A fast and flexible stochastic dynamic programming model of the electricity market

VTT-EMM – structure and use

Eero Tamminen | Veikko Kekkonen | Göran Koreneff | Tiina Koljonen

VTT

A fast and flexible stochastic dynamic programming model of the electricity market

VTT-EMM – structure and use

Eero Tamminen, Veikko Kekkonen, Göran Koreneff & Tiina Koljonen
VTT

Cover image: Göran Koreneff

Preface

This report presents the Nordic electricity market model VTT-EMM, developed at VTT Technical Research Centre of Finland. All aspects of the model are discussed: the objectives of the development work, the mathematical approach, the overall structure, the detailed representation of the different sectors of the production system, approximations applied, generation of price forecasts and other results, the computer system, and applications of the model.

The original development work was done in 2000–2004, and the model has since been used in a large number of studies and analyses. The development work has previously only been documented in a few research reports and memoranda. This is the first comprehensive presentation of the model.

The original development of the model was financed by VTT Technical Research Centre of Finland, Tekes – the Finnish Funding Agency for Innovation, SVK-Pooli (the Development Pool for Electric Power Technology of the Finnish Energy Industries), Helsingin Energia and Turku Energia.

Contents

1. Introduction

1.1 Objectives of the development of the VTT electricity market model

This report presents the aggregated Electricity Market Model of VTT Technical Research Centre of Finland, VTT-EMM. The model is a mathematical representation of the dynamics of the demand, production and market price of electricity in the power system of the Nordic countries, co-ordinated by the Nord Pool electricity market. The objective of the work was to develop a robust, agile, transparent and fast mathematical model of the stochastic development of the spot price of electricity on the Nord Pool market in the medium to long term. The model was intended for electricity price estimation, market analyses and energy policy studies. In policy studies, the model is required to be able to analyse the sensitivity of the price level to different policy measures in a time horizon that extends from the present time to 20–30 years into the future. Short-term (days to weeks) questions are outside the scope of this model.

1.2 The Nordic power system and market

The Nordic electricity system consists of Finland, Sweden, Norway, and Denmark, which have a common market place. The system is versatile and dominated by hydropower. The large share of hydropower is a most important property of the Nordic power system. Hydropower accounts for more than half of the total production of electricity in the system. The capacity of water reservoirs is large, about 60% of the yearly average hydropower production. The optimal use of hydro reserves over time is a central dynamic decision problem in the operative planning of the power producers.

Most of the electricity produced and consumed goes through the Nord Pool Spot exchange. The Nord Pool Spot is a day-ahead market on which the system spot price for each hour is determined by the intersection of all the demand and supply bids (Nord Pool 2013). There are about 360 buyers and sellers in the market in 2013, forming a sufficient basis for a well-working open electricity market. Internal bottlenecks can split the Nordic market into pre-defined market areas with individual prices, but this study focuses on the system spot price, which is always

formed and is the basis of all power market derivatives. Cross-border transmission lines within the market are implicitly in Nord Pool market use. Cross-border connections to Russia, Poland, Germany, Estonia and the Netherlands allow for imports and exports to and from the Nordic market. Nord Pool Spot has also expanded in recent years to the Baltic market, Estonia (2010) and Lithuania (2012).

1.3   Related research

The theory and praxis of mathematical optimisation underwent rapid development in the decades following the Second World War. New results were obtained, new approaches to optimisation problems were developed, and the emerging computers made it possible to apply the new methods in practice in many industrial sectors.

In the electric power sector, production optimisation models based on linear, non-linear and non-convex programming, control theory and the Pontryagin maximum principle, and dynamic programming (DP, the Bellman principle of optimality) were developed in co-operation between power producers, universities and research institutes. Valuable results were obtained, and the methods developed were implemented by the producers.

However, some of the most interesting problems proved too extensive for a direct integrated solution at that time. This was especially the case for the problem of optimal dynamic operation of large hydropower systems over medium-long (typically months to a year) time periods, when the essential stochastic nature of the inflow of water into the system, and of the demand, was taken into account. The problem can be formulated as an optimisation problem for a stochastic decision process and solved, in principle, by applying stochastic dynamic programming (SDP). In practice, the solution time was much too long for these dynamic and stochastic problems for large production systems. This was, and to a certain extent still is, the curse of dimensionality in the context of DP.

As the capacity of computers grew, more extensive applications became feasible. The experiences and results were mainly documented in laboratory reports and working papers. For surveys of this work, see Yakowitz (1982), Yeh (1985) and Stedinger (1998). After the creation of electricity markets, the same approaches and models were applied in price forecasting, and, in this context, they are called fundamental models.

The time-consuming task in DP is the construction of the value functions. A straightforward solution via the principle of optimality constructs the value functions state by state and time step by time step. New algorithmic approaches to the problem were developed; see Archibald et al. (1999). The stochastic dual dynamic programming (SDDP) algorithm of Pereira and Pinto (1991) has proven successful and has been applied to a variety of SDP problems. For applications in the Nordic power system, see Gjelsvik et al. (2010).

In Scandinavia, perhaps the most active institution in the power planning area has been the Norwegian Electric Power Research Institute, EFI. EFI developed methods and models systematically for short-, medium- and long-term optimisa-

tion of large power systems with water reservoirs. The EFI power planning model Samkjöringsmodellen is widely used. After a reorganisation in the Norwegian research field, these models and methods were further developed and made available by the Norwegian research organisation SINTEF. Today, the SINTEF EMPS model computes the Nord Pool system price, applying stochastic dynamic programming and a one-reservoir hydro model. It further considers several reservoirs and areas connected by transmission lines with limited capacity and uses a heuristic iterative procedure in order to obtain an overall power balance and market area prices. For information about the EMPS model, see the reference SINTEF 2014.

Compared with the earlier work, VTT-EMM does not contain any new mathematical methods or new applications on power market and production problems. VTT-EMM is a straightforward application of Bellman's principle of optimality on a stochastic decision process. The goal for the development work was to make this powerful approach easily accessible to the user by constructing a simple, transparent and fast model for Nord Pool systems price forecasting. To achieve this goal the representation of hydropower and thermal power is as simple as possible, approximations are made in the recurrent optimisation subtask, and fast programs and efficient task organisation are applied in the execution of the optimisation subtasks.

A short computing time is not a goal in itself. Fast execution of the dynamic programming computations makes it possible to extend the model with an additional state variable or stochastic element, and allows extensive scenario analyses. A simple model is also transparent, and this is a great benefit for the user. For every model, the numerical output is only part of the results. The model should also help the analyst to form a clear understanding of the structure and the logic of the problem, i.e. to help to understand clearly and quantitatively how the results depend on the elements and data of the problem. This latter result, if achieved, may even be more important than the numerical results as such.

The recent interest in hedging against electricity price and volume risks has motivated industry and researchers to develop viable models on a broad scale describing electricity price behaviour. The reported approaches vary with respect to application and in terms of the time horizon of price forecasting. Most of the existing literature focuses on developing realistic spot price models based on mathematical finance. However, many statistical models have been created to characterise not only spot price dynamics but also electricity derivatives. During this century, stochastic models for spot prices and power derivatives have been reviewed and further developed by, for example, Skantze et al. (2000a), Lucia and Schwartz (2002), Bunn and Karakatsani (2003), Fleten and Lemming (2003), Weron (2005), Deng and Wenjiang (2005), Vehviläinen and Pyykkönen (2005), and Benth and Koekebakker (2008). These models usually require estimations of many parameters, which may be difficult even with long time series available. The available data on the Nordic electricity market are insufficient due to the strong seasonal and yearly variations and because of several structural and political changes that have influenced the price dynamics since the market opening in 1996. As statistical

models are not able to describe longer term dynamics of electricity prices, these models should be used for short-term intervals. Medium- and long-term expectations of spot prices are often formed by fundamental models, as these models usually include a detailed technical description of generation, transmission and distribution, as well as extensive data sets of hydrological conditions, fuel prices, outdoor temperatures, etc. The main drawback of the fundamental models is that they do not capture the price of risk determined by the market forces. More information on the fundamental models can be found in Fleten and Wallace (1998), Skantze et al. (2000a, 2000b), and Tamminen and Kekkonen (2001a, 2001b). As for the functioning of the Nordic market, Amundsen and Bergman (2006) present an in-depth analysis of why it has worked so well.

1.4 Overview of the model and the solution method

The operation of the Nordic power production system is modelled as a stochastic decision process, controlled by active production decisions made on the electricity market and by exogenous stochastic processes, such as the inflow of water into the hydro system and the demand. It is assumed that the bidding and decision making process on the Nord Pool market leads to a cost-optimal allocation of production to the total demand, where the marginal production cost at optimum is the estimate for the spot price of electricity. The optimisation problem is solved with the model, which, for any given initial state of the system and for any realisation of the input random process, computes the corresponding realisation of the modelled price to serve as an estimate for the real market price in the actual case. The structure of the model is completely general, and the model is easily adapted and extended to represent also other production systems and markets, when the technical and economic data required are available.

The problem set up and solved with the model was the minimisation of the total expected production costs of the Nordic power system over the study period. The length of a study period (a model run) is from one to a few years, divided into weeks in the model. A week is the time step of the discrete model of the stochastic decision process. The stochastic input processes are treated as weekly sequences, and the production allocation decisions proceed in weekly steps.

The principle of optimality of Bellman is applied, and the problem is solved through stochastic dynamic programming. The solution is in the form of an optimal production policy and the corresponding weekly dynamic programming (DP) value functions for the whole period. The value functions essentially define the value of water in the hydro reservoir of the model in future electricity production.

The decisions forming the optimal policy, and the value functions, are solved recursively week by week for the whole study period, starting from the end of the period. For each week the model sets up an optimisation problem representing the balance between demand and production in the week. In every possible situation the corresponding optimal weekly decision allocates the demand to the capacity so that total expected production costs from the beginning of the actual week to

the end of the study period are minimised. The expected value of the hydro reserves left over for future production, solved earlier during the recursion, is included in the objective function. In order to obtain the overall solution, the weekly optimisation task is solved over and over again with different parameters. The applicability of the model is essentially determined by the efficiency of solving this task.

The weeks are further divided into hours in the model, but the weekly production allocation problems do not have any genuine internal dynamic structure, and the division into hours is only done in order to model the variation in demand, production, production costs and price within the week. Other subdivisions of time into load segments have also been applied in the weekly model.

