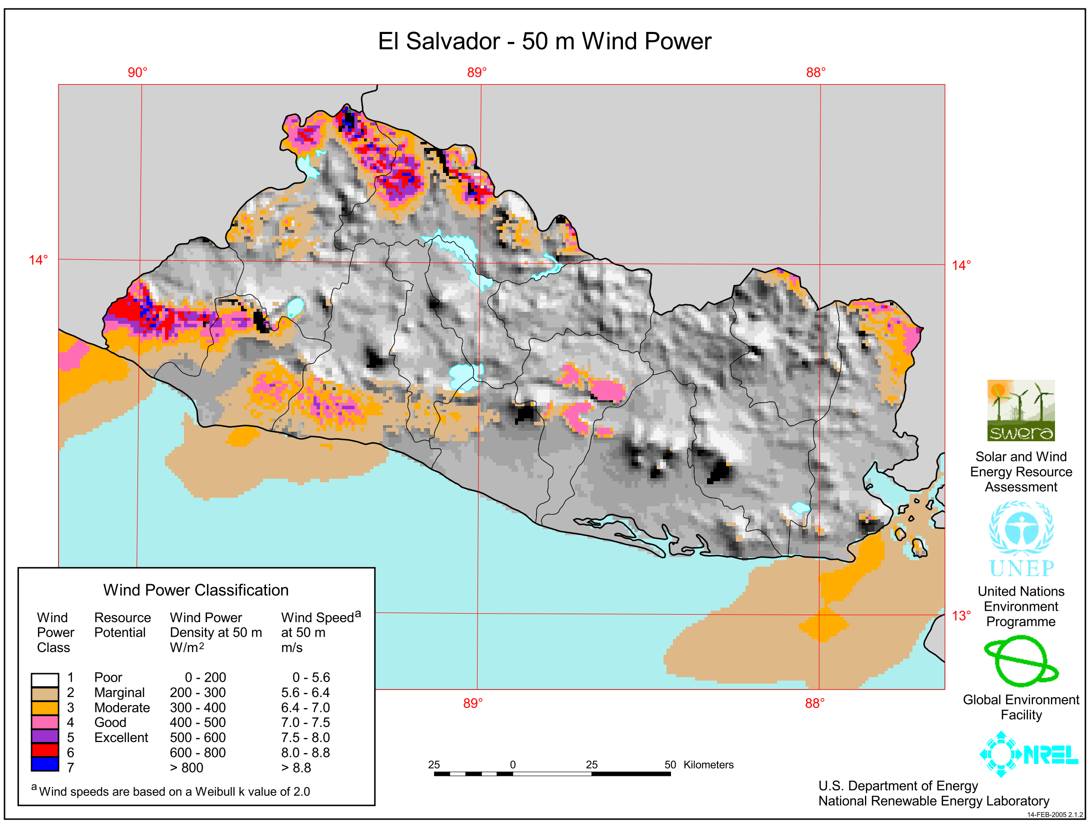


Figure E.1: Annual average wind speed for El Salvador [34]