The model is based on a highly aggregated representation of the power system. The development of the total yearly demand in the system is defined as a discrete stochastic process. The demand for each week and hour is computed from the total demand with the help of statistical index series. All hydropower production in the Nordic area is aggregated into one run-of-river hydro plant and one power plant with a water reservoir. The total inflow of water into the system is defined explicitly as a weekly stochastic sequence. The running production costs of hydropower are taken to be zero. All other production, i.e. thermal and wind power, is also aggregated into a single plant called the thermal plant. The thermal plant is represented by a total cost function defined as the expected minimum production costs per time unit with the available capacity of the system as a function of power produced. The available capacity is always dispatched optimally in the order of rising production costs. The expected costs are computed over the probability distributions of the available capacities of different plants and production classes of the system. The cost functions are computed for each time interval by efficient auxiliary programs on the basis of plant or production class data for the actual time.

The approach leads to a recurrent weekly optimisation problem with linear constraints and a convex objective function. In the model, the objective function is approximated with a piecewise linear function, leading to a linear programming (LP) problem. The efficient programs available for LP problems can then be used for the solution of the recurring problem. The closeness of the approximation to the original convex function can be chosen freely.

2. Mathematical structure of the model

2.1 Stochastic decision processes

2.1.1 State variables, stochastic variables and decision variables

We examine dynamic stochastic decision processes with the following structure: Time t is a discrete variable taking values (steps) $t = 0,1,..., T$. At every time step t there are three vector variables associated with the process: the state variable, the stochastic variable and the decision variable. The state variable $x(t)$ expresses the state of the system at time t (at the end of step t). The stochastic variable $S(t)$ at t is a random vector variable with possible values (realisations) $s(t)$. The decision (control) variable $u(t)$ is the operative decision at time t (during step t). The states $x(t)$, the values $s(t)$ of the stochastic element $S(t)$ and the decision variables $u(t)$ are vectors in finite dimensional real vector spaces.

The sequence of states $x = \{x(0), x(1),...,x(T)\}$ is called the trajectory of the system. The sequence of stochastic variables is a random process $S = \{S(1), S(2),..., S(T)\}$ with realisations $s = \{s(1), s(2),..., s(T)\}$ called the stochastic control. The sequence of decisions $u = \{u(1), u(2),..., u(T)\}$ is the active operational control of the process.

2.1.2 Evolution of the system and the dynamic equation

The evolution of the system over time is determined uniquely by the starting point (initial state) of the trajectory, the realisation of the stochastic control, the chosen operational control and the law of motion (the dynamic equation) of the system. The initial state at $t = 0$ is usually given by $x(0) = q$. The state $x(t-1)$ reached at the end of step $t-1$ is the starting point for the continuation of the trajectory at step t. At the beginning of step t, the random control element $S(t)$ is realised first; let the realisation be $s(t)$. Then, an active control element $u(t)$ is chosen. The process moves from state $x(t-1)$ to state $x(t)$ at the end of step t given by the dynamic equation (2.1).

$$x(t) = \Phi_t\big[x(t-1), s(t), u(t)\big] \quad for\ all\ t=1,..., T \tag{2.1}$$

In the dynamic equation, $\Phi_t[.,.,.]$ is a given function between the appropriate spaces. For brevity, the possible dependence of t is often not explicitly indicated. To emphasise the stochastic nature of the motion of the system, we also write the dynamic equation (2.1) as $X(t) = \Phi_t[X(t-1), S(t), u(t)]]$, $t=1,2,\ldots,T$ with the random variable $S(t)$, instead of the realisation $s(t)$, as one of the arguments. Consequently, state $X(t)$ is also a random variable that depends on the previous state $X(t-1)$ of the stochastic control element $S(t)$ and the active control $u(t)$. The equation of motion (2.1) gives the exact definition.

A triple $\{x,s,u\}$, consisting of a realisation s of the random process, a control u and the corresponding trajectory x according to the dynamic equation (2.1), is called a realisation of the decision process. Partial realisations of the process, i.e. realisations restricted to some subinterval of time, are also frequently considered.

The state trajectory, the stochastic control, the active control and the equation of motion form the basic structure of the stochastic decision process. Note again that the state trajectory X of the decision process is a stochastic process that depends on the process S and the control u. The definition of the process is completed by adding the probability distribution of the stochastic process S, constraints on the variables and an objective (cost) function. Minimisation of the objective function can then be applied as a criterion in making the operative decisions.

2.1.3   Information and decision structure of the process

The dynamic structure formed by the realisation of the stochastic control, unfolding step by step, and the active controlling decisions made between the successive stochastic elements is an essential characteristic of the process. An important aspect is the accumulation of information about the development of the process and the opportunity to utilise the accumulated information in forecasting future development. At the beginning of step t, the history of the whole decision process during the time interval $[1, t-1]$ is known. This includes, especially, the partial realisation $\{s(1), s(2),\ldots, s(t-1)\}$ of the random control process S and state $x(t-1)$. The probability distribution of the stochastic element $S(t)$, which corresponds to the 'information state', is its conditional distribution assuming the history of the process. Then, $S(t)$ is realised; let the realisation be $s(t)$. The state $x(t-1)$, reached at the end of step $t-1$, and the realisation $s(t)$ uniquely determine the set Ω_t of feasible (allowed) decision alternatives at step t; see Section 2.1.5. Then, a decision $u(t) \in \Omega_t$ is chosen, and the process moves to the state $x(t)$ given by the dynamic equation (2.1).

At any point t in the decision process, set Ω_t of allowed decision alternatives is dependent on the history of the decision process, including the realisation $s(t)$ of the stochastic control element $S(t)$. The objective applied in making the decision $u(t)$ is a minimisation of the expected costs during the remaining period $[t,T]$. This optimisation problem again depends on both the feasible set Ω_t, and the conditional distribution of the stochastic process S during the period $[t+1,T]$, assuming its history up to and including $s(t)$.

2.1.4 Probability distribution of the stochastic control

The model is used for studying the behaviour of the process through simulation. As the realisations of the decision process are computed step by step, the probability distribution of the stochastic element $S(t)$ is needed at each step t. It is assumed that all these distributions are available. The mathematical structure of the decision process does not set any special requirements for the probability distribution of the control process S. If the process is completely defined mathematically, then the distributions are also given. If S is a model of some real-world process, these distributions have to be estimated statistically.

At step t the partial realisation $\{s(1), s(2),...,s(t-1)\}$ of the process S during the time interval $[1, t-1]$ is known. At that point of the process, the proper distribution of $S(t)$, which corresponds to the accumulated information, to be used in the computations is its conditional distribution, assuming the realisation $\{s(1), s(2),...,s(t-1)\}$. However, it may be difficult to obtain the distributions. For example, the stochastic control S can be a mathematical model of some economical or weather process. For such phenomena, it is often the case that even if there is a correlation between the successive elements $S(t)$, there is not enough theoretical knowledge or statistical data available to allow the conditional distributions of the elements $S(t)$ to be constructed and estimated reliably.

Furthermore, in the dynamic programming iteration, all the information about the history of the process used in the continuation of the process in step t has to be contained in the state $x(t-1)$. In other words, information about the correlation between the stochastic elements $S(\tau)$ and $S(t)$ for $\tau < t$ has to be carried by the state $x(t-1)$. Logically, this requirement could easily be satisfied by including the whole history in the state. This again is impossible in practice due to the well-known curse of dimensionality. For computational reasons, the state cannot have too many dimensions.

For computational reasons, the probability distributions of the stochastic control elements $S(t)$ are also treated as discrete approximations in the applications. Note that this is a purely numerical and computational approximation. We assume that for every $t = 1,2,...,T,$ the random vector $S(t)$ has a discrete distribution:

$$S(t) = s(t,i) \; with \; the \; probability \; \pi(t,i), \; where \; i \in I(t) = \left\{1,...,I_t\right\} \quad (2.2)$$

and for all $t,$ the possible realisations $s(t,i)$ and their probabilities $\pi(t,i)$ are available when needed in the solution process. Furthermore,

$$\sum_{i \in I(t)} \pi(t,i) = 1 \qquad\qquad\qquad\qquad (2.3)$$

As discussed above, if the correlation between successive stochastic control elements is modelled, then the probabilities $\pi(t,i)$ depend on the state $x(t-1)$. As we cannot usually model this correlation and, in order to keep the notation simple, this dependence is not explicitly indicated.

2.1.5　Constraints

At each time step *t*, the decision variables *u(t)* are required to satisfy constraints
(2.4), where *s(t)* again is the realisation of the stochastic element *S(t)*. The con-
straint set Ω_t in (2.4) is usually defined through equalities and inequalities. After
the realisation of *S(t)*, the constraint set Ω_t is uniquely determined.

$$u(t) \in \Omega_t\big[x(t-1), s(t)\big], \quad for \ \ all \ \ t = 1,\dots,T \tag{2.4}$$

In addition to the control constraints (2.4) that depend on the previous state *x(t-1)*,
there may be control constraints (2.5) that depend on the next state *x(t)* and con-
straints (2.6) on the next state *x(t)* in the formulation of the process:

$$u(t) \in \Xi_t\big[x(t), s(t)\big], \quad for \ \ all \ \ t = 1,\dots,T \tag{2.5}$$

and

$$x(t) \in \Psi_t, \quad for \ \ all \ \ t = 1,\dots,T \tag{2.6}$$

where Ξ_t *[x(t), s(t)]* and Ψ_t are given constraint sets. The constraints of type (2.5)
and (2.6) are rewritten as control constraints of type (2.4) that depend on the pre-
vious state *x(t-1)* and the realisation *s(t)*, by solving the state *x(t)* from the dynamic
equation (2.1). Through this transformation, constraints of type (2.5) and (2.6) are
included in the set (2.4) of control constraints.

2.1.6　The cost function

For the time period *[1,t]*, for any trajectory *x*, any realisation s of the stochastic
control and any control *u,* the costs are given by the sum

$$\sum_{t=1}^{T} c_t\big[x(t), s(t), u(t)\big] \tag{2.7}$$

where for every *t*, c_t*[.,.,.]* is a given real function of its arguments. The expected
costs in the period of the trajectory *x* and the control *u* are given by

$$E_S \sum_{t=1}^{T} c_t\big[x(t), S(t), u(t)\big] \tag{2.8}$$

The objective applied in choosing between the different controls u*(t)* for all *t* is
minimisation of the expected value (2.8) of the objective function. Note that the
expected value is determined when a trajectory *x*, a control *u* and the distribution
of the stochastic control *S* are given.

2.1.7   The stochastic decision problem

We have now completed the mathematical definition of a stochastic decision process. The tools are applied to modelling the electricity market decision process. The following problem is a basic sequential optimisation problem for the stochastic decision process under study:

$$\min_u E_S\left\{\sum_{t=1}^{T}c_t\big[x(t),S(t),u(t)\big]\right\} \tag{2.9}$$

under the constraints:

$$x(t)=\Phi_t\big[x(t-1),S(t),u(t)\big],\ for\ all\ t=1,...,T \tag{2.10}$$

$$x(0)=q \tag{2.11}$$

$$u(t)\in\Omega_t\big[x(t-1),S(t)\big],\ for\ all\ t=1,...,T \tag{2.12}$$

The objective is to minimise the expected costs (2.9). To indicate the stochastic nature of the process, we have again written the random variable $S(t)$, instead of its realisation $s(t)$, as an argument in the dynamic equation (2.10) and in the constraint set (2.12). Somewhat inconsequently, we have used, for the state trajectory, the symbol for its realisation $x(t)$, but this corresponds exactly to the use of the formulae (2.9)-(2.12) in computations. The stochastic control elements $S(t)$ are treated explicitly as random variables, but for the state trajectory X individual realisations x are computed according to the law of motion (2.10) and the constraints (2.11)-(2.12).

2.1.8   Solution to the problem: an optimal decision policy

Next, we take a closer look at the dynamics of the stochastic decision process (2.9)-(2.12) and at the nature of the optimisation problem. The process starts from the given initial point (2.11). Then, the first random element $S(1)$ is realised as $s(1)$, which is inserted into the dynamic equation (2.10) and the constraint set (2.12). Then, we choose some $u(1)$ satisfying the constraints (2.12) and obtain the next state $x(1)$ from the dynamic equation (2.10). The first term of the cost function in (2.9) is now determined and can be computed. The process is repeated for every step $t = 2,...,T$.

In the decision problem, the objective is minimisation of the expected costs (2.9) over the whole time period, when every control element $u(t)$ is fixed in turn in the defined time order. The solution to this problem is not a control $u = \{u(1), u(2),..., u(T)\}$ fixed at $t=0$. The solution is an optimal decision policy, with the property that every decision $u(t)$ is made optimally at the actual situation reached,

defined by the current state $x(t-1)$ and the realisation of the stochastic control process up to and including $s(t)$. When choosing $u(t)$, the objective is to minimise the expected value of the costs over the remaining time period $[t,…,T]$. At this decision point, only the immediate contribution $c_t[x(t), s(t), u(t)]$ can be computed for every choice of $u(t)$ by solving $x(t)$ from the dynamic equation. For the costs during the remaining steps $t+1,…,T$, we use the expected value. The choice of $u(t)$ has to be made so as to minimise the sum of the immediate contribution from step t and the expected value for the remaining process. Information gained during the period $[0,…,t]$, i.e. the partial realisation $[s(1),…,s(t)]$, can be used in the estimation of the probability distribution of the continuation $[S(t+1),…,S(T)]$ of the process S.

2.2 Solving the stochastic decision problem by dynamic programming

2.2.1 Solving the optimal decision policy

We have formulated the stochastic decision problem associated with the model so that Bellman's principle of optimality and the dynamic programming method can be applied to its solution. For the problem (2.9)-(2.12), Bellman's principle can be stated as follows:

An optimal decision policy has the property that whatever the initial state and the initial decision, the remaining process (decision policy) is optimal in the remaining problem starting from the state resulting from the first decision.

We define the value functions G_t for all time steps t and for all feasible states y at the beginning of step t (at the end of step $t-1$), before the random vector $S(t)$ is realised and before the decision $u(t)$ is taken as follows:

$G_t(y) =$ Expected minimum costs from the beginning of step t to the end of the horizon T (to the end of step T), when the initial state at the beginning of step t (at the end of step $t-1$) is y.

Then, we start from state y at the beginning of step t. First the random vector $S(t)$ is realised, $S(t) = s(t,i)$ with probability $\pi (t,i)$. Then, whichever $s(t,i)$ was realised, $u(t)$ is chosen, in every case, so as to minimise the expected (optimal) costs from the beginning of step t to the end of the horizon T. If we choose $u(t)=w$, then at the end of the step, the state $x(t)=\Phi_t [y, s(t,i), w]$. The contribution to the cost function from step t is $c_t [\Phi_t [y, s(t,i), w], s(t,i), w]$, and the minimum expected contribution from the steps $(t+1)$ to T, according to the definition of the value function, is equal to $G_{t+1} [\Phi_t [y, s(t,i), w]]$. Taking an optimal decision w in each case (for each $s(t,i)$) and forming the expected value over the distribution of $S(t)$, we obtain the fundamental recursive relation for the stochastic decision problem (2.9)-2.12):

$$G_t(y) = \sum_{i \in I(t)} \pi(t,i) \min_w \left\{ c_t \big[\Phi_t [y, s(t,i), w], s(t,i), w \big] + G_{t+1} \big[\Phi_t [y, s(t,i), w] \big] \right\} \tag{2.13}$$

$$w \in \Omega_t [y, s(t,i)]. \tag{2.14}$$

By applying the recursion (2.13)-(2.14), the value function G_t can be computed, if the function, one time step later, G_{t+1} is known. The iteration can always be started backwards in time from:

$$G_T(y) = \sum_{i \in I(T)} \pi(T,i) \min_w \left\{ c_T \big[\Phi_T [y, s(T,i), w], s(T,i), w \big] \right\} \tag{2.15}$$

$$w \in \Omega_T [y, s(T,i)]. \tag{2.16}$$

Starting from (2.15)-(2.16), the iteration (2.13)-(2.14) is carried out for every t = T-1, T-2,…,1. In this iteration, the optimisation subtask (2.17)-(2.18) is a fundamental building block of the model.

$$\min_w \left\{ c_t \big[\Phi_t [y, s(t,i), w], s(t,i), w \big] + G_{t+1} \big[\Phi_t [y, s(t,i), w] \big] \right\} \tag{2.17}$$

$$w \in \Omega_t [y, s(t,i)] \tag{2.18}$$

The optimal solutions $\hat{w}_t[y, s(t,i)]$ to this problem for all steps t, all possible states $x(t-1) = y$ and all possible realisations $s(t,i)$ of $S(t)$ constitute the complete solution, the optimal policy for the problem. In the course of solving the overall problem by dynamic programming, the optimisation subtask will be solved over and over again with different parameters. The applicability of the model is essentially determined by the effectiveness of solving this basic problem.

2.2.2   Generating optimal realisations of the stochastic decision process

The value functions G_t for all t give a complete solution to the optimisation problem (2.9)-(2.12). By using these, optimal realisations of the decision process can be generated. We start from $x(0) = q$ and take a realisation $s = \{s(1), s(2),…, s(T)\}$ of the stochastic control. The first optimal decision $\hat{u}(1)$ is solved from problem (2.17)-(2.18) with the parameters $y=q$, $s(t,i)=s(1)$ and t=1. Then, $x(1)=\Phi_t [q, s(1), \hat{u}(1)]$, and we proceed to the next step and generate the whole optimal realisation of the decision process for the realisation s of the stochastic control. Repeating the computation for a representative sample of realisations of the process S, a corresponding sample of realisations of the overall process can be computed and estimates of various expected values etc. obtained.

2.3 Model of the electricity market and the production system

2.3.1 The overall model and the weekly optimisation subtask

The stochastic dynamic programming approach to modelling decision processes, presented in Sections 2.1–2.2, is now applied to the construction of a model of the Nordic electricity production system, co-ordinated by the Nord Pool electricity market. A large share (about 50%) of hydropower, stochastic inflow of water into the reservoirs, and economically optimal use of water are essential characteristics of the system to be modelled.

The model is used to determine operation policies for the production system that minimise the expected total production costs by using the whole available production capacity and the hydro reservoirs optimally. We assume that the electricity market realises this cost-optimal allocation of the demand for electricity to the available production resources and use the model to simulate the behaviour of the market.

The time steps t of the model are weeks, and for every week t the minimum expected production costs over the remaining study period (from the beginning of week t to the end of the study period) are computed with the model. The marginal cost, i.e. the derivative of the expected remaining total costs with respect to the demand for electricity during week t, gives the estimate of the spot market price of electricity in the week. If the market is ideal, it will realise this price.

For clarity, we first present the basic structure of the model in Chapter 3, in aggregated form, with all the essential properties of the detailed model used in applications. The aggregated form does not model the internal dynamics within the weeks, i.e. its variables are time averages over the week.

The detailed model, presented in Chapter 4 is technically and computationally more complicated. It has more decision variables and more detailed constraints. The weeks are divided into hours and the internal dynamic development is modelled approximately. The detailed model also makes use of many approximations in the numerical computations. The value function iteration (2.13)-(2.14) always proceeds in weekly steps however.

3. The overall model

3.1 Variables

At every week t, the model contains the following vector variables:

The state variable $x(t)$ has two components, $x(t) = [x_1(t),\ x_2(t)]$, where
   $x_1(t)=$ Amount of water in the aggregated hydro reservoir (in energy units) at the end of week t, and
   $x_2(t)=$ Relative level of total yearly demand for electricity reached at week t.

The stochastic element $S(t)$ also has two components $S(t) = [S_1(t),\ S_2(t)]$, where
   $S_1(t)=$ Inflow of water into the hydropower system during week t, and
   $S_2(t)=$ Additive change in the relative level of total yearly demand for electricity from week $t-1$ to week t.

The decision (control) variable $u(t)$ is a 3-vector $u(t) = [u_1(t),\ u_2(t)\ ,\ u_3(t)]$, where
   $u_1(t)=$ Production of hydropower (electricity) during week t,
   $u_2(t)=$ Production of thermal power during week t, and
   $u_3(t)=$ Release of water from the reservoir past the turbines during week t.

The components of the state $x(t)$ and the control $u(t)$ are real variables, and the components of the stochastic element $S(t)$ are real random variables with given distributions, which may depend on the state $x(t-1)$. $S_1(t)$ is the total inflow into the hydro system; it contains both the storable inflow into the aggregated hydro reservoir and the run-of-the river inflow, which has to be used in power generation as it flows into the rivers.

3.2 The dynamic equation

The difference equations (3.1)-(3.2) give the law of motion of the system:

$$x_1(t) = x_1(t-1) + S_1(t) - u_1(t) - u_3(t) \tag{3.1}$$

and

$$x_2(t) = x_2(t-1) + S_2(t) \tag{3.2}$$

for all t = 1, ..., T.

The equations are of the simplest possible type, linear with coefficients constant in time. The development of the demand *x2(t)* depends only on the second component $S_2(t)$ of the stochastic control. The equations (3.1)-(3.2) can be written as a single linear vector equation (3.3), corresponding the dynamic equation (2.10), *x(t)=Φ_t[x(t-1), S(t), u(t)], t= 1, 2,…,T*, for the system.

$$x(t) = x(t-1) + S(t) + Hu(t), \quad for \ \ all \ \ t = 1,...,T \tag{3.3}$$

where the constant matrix

$$H = \begin{vmatrix} -1 & 0 & -1 \\ 0 & 0 & 0 \end{vmatrix} \tag{3.4}$$

The initial state is given by:

$$x(0) = \left[x_1(0), x_2(0) \right] = q = \left[q_1, q_2 \right] \tag{3.5}$$

where q_1 is the initial level of the hydro reservoir and q_1=1, and the initial relative level of the total yearly demand is defined to be 1.

3.3 The probability distributions

The distribution of the stochastic control vector *S(t)* is defined as discussed in Section 2.1.4, *S(t) =s(t,i)*, with probability *π(t,i)*, where *i ∈ I(t)={1,…,I_t}*. The two components of the control process *S,* the scalar stochastic processes S_1 and S_2, are statistically independent. Their distributions may, in principle, depend on the state *x(t-1)*. In the basic formulation of the model, the inflow random variables *S_1(t)* for different weeks *t* are statistically independent, and their distributions depend only on the week *t.* The additions to the demand *S_2(t)* may be correlated, as the demand is the second component *x_2(.)* of the state, i.e. the distribution of *S_2(t)* may depend on the level of demand *x_2(t-1)* reached at the end of step *t-1.* For simplicity, in this presentation we assume that the distribution of each *S_2(t)* is a function of week *t* only.

3.4 Constraints

In the aggregated model, the variables satisfy the following constraints (3.6)-(3.9):

All state and control variables are nonnegative by definition (3.6). This is assumed and usually not repeated when discussing the constraints of the model.

$$0 \leq x_1(t), x_2(t), u_1(t), u_2(t), u_3(t), \; for \; all \; t = 1,...,T \tag{3.6}$$

The contents of the aggregated water reservoir satisfy the given time-dependent lower and upper bounds (3.7).

$$a(t) \leq x_1(t) \leq b(t), \; for \; all \; t = 1,...,T \tag{3.7}$$

The production of hydropower is limited by lower and upper bounds (3.8), which may depend on the state and the realisation of the inflow. The total weekly demand $d(t)$ for electricity has to be covered (3.9).

$$g(t) \leq u_1(t) \leq h(t), \; for \; all \; t = 1,...,T \tag{3.8}$$

$$d(t) \leq u_1(t) + u_2(t), \; for \; all \; t = 1,...,T \tag{3.9}$$

The weekly demand is given by:

$$d(t) = D_0 \eta_{1t} x_2(t), \; for \; all \; t = 1,...,T \tag{3.10}$$

where D_0 is the initial level of the yearly demand, and $\{\eta_{1t}\}$ is the weekly index series for the variation in the demand; see Chapter 9.

We insert (3.1) into (3.7), and (3.2) and (3.10) into (3.9) and obtain these constraints in the standard form corresponding to the definition (2.12):

$$x_1(t-1) + S_1(t) - b(t) \leq u_1(t) + u_3(t) \leq x_1(t-1) + S_1(t) - a(t), \; for \; all \; t = 1,..,T \tag{3.11}$$

$$D_0 \eta_{1t} \left[x_2(t-1) + S_2(t) \right] \leq u_1(t) + u_2(t), \; for \; all \; t = 1,..,T \tag{3.12}$$

3.5 The objective function and the market equilibrium (production optimisation) problem for the overall model

The objective function is the total operative production costs over the study period. Hydropower production costs are taken to be zero. For thermal power, we define the cost functions for each time step:

$c_t(f)$ = operative thermal production costs in week t as a function of to-
tal thermal production f during the week. (3.13)

It is assumed that the costs $c_t(f)$ are convex functions of production f, for all t. They are uniquely determined by the properties of the thermal capacity available for production during week t.

Now, we can state for the aggregated model the basic market equilibrium (production optimisation) problem in the form corresponding to the general sequential problem (2.9)-(2.12) for stochastic decision processes defined in Chapter 2.

The objective (2.9) has the form

$$\min_u E_S\left\{\sum_{t=1}^{T} c_t[u_2(t)]\right\}$$ (3.14)

The dynamic vector equation (2.10), $x(t)=\Phi_t[x(t-1), S(t), u(t)]$, $t=1, 2,...,T$, is given by the difference equations (3.15)-(3.16) for all $t = 1, ..., T$,

$$x_1(t) = x_1(t-1) + S_1(t) - u_1(t) - u_3(t)$$ (3.15)

$$x_2(t) = x_2(t-1) + S_2(t)$$ (3.16)

The initial conditions (2.11) at the time $t = 0$, $x(0) = q$ are

$$x_1(0) = q_1, \text{ and } x_2(0) = q_2 = 1$$ (3.17)

and the constraints (2.12), $u(t) \epsilon \Omega_t [x(t-1), S(t)]$, $t= 1,...,T$, are given by (3.18)-(3.21) for all $t = 1,..., T$.

$$g(t) \le u_1(t) \le h(t)$$ (3.18)

$$x_1(t-1) + S_1(t) - b(t) \le u_1(t) + u_3(t) \le x_1(t-1) + S_1(t) - a(t)$$ (3.19)

$$D_0\eta_{1t}[x_2(t-1) + S_2(t)] \le u_1(t) + u_2(t)$$ (3.20)

$$0 \le x_1(t), x_2(t), u_1(t), u_2(t), u_3(t)$$ (3.21)

The recurring minimisation sub-problem, (2.17)-(2.18) in the general formulation, now reads, for the aggregated model with parameters t, $x(t-1)=y=(y_1,y_2)$ and $S(t)=s(t,i)=(s_1(t,i), S_2(t,i))$, as follows:

$$\min_w \left\{ c_t(w_2) + G_{t+1} \left[\Phi_t \left[y, s(t,i), w \right] \right] \right\}, \text{ where} \tag{3.22}$$
$$\Phi_t \left[y, s(t,i), w \right] = \left[y_1 + s_1(t,i) - w_1 - w_3, y_2 + s_2(t,i) \right]$$

and the following constraints are satisfied:

$$g(t) \leq w_1 \leq h(t) \tag{3.23}$$

$$y_1 + s_1(t,i) - b(t) \leq w_1 + w_3 \leq y_1 + s_1(t,i) - a(t) \tag{3.24}$$

$$D_0 \eta_{1t} \left[y_2 + s_2(t,i) \right] \leq w_1 + w_2 \tag{3.25}$$

$$0 \leq w_1, w_2, w_3 \tag{3.26}$$

The sub-problem (3.22)-(3.26) has simple linear constraints and a convex objective function, which will be approximated with a piecewise linear function in numerical applications. The problem can be solved very fast. The optimal solutions $(\hat{w}_1, \hat{w}_2, \hat{w}_3) = (u_1(t), u_2(t), u_3(t))$ to the sub-problems (3.22)-(3.26) constitute the optimal policy for the overall problem (3.14)-(3.21).

3.6 Solution to the market equilibrium problem through dynamic programming

The solution to the basic optimisation problem (3.14)-(3.21) is now obtained by applying the dynamic programming recursion (2.13)-(2.14), which leads to the following procedure: For every $t = 1, \ldots, T$,

$$G_t(y) = \sum_{i \in I(t)} \pi(t,i) \min_w \left\{ c_t(w_2) + G_{t+1} \left[\Phi_t \left[y, s(t,i), w \right] \right] \right\} \tag{3.27}$$

$$w \in \Omega_t \left[y, s(t,i) \right] \tag{3.28}$$

The state transformation Φ in (3.27) and the feasible set Ω in (3.28) were defined in Section 3.5. The recurring minimisation sub-problem (3.22)-(3.26) in the short-hand notation reads:

$$Z^*_t \left[y, s(t,i) \right] = \min_w \left\{ c_t(w_2) + G_{t+1} \left[\Phi_t \left[y, s(t,i), w \right] \right] \right\} \tag{3.29}$$

$$w \in \Omega_t \left[y, s(t,i) \right] \tag{3.30}$$

As usual in the treatment of the dynamic programming recursion, we have used the letters y and w for the vectors $x(t)$ and $u(t)$ in their roles as general variables in

optimisation tasks. The optimal value in (3.29)-(3.30) at time t, for $x(t-1)=y$, and $S(t)= s(t,i)$ is denoted by $Z^*_t [y, s(t,i)]$.

If the objective is to minimise the production costs strictly within the period under study, then nothing that happens after the period has any weight in the decision problem. The recursion is started at $t=T$, and $G_{T+1}[\Phi_T [y, s(t,i), w]] \equiv 0$, in (3.27). This gives an optimal solution, which uses as much of the hydro reserves as possible. If the constraints allow, it is optimal to use them all. This does not usually correspond to the real decision situation. In order to avoid excessive use of hydro reserves within the period, the value of water in storage at the end T of the period has to be credited in the objective function.

Let the value of water stored at the end of week T be $V_{T+1} (y)$, where $y =(y_1, y_2)$, and y_1 is the amount of water stored and y_2 is the demand. Then, the recursion is started at $t=T$ from:

$$G_T(y) = \sum_{i \in I(T)} \pi(T,i) \min_w \left\{ c_T(w_2) - V_{T+1}\left[\Phi_T\left[y, s(T,i), w\right]\right] \right\} \quad (3.31)$$

$$w \in \Omega_T\left[y, s(T,i)\right] \quad\quad\quad\quad\quad\quad (3.32)$$

As soon as the function $V_{T+1} (y)$ is known, the recursion can actually be started from (3.31)-(3.32). In applications, the function V is usually unknown, and it has to be estimated before the recursive computations can begin.

3.7 Estimation of the water value function

The water value function V needed to start the dynamic programming iteration (3.31)-(3.32) was defined as $V_{T+1} (y)= V_{T+1} (y_1, y_2)$ = value of water stored in reservoirs as a function of the amount y_1, stored and of the level of demand y_2, both at the end of week T (at the beginning of week $T+1$). The value of water $V_{T+1}(y)$ thus depends only of what happens after time T, and it is not determined by data on the study period $[0,T]$. Conceptually, the water value function V is an additional, independent element in the problem, which has to be estimated and given by the planner.

The dynamic programming model can be used to generate water value functions however. When the model is used for this purpose, the basic assumptions of the analyses with the model are also made by the analyst.

The following approach can be applied: (1) extend the optimisation problem beyond time T to T', (2) collect data and generate forecasts for the stochastic elements for the extension $[T,T']$, (3) estimate the value of water at $T'+1$, and (4) solve the optimisation problem for the whole period $[0,T']$. The solution also gives the value of water at time $T+1$. If the extended period is long enough, and if the water value function at $T'+1$ is reasonable, then the value at $T+1$ is not very sensitive to the choice of the value at $T'+1$, and the former may be used in further studies where the value $V_{T+1} (y)$ is needed.

Another possibility is to assume that the planning situation remains stationary after T. Suppose $T = 1$ year. Then, we proceed as follows: (1) construct a stationary model for the stochastic demand where the demand remains at the level reached at T; (2) make an initial estimate for the water value function $V^0_{T+1}(y)$; (3) use the constructed stationary demand, the original water inflow model, and the initial estimate for the water value to solve the optimisation problem for the period; (4) on the basis of the results, compute the water value function $V^0_1(y)$; (5) repeat (3) using this value function as the value of water at $T+1$; and (6) iterate until the sequence of water value functions converges.

In practice, both quantitative analyses and practical experience are used in the choice of a suitable water value function. Finally, we note the relation (3.33) between the water value function V and the cost value function G for the problem, based directly on the definitions of these functions (y_1 is the amount of water and y_2 the level of demand):

$$V_t(y_1, y_2) = G_t(0, y_2) - G_t(y_1, y_2), \text{ for all } y_2 \text{ and } t \qquad (3.33)$$

4. The weekly optimisation subtask

4.1 Introduction

The weekly optimisation subtask (3.22)-(3.26) consists of allocating the weekly demand of electricity optimally to the available thermal and hydro capacity, taking into account the possibilities of storing water in hydro reservoirs. In general, it is a non-linear and possibly non-convex optimisation problem. In the present work, we have approximated the problem with a linear programming (LP) problem so that we can benefit from the efficiency with which LP problems can be solved. In this chapter, we present the weekly allocation problem in detail and its linear programming approximation applied in the model.

4.2 Time steps and indices

4.2.1 Weeks and hours

The time period under study is typically one or a few years. It is divided into weeks $t = 1,\ldots,T$ and further into hours $h = 1,\ldots,H$. Weeks are the time steps of the stochastic decision process and the dynamic programming model. Hours are internal details in the weekly minimisation sub-problem (2.17)-(2.18) and (3.22)-(3.26). In our formulation, the detailed weekly sub-problems do not have any genuine dynamic structure. The hours h are different loading situations, the order of which has no bearing on the results. This formulation includes the case in which the variation in power demand within a week is represented by a step function approximation of the duration function for the demand, and the index h labels its steps. The lengths of the hours (time steps within a week) h are denoted by $l(h)$.

As before, we make the notational convention that the state vector at the right end point of week t is given index t and denoted by $x(t)$. The state at the beginning and the end of the year are thus denoted by $x(0)$ and $x(T)$, respectively. The decision in week t is $u(t)$. These definitions are illustrated in Figure 4.1.
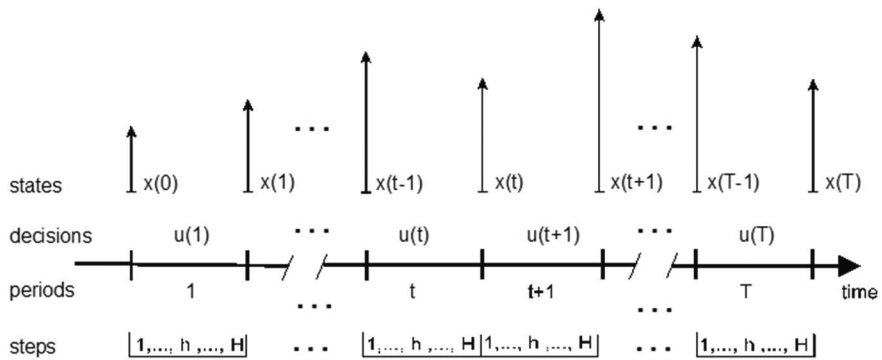
Figure 4.1. States x(t), decisions u(t), weeks t and hours h of the process.

4.2.2 Indices of the segments of the piecewise linear approximations of cost and value functions

Convex cost and value functions are approximated with piecewise linear functions in the model; see Figure 4.2 and Figure 4.3. Their segments are indexed as follows:

$\gamma = 1,\dots, \Gamma$ An additive component of thermal power production corresponding to a linear segment of the piecewise linear approximation of thermal production costs as a function of power produced. Typically $\Gamma = 5\dots10\dots$

$\lambda = 1,\dots, \Lambda$ Linear segment of the piecewise linear approximation of the water value function and of the total cost function of the problem. Typically $\Lambda = 10\dots$

- additive components $u_2(t,h,\gamma)$ of the thermal effect $u_2(t,h)$, and their upper bounds $f(t,\gamma)$.
- costs coefficient $c(t,\gamma) = \tan \alpha_\gamma(t)$ for the components $u_2(t,h,\gamma)$.
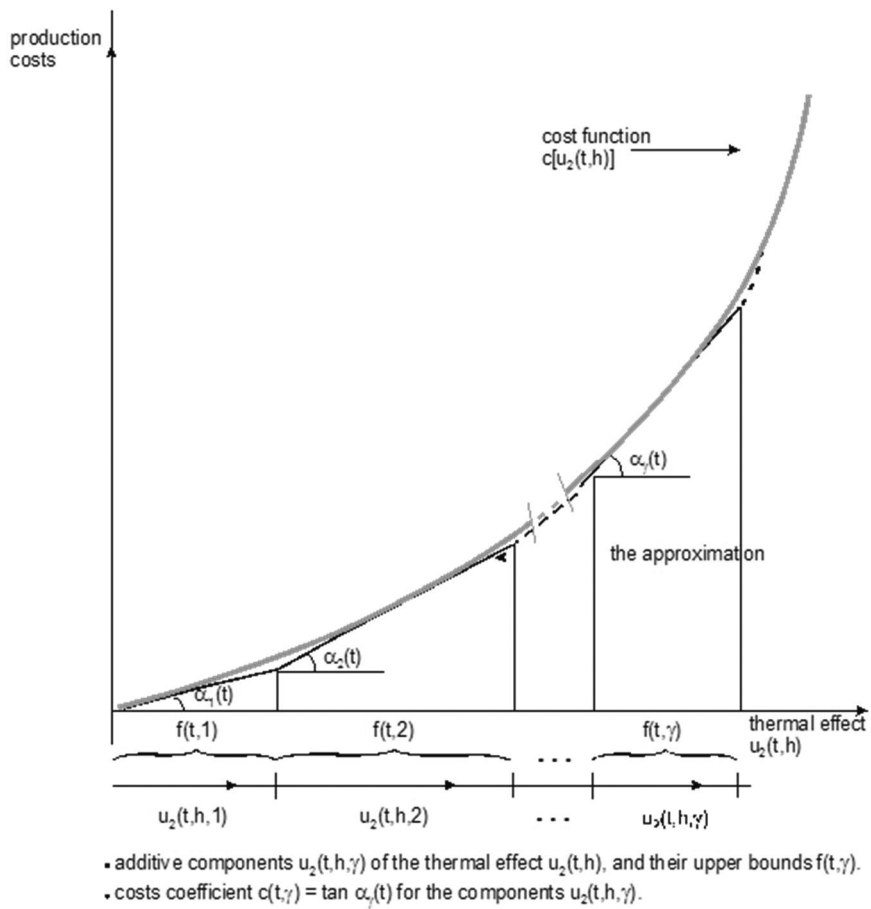
Figure 4.2. Piecewise linear approximation of the thermal production cost function.

- additive components $x_1(t,\lambda)$ of the contents $x_1(t)$ of hydro reservoirs, and their upper bounds $k(t,\lambda)$.
- costs coefficients $g(t,\lambda) = \tan \beta_\lambda(t)$ for the components $x_1(t,\tau)$.
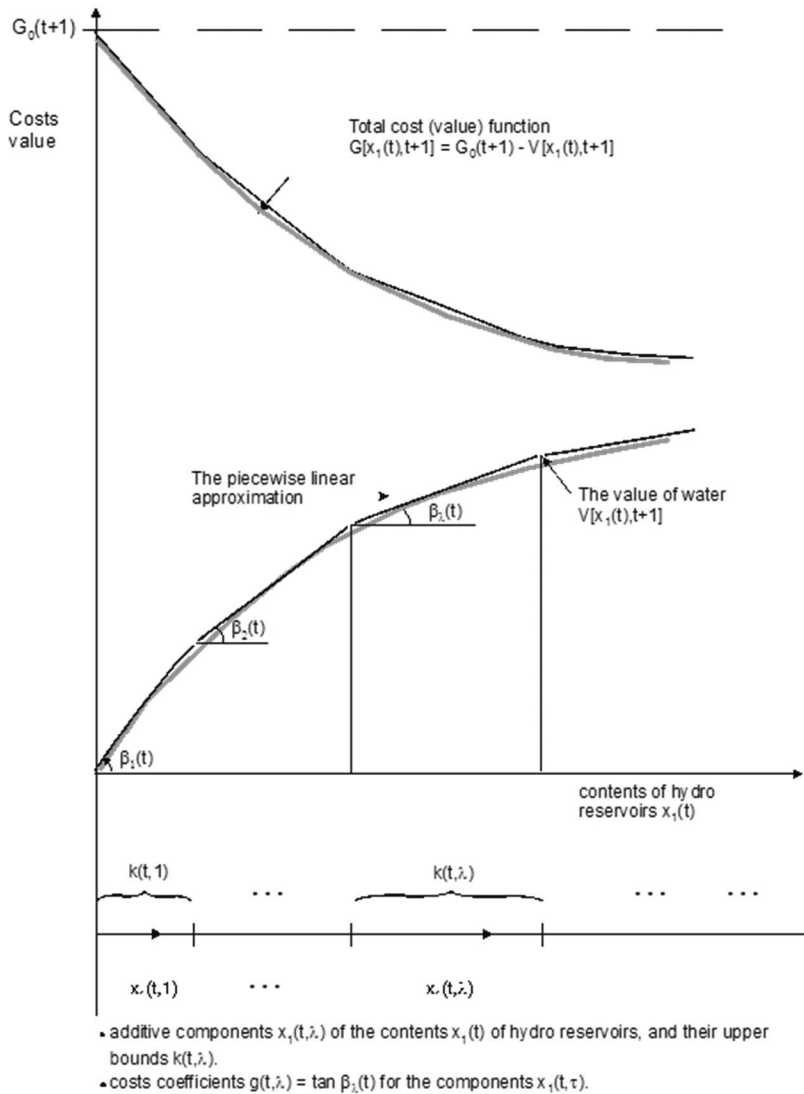
Figure 4.3. Piecewise linear approximation of the water value function and the total cost function.

4.3 **Variables in the weekly optimisation task**

4.3.1 Production (decision) variables for each week

For each week t, the control or decision vector, denoted by $u(t)$ cf. 2.1.1 and 3.1, has the following components:

$u_1(t,h)$	Hydropower effect in hour (sub-step) h of week t,
$u_1(t)$	Total hydro energy production in week t,
$u_3(t)$	Total bypass discharge of water from reservoirs in week t, measured in energy units,
$u_2(t,h)$	Thermal effect in hour h of week t,
$u_2(t,h,\gamma)$	Component γ of thermal effect in hour h, week t, and
$u_2(t)$	Total thermal energy production in week t.

Here, h = 1,…,H and γ = 1,…,Γ. If required, the constants H and Γ are allowed to depend on week t.

The basic weekly decision variables are the hydro effect $u_1(t,h)$ and the thermal effect $u_2(t,h)$ in each hour h of week t, and the total bypass discharge of water $u_3(t)$ in week t. The components of thermal power $u_2(t,h,\gamma)$ are auxiliary variables required to model the piecewise linear approximation of the thermal production costs; see Figure 4.2. The total weekly hydro and thermal energy production variables $u_1(t)$ and $u_2(t)$ are computed as sums over hourly values, taking into account the lengths of the hours (sub-steps) h. (In principle, they may be of unequal length.)

4.3.2   State variables for each week

For each week t, the state vector of the decision process, denoted by $x(t)$, cf. 2.1.1 and 3.1, has the following two components:

| $x_1(t)$ | Amount of water in reservoirs at the end of week t, measured in energy units, and |
| $x_2(t)$ | Relative level of total yearly electricity demand at the end of week t. |

These entities enter the detailed (lower level) optimisation model, the former as a variable determined by the production decisions and the latter as a constant in the optimisation problem.

The optimisation model also contains the following auxiliary variables:

| $x_1(t,\lambda)$ | Additive component λ = 1,…, Λ of the state of the water reservoir, corresponding to a segment of the piecewise linear approximation to the value function $G_{t+1}[x_1(t),\ x_2(t)]$ for a constant $x_2(t)$.
Unit: energy. |

The real variables $x_1(t)$ and $x_2(t)$ are thus the state variables in the dynamic programming recursion. Only $x_1(t)$ is a variable in the optimisation sub-problem. The

dynamic equation for the second component of the state, the relative level of the demand $x_2(t)$, reads $x_2(t) = x_2(t-1) + s_2(t,i)$. Consequently, $x_2(t)$ is determined by the previous state and the random decision and it is constant in the optimisation sub-problem; cf. Section 3.4. The components $x_1(t, \lambda)$ of the contents of the water storage are auxiliary variables in energy units required to represent the piecewise linear approximation of the value function; see Figure 4.3.

4.4 Constraints

4.4.1 Constraints containing only production variables

Lower and upper bounds:

For each week t and each hour h, the hydro effect has a given lower bound and upper bound. The bounds may depend on the week, the hour, the contents of the hydro reservoir and the realisation of the stochastic element $s(t)$.

$$a(t) \leq u_1(t,h) \leq b(t), \; for \; all \; h=1,..,H \tag{4.1}$$

The components of thermal power have upper bounds (4.2) determined by the piecewise linear approximation to the cost function (the bound is the length of the corresponding linear segment of the approximation). The component with the highest cost, $u_2(t,h,\Gamma)$, is unbounded, so that the optimisation problem always has a feasible solution.

$$u_2(t,h,\gamma) \leq f(t,\gamma), \; for \; h=1,...,H \; and \; \gamma = 1,...,\Gamma-1 \tag{4.2}$$

Functional constraints:

The production covers the demand for electricity every hour:

$$u_1(t,h) + u_2(t,h) \geq d(t,h), \; for \; h=1,..,H \tag{4.3}$$

where $d(t,h)$ is the hourly demand. It is a constant parameter in the optimisation problem, determined by the level of the yearly demand reached in week t.

Weekly hydro energy production is the sum of its components:

$$u_1(t) = \sum_{h=1}^{H} l(h)u_1(t,h) \tag{4.4}$$

where $l(h)$ is the length of hour h.

Hourly thermal power is the sum of its components,

$$u_2(t,h) = \sum_{\gamma=1}^{\Gamma} u_2(t,h,\gamma), \ \ for \ h=1,...,H \tag{4.5}$$

and the weekly thermal energy production is computed as the sum

$$u_2(t) = \sum_{h=1}^{H} l(h)u_2(t,h) \tag{4.6}$$

4.4.2 Constraints that also contain state variables

Lower and upper bounds:

The contents of the hydro reservoir at the end of week t have a lower and an upper bound:

$$A(t) \le x_1(t) \le B(t), \ \ for \ all \ t=1,..,T \tag{4.7}$$

These bounds can be physical, and they can also be used to control the development of the solution.

The components of the state have upper bounds, the lengths of the segments of the piecewise linear approximation of the value function:

$$x_1(t,\lambda) \le k(t,\lambda), \ \ for \ \ \lambda=1,...,\Lambda-1 \tag{4.8}$$

Functional constraints:

The dynamic equation for the development of the state of the water reservoir:

$$x_1(t) = x_1(t-1) - u_1(t) - u_3(t) + s_1(t,i) , \ or \ by \ arranging \ the \ terms$$

$$x_1(t) + u_1(t) + u_3(t) = x_1(t-1) + s_1(t,i), \ for \ all \ t=1,..,T \tag{4.9}$$

where $s_1(t,i)$ is the realisation of the total inflow of water into the reservoirs in week t. Note that in the latter form, the variables in the dynamic programming optimisation problem for week t are on the left-hand side of (4.9), and the right-hand side is constant in the optimisation task.

The state is the sum of its components:

$$x_1(t) = \sum_{\lambda=1}^{\Lambda} x_1(t,\lambda) \tag{4.10}$$

4.5   The objective function and the optimisation sub-problem

The objective of the optimisation sub-problem to be solved in the dynamic pro-gramming iteration for every week t, for every feasible starting state $x(t\text{-}1)$ at the beginning of the week and for every realisation $s(t,i)$ of the random vector associated with week t is a piecewise linear approximation of the convex objective function (3.21) and reads:

$$\min\left\{\sum_{h=1}^{H} l(h)\sum_{\gamma=1}^{\Gamma} c(t,\gamma)u_2(t,h,\gamma) + G_0 - \sum_{\lambda=1}^{\Lambda} g(t,\lambda)x_1(t,\lambda)\right\} \qquad (4.11)$$

where $c(t,\gamma)$ and $g(t,\lambda)$ are cost coefficients. The coefficients $c(t,\gamma)$ are determined by the piecewise linear approximation to the thermal production cost function $c_t[u_2(t,h)]$ for week t, which again is determined by the thermal capacity available in that week; see Figure 4.2. The cost coefficients $g(t,\lambda)$ are the coefficients of the piecewise linear approximation to the water value function at the beginning of the next week, $V_{t+1}[x_1(t)]$; see Figure 4.3. The coefficients are determined when the optimal cost function for the week $t+1$ has been computed. The function $V_{t+1}[x_1(t)]$ and the coefficients $g(t,\lambda)$ depend on the level of the demand at the end of the week, $x_2(t)$.

G_0 is the value of $G_t[x_1(t), x_2(t)]$ at the point $x_1(t) = 0$, i.e. G_0 equals the minimum of the partial objective function, from the beginning of period $t+1$ to the end of the period, if the initial hydro reserve is empty. It is constant in the optimisation, and it depends on the demand $x_2(t)$.

The optimisation sub-problem is a linear programming problem with (4.11) as the objective function and the constraints (4.1)-(4.10).

5. Hydropower

5.1 The basic model of hydro production

The treatment of hydropower was presented in Chapter 3 as part of the overall market model. In this chapter, the modelling of hydropower is looked at in more detail. For clarity, the straightforward model structure required to represent the subdivision of weeks into hours is omitted in the presentation.

All of the hydro production in the Nord Pool area is aggregated into a single system with the following variables (all in energy units):

$x_1(t) =$ Amount of water in the aggregated hydro reservoir at the end of week t,

$S_1(t) =$ Inflow of water into the hydropower system during week t,

$u_1(t) =$ Production of hydropower (electricity) during week t, and

$u_3(t) =$ Release of water from the reservoir past turbines during week t.

The level of the reservoirs $x_1(t)$ is a state variable in the dynamic programming iteration. $S_1(t)$ is a stochastic control variable, the total inflow of water into the hydro system, i.e. it contains both the storable inflow into the aggregated hydro reservoir and the run-of-the-river inflow, which has to be used in power generation as it flows into the rivers or is released past the turbines. The production decision variables $u_1(t)$ and $u_3(t)$ form the active control of the hydro system.

The dynamic development of the system is governed by the difference equation (5.1)

$$x_1(t) = x_1(t-1) + S_1(t) - u_1(t) - u_3(t), \text{ for all } t = 1, \dots, T \tag{5.1}$$

The inflow random variables form a stochastic process $S_1 = \{S_1(t), t = 1, \dots, T\}$. In the basic formulation of the model, the successive inflow random variables $S_1(t)$, for different weeks t, are statistically independent, and their distributions depend only on time t. In principle, the distribution of $S_1(t)$ may depend of the state $x(t-1)$.

The variables satisfy a number of constraints. The initial state of the reservoir is given, and all the variables are non-negative. The contents of the aggregated water reservoir satisfy the lower and upper bounds (5.2).

$$a(t) \leq x_1(t) \leq b(t), \ for \ all \ t = 1,...,T \qquad (5.2)$$

The production of hydropower has an upper bound (5.3),

$$u_1(t) \leq r(t), \ for \ all \ t = 1,...,T \qquad (5.3)$$

and the total flow in the aggregated hydro system has a lower bound (5.4),

$$d(t) \leq u_1(t) + u_3(t), \ for \ all \ t = 1,...,T \qquad (5.4)$$

In principle, the upper and lower bounds in (5.2)-(5.4) depend on the time t of the state $x_1(t)$ and the realisation $s_1(t)$ of the stochastic element $S_1(t)$. These bounds and the distribution of the inflow process S_1 constitute the most important basic data for the system. Owing to the high level of aggregation, the estimation of the distribution of inflow and the upper and lower bounds in (5.2)-(5.4), and the validation of the model constitute extensive research tasks. For example, the lower level $d(t)$ of the total flow in the system is determined by the run-of-the-river flow and hydrological regulations within the whole production system.

5.2 Water flow data

The numerical parameters of the aggregated model of the Nordic hydro system are based on the integrated behaviour and properties of a large number of partially interconnected individual systems and components: power plants, water reservoirs and rivers. The structural properties and interconnections of the component systems, the capacities of the power plants and reservoirs, regulations on water levels in the reservoirs and water flows in the rivers determine the operational properties of the subsystems and the aggregated system. For a relatively small hydro system, the overall operational parameters, such as total available power capacity or minimum flow, are closely based on the technical and operational characteristics of the system components. For the aggregated Nordic system, the parameters have to be estimated on the basis of production statistics.

In principle, the input data on the flow of water used in the model are the probability distribution of the discrete stochastic inflow process $S_1 = \{S_1(t), t = 1,...,T = 52\}$ formed by the sequence of total weekly inflow random variables $S_1(t)$ over a year. One possibility, which we have frequently used, is to define the distribution approximately through a representative sample of N equally probable realisations of the yearly flow process:

$$\{s_1(n,t), t = 1,...,T = 52\}, n = 1,...,N \qquad (5.5)$$

i.e. the input is N weekly inflow time series $\{s_1(n,t), t = 1, ... ,52\}$, each with the probability 1/N.

However, this is only one very straightforward definition of the inflow process. In the following, we examine briefly the problem of estimation and the use of inflow data in the model.

In the optimisation phase (computing the value functions), for each week *t,* the input data used is the following discrete approximation of the probability distribution of total water inflow into the system during the week:

$$S_1(t) = s_1(i,t) \text{ with probability } \pi(i,t), \text{ for } i = 1,...,I, \text{ and for all } t = 1,...,T \quad (5.6)$$

i.e. the input data consist of the flow values *s₁(i,t)* and the probabilities *π(i,t)* for all *i = 1,...,I,* and for all *t = 1,...,T.* Clearly, the distributions of the weekly flows formed directly from the sample (5.5) are a special case of (5.6).

In the simulation phase, realisations of the market price are generated, and these simulations are based on a representative sample of the realisations of the inflow process. Here, we can use the sample (5.5), a modification or subset of the sample (e.g. corresponding to a wet or a dry year), or we can use a sample generated with a mathematical model of the inflow process.

It is an important research task to collect information about the weekly distribution of the inflow of water into the Nordic system to form the discrete approximations and to generate representative samples of the yearly random inflow process. In the present version of the model, the correlation between the inflows of water in successive weeks *t* is not represented in the optimisation phase, and the weekly inflows are statistically independent of each other. However, the realisations (5.5) based on the actual flow data used in the simulation phase show proper correlation in time between the realisations of the weekly inflow variables.

As stated earlier, for the weekly distributions (5.6), we can use the distribution contained in the sample directly or apply the following approach:

1. Analyse the available national and regional weekly inflow data. For the statistical region under study, for every week *t* of the year, compute the mean values $m_s(t)$ and the variances $\sigma_s(t)$ in the weekly inflows and the mean value M_s of the yearly inflow. The subscript *s* indicates that these parameters are computed on the basis of a statistical sample. A closer analysis shows that the distribution of the weekly inflow can often be approximated with a log-normal distribution. Approximation with a normal distribution is usually also possible.

2. Form the index series

$$\alpha_s(t) = m_s(t) / M_s, t = 1,...,T \qquad (5.7)$$

$$\beta_s(t) = \sigma_s(t) / m_s(t), t = 1,...,T \qquad (5.8)$$

The index series $\alpha_s(t)$, $t = 1,...,T$ gives the ratio of the expected weekly and the expected yearly inflow. The index series $\beta_s(t)$, $t = 1,...,T$ gives the ratio between the variance and the mean value of the weekly inflow.

3. The index series based on a sample shows random fluctuations. Smooth out the index series $\alpha_s(t)$ and $\beta_s(t)$ in an appropriate way and obtain smoothed index series $\alpha(t)$ and $\beta(t)$. In smoothing out the series, the condition $\Sigma_{t=1}^T \alpha(t) = 1$ must always be satisfied.

4. Estimate the mean (expected) value M of the total yearly inflow of water into the aggregated Nordic system. Assume that the total Nordic inflow has the same statistical distribution as the sample. The inflow $S_1(t)$ in week t is assumed to have a normal or, alternatively, a log-normal distribution with mean $\alpha(t)M$ and standard deviation $\beta(t)\alpha(t)M$. Finally, generate an appropriate discrete approximation of the distribution of the flow to be used in the model.

We have also tested various mathematical models of the stochastic inflow process in the construction of samples of realisations of the inflow. A completely satisfactory solution has not been found.

5.3 Modelling correlation between successive water inflow terms

One of the restrictions of the aggregated model in Section 5.1 is the fact that the correlation between successive weekly inflows $S_1(t)$ cannot be taken into account in the model with only one state variable associated with hydropower. However, it has to be remembered that this does not mean that the model would be based on incorrect data about the distribution of inflow. The distribution of every inflow variable is correct, but the model does not use additional information about the simulated partial realisation of the stochastic inflow process to update and modify the distribution in the dynamic programming iteration. The results are not incorrect but they can be characterised as more average than what the actual state of information would allow. What is said above concerns the stochastic dynamic optimisation (computation of the value functions) phase only. When the model and the computed value functions are used in the simulation mode to generate realisations of the price of electricity, then these can be based on realisations of the inflow process with proper correlation between the successive terms. The development, estimation and validation of such a model for the aggregated inflow are an independent research task.

The correlation between successive weekly inflow values can be modelled as follows. We augment the system with additional storage consisting of snow and water on the way to the hydro system and define additional variables:

$x_s(t) =$ Amount of snow and water in storage and on the way to the hydro-power system at the end of week t, measured as net electrical energy in production, and

$S_s(t) =$ Precipitation of water and snow into the reserve $x_s(t)$ during week t, in the same units.

The amount $x_s(t)$ is a state variable and $S_s(t)$ is a stochastic control variable. $S_s = \{S_s(t), t = 1, \dots, T\}$ is a stochastic process with statistically independent terms $S_s(t)$. The dynamic development of the snow and water storage is now governed by the state equation:

$$x_s(t) = x_s(t-1) + S_s(t) - S_1(t), \ for \ all \ t = 1,\dots,T \qquad (5.9)$$

The inflow of water into the hydro system $S_1(t)$ is no longer an independent component of the stochastic control vector. It is a dependent variable determined by the previous state and time t. It will be assumed that a time-dependent portion of the snow and water storage will be converted into flow of water into the hydro system:

$$S_1(t) = \gamma(t)x_s(t-1), \ for \ all \ t = 1,\dots,T \qquad (5.10)$$

This simple augmented model can be motivated on physical grounds. It generates correlated inflows into the hydro system. The estimation of the coefficients $\gamma(t), t = 1, \dots, T$ and the distribution of the process $S_s = \{S_s(t), t = 1, \dots, T\}$ are the essential difficulty in applying the model to the aggregated Nordic hydro system, due to the limited availability of relevant data. It would be easier to validate such a model first for a small hydro system with uniform weather conditions in the precipitation area and good weather and water inflow statistics.

6. Thermal power

6.1   Modelling thermal power in the electricity market simulation model

We consider the operation of thermal power in the system during the weeks. The thermal power system consists of a large number of plants that may be grouped into production classes consisting of similar plants. In the model of the electricity market it is always assumed that the production of the available thermal power plants is offered to the market at given, plant-dependent prices and that the market leads to cost-optimal load dispatching in the system. Note that the cost parameters can be freely chosen, plant by plant and week by week, if needed. In fact, the cost parameters may, in principle, also vary within the week. That gives the analyst considerable freedom in modelling the market behaviour of different producers.

   We start with complete lists of power plants for each week, giving all the plants in the system that are manned and ready for operation during the week. These data are based on the production plans of the plants, which are assumed to be known. The operation of the plants is subject to random disturbances and, as a result of these, all plants that have planned to be ready for operation are not actually available for production during the week. The total available capacity of the system is a random variable, and the minimal production costs with the available plants as a function of power produced is a random function. The collection of available plants is determined by random choice, according to the availabilities of the plants, from the set of plants ready for production during the week. Once this choice has been made, the total available capacity (a real scalar) and minimal production costs as a function of power produced (a real-valued function) can be computed for this realisation of available plants in the system.

   The probability distributions of the available capacity and minimal production cost function constitute a complete aggregated stochastic production model of the power system. The most central characteristics of the model are expected minimal production costs $G(y)$ and marginal expected production costs $F(y) = G'(y)$ for the system, as functions of power produced y. The present version of the market simulation model is based on these expected cost functions.

In Sections 6.2–6.3, we first derive the probability distributions of the available capacity and optimal production costs on the basis of plant data. In Section 6.4, the expected optimal production costs for a system of plants are computed recursively, applying dynamic programming. In Section 6.5, we compute expected optimal cost functions for a production system consisting of production classes, where each class is characterised by a common production cost and the probability distribution of its available capacity.

6.2 Computation of the available capacity and optimal production costs of a thermal power system on the basis of plant data

The system consists of a number N of plants. Each plant is either available for production or out of production during the whole time step. Partial unavailability is thus not considered. The probability p_n of outage of the plant n is called the forced outage rate, and the complementary probability $1 - p_n$ is called availability. The installed capacity of the plant is K_n, and the production cost of each plant is a linear function of power produced with a cost coefficient c_n.

The available capacity of plant n is a random variable X_n, taking the values 0 and K_n with probabilities p_n and $1 - p_n$. The expected capacity of the plant n is $E\{X_n\} = (1-p_n)K_n$. The total available capacity X of the system is the sum of the available capacities X_n, and without any further assumptions it follows that the expected value of X is

$$E\{X\} = \sum_{n=1}^{N} E\{X_n\} = \sum_{n=1}^{N} (1 - p_n)K_n \tag{6.1}$$

The probability distribution of the total capacity X is determined by the joint distribution of the available capacities X_n of the individual plants n. If the outages of the plants n are causally independent, then the random variables X_n are statistically independent. The distribution of X is in this case a modified binomial distribution, and it can often be approximated with a normal distribution. In case of statistical independence, the variance of the total capacity is the sum of the variances in the plant capacities (6.2).

$$\sigma^2(X) = \sum_{n=1}^{N} \sigma^2(X_n) = \sum_{n=1}^{N} p_n(1 - p_n)K_n^{\,2} \tag{6.2}$$

It is assumed that each plant n, available for production, can be operated during the whole time step at any power level between 0 and the plant capacity K_n. In the optimal operation of the system, most of the plants either produce at full capacity or are shut down, and only one plant is operated at an intermediate power level between its minimum and maximum capacity. Thus, the approximation, that the minimum power levels of the plants equal zero, is acceptable.

The set-up of the computational problem is thus as follows: each plant is available for production during the time interval considered with a given probability. First, random choice, according to these probabilities, determines which plants in the system are actually available for production during the time step. These plants form the collection (configuration) E of available plants. Then, the demand y is allocated optimally to these plants, and optimal (minimal) production costs are computed for the collection E as a function of the demand y. By repeating the analysis for all possible configurations E, or for a statistically representative sample of configurations, and considering their probabilities of realisation, the probability distributions of the available capacity and minimal production costs as a function of demand are obtained.

6.3 Direct computation of the probability distributions of the available capacity and production costs

The probability P_E of the realisation of the configuration E, where plants *a, b, c,* ... are available, the plants *u, v, w,* ... are not, and the outages of the plants are not correlated, is given by

$$P_E = (1 - p_a)(1 - p_b)(1 - p_c)...p_u p_v p_w...$$ (6.3)

The available capacity of the configuration E, $K_E = K_a + K_b + K_c + ...$, and the probability distribution of the available capacity of the system are determined completely by the probabilities P_E and the corresponding available capacities K_E for all the configurations E. A good approximation can be generated fast with a simple stochastic simulation system. The expected value and the variance in the available capacity are given by (6.1) and (6.2), respectively.

Minimal production costs for the configuration E and for any demand y are determined by loading the available plants of the configuration in the order of rising production costs c_i. The minimal costs as a function of the demand y are thus given by a piecewise linear convex function. These functions for all the configurations E, together with their probabilities P_E, completely determine the distribution of the production costs of the power system for any demand.

The expected minimal cost function $G(y)$ for the whole system is the weighted average of the cost functions of the different configurations E with the weights P_E, (6.3). As the number of different configurations for a system of N plants is 2^N, it follows that direct computation is only possible for relatively small production systems. Approximations with a desired accuracy can be generated through random sampling.

The fact that the load y can exceed the available capacity requires special attention. In that case, the load cannot be covered at any cost, i.e. the cost is infinite. As there is a finite (usually small) probability that all the plants in the system are out of production, the expected cost according to the straightforward computation would be infinite for every load y. Thus, at each load level y, we ought to con-

sider separately the event of insufficient capacity to meet the load and the event that capacity covers the load. Expected costs were only defined in the latter case.

However, these complications can technically be avoided if we augment each production system with a back-up plant with infinite capacity, zero forced outage rate and a unit production cost higher than the costs of all other plants in the system. Note that the production cost of a plant with infinite capacity and zero outage rate is always an upper limit of the marginal optimal production cost of the whole system, as no plant with higher production costs ever contributes to an optimal production programme of the system.

6.4 Computation of expected production costs through dynamic programming on the basis of plant data

Applying Bellman's principle of optimality (dynamic programming), the expected total and marginal production costs as a function of the demand can be computed exactly and fast for a production system. We examine the system defined above. Data for the system are given in Table 6.1.

Table 6.1. Power plant data for a production system.

Plant	Prod. cost	Max. capacity	Forced outage rate
1	c_1	∞	0
2	c_2	K_2	p_2
...	...	...	...
N	c_n	K_n	p_n
...	...	...	...
N	c_N	K_N	p_N

When applying dynamic programming in the computation, the following two facts are essential: (i) the plants are arranged in the order of decreasing (non-increasing, to be exact) production costs, and (ii) the first plant, with the highest production costs, has an infinite maximum capacity and zero forced outage rate. It is also assumed that the lower bound of the production power of each plant is 0. As discussed earlier, this approximation has only a small effect on the results. The production cost parameters c_i are defined as costs per power and time unit.

To start the iteration of expected minimal costs (value functions in the language of dynamic programming), we thus assume that in the production system there is always a source of supply that can deliver any amount of electric power, and this source has a higher production cost than any other plant in the system. We arrange the loading of the plants and the computation of the expected minimal production costs as follows: first we consider the subsystem consisting of only the first most expensive plant, next the system consisting of the first two most expensive

plants, etc. Plants are added to the system in the order of decreasing (non-increasing) costs.

We define the expected cost functions G_n and F_n for any power level y in the infinite interval $[0, \infty)$ as follows:

$G_n(y)=$ expected minimal costs in the time unit for the production of a total effect y with the plants $1, \dots, n$ in the system, and

$F_n(y)=$ marginal expected minimal costs in the time unit at effect y when that effect is covered optimally with the plants $1, \dots, n$ in the system.

$F_n(y)$ is the derivative of G_n at y if the latter is differentiable at that point. It can easily be seen that there is only a finite but, in general, large number of points where G_n fails to be differentiable. It has left and right derivatives, however, at points y where it is not differentiable. At such a point, $F_n(y)$ is discontinuous, and it has left and right limits equal to the left and right derivatives of G_n at y.

Assume that the function G_{n-1} is known for all y. Then, we can construct the function G_n as follows: as the plant n is added to the system, the probability that this plant is not available is p_n. In that case, for every y, the expected optimal production costs $g_{n1}(y)$ are the same as with the previous system consisting of the plants $1, 2, \dots n - 1$ and given by

$$g_{n1}(y) = G_{n-1}(y) \tag{6.4}$$

The probability that the nth plant is available is $(1 - p_n)$. It can be operated at any power between 0 and the capacity K_n. As its costs are lower than (or equal to) the costs of any plant in the previous system with expected optimal costs $G_{n-1}(y)$, then it is optimal to load plant n first, and the expected optimal costs $g_{n2}(y)$ are as follows:

$$\begin{aligned} g_{n2}(y) &= c_n y, & \text{if } 0 \le y \le K_n \\ &= c_n K_n + G_{n-1}(y - K_n), & \text{if } K_n < y \end{aligned} \tag{6.5}$$

Further, directly from the definitions it follows that:

$$G_n(y) = p_n g_{n1}(y) + (1 - p_n) g_{n2}(y) \tag{6.6}$$

Inserting (6.4) and (6.5) into (6.6), we get

$$\begin{aligned} G_n(y) &= p_n G_{n-1}(y) + (1 - p_n) c_n y, & \text{if } 0 \le y \le K_n \\ &= p_n G_{n-1}(y) + (1 - p_n)[c_n K_n + G_{n-1}(y - K_n)], & \text{if } K_n < y \end{aligned} \tag{6.7}$$

By applying the recursion (6.7), the cost functions G_n are computed for any $n = 2, \dots, N$.

According to the definition of the problem, the first cost function G_1 has the simple form, giving a starting point for the recursion:

$$G_1(y) = c_1 y, \ for \ all \ y \tag{6.8}$$

The same argument can be repeated for the marginal cost functions F_n, or we can take the derivative of (6.7) on both sides and, for marginal cost functions, we obtain the recursion formula:

$$\begin{aligned} F_n(y) &= p_n F_{n-1}(y) + (1 - p_n)c_n, &&if \ 0 \le y \le K_n \\ &= p_n F_{n-1}(y) + (1 - p_n)F_{n-1}(y - K_n), &&if \ K_n < y \end{aligned} \tag{6.9}$$

The iteration begins with the first marginal cost function F_1 having the simple form:

$$F_1(y) = c_1 \tag{6.10}$$

The cost functions G_n and F_n computed by the recursive method of dynamic programming are exact, i.e. no further approximations or assumptions in addition to those explicitly stated above, and made in the derivation of the functions. Note that the marginal cost functions F_n are also called EIC functions in the literature (from Expected Incremental Cost).

6.5 Computation of expected costs for a power system on the basis of production class data

6.5.1 Basic data for homogenous production classes

In the computation of expected costs for a production system in the previous section, every power plant was defined individually by its capacity, outage rate and costs. In this chapter, we derive expected cost functions for more aggregated production models consisting of production classes composed of similar plants with common parameters such as coal-fired condensing power, gas turbines, etc. Wind power can be modelled in a similar way.

We consider production classes (sources of electricity) $n = 1,...,N$. Each class n has a total available capacity X_n, which is a real-valued, non-negative random variable with probability density function $f_n(x)$ and distribution function $\Phi_n(x)$ defined on the non-negative real line $[0, \infty)$. The production costs within a class are homogenous, and the common unit production cost (i.e. cost per power and time unit) for the class n is c_n. Demand y is allocated to the production system. As earlier, in the system there is a backup source with unlimited capacity and a unit production cost that is higher than the cost of any other source in the system.

6.5.2　Computation of expected costs for a single production class

First, we examine a production system consisting of a single production class and a back-up source and drop the subscripts here. Data for the class are given by X, $f(x)$, $\Phi(x)$ and c. For any demand y, if the capacity of the class $X = x$ and $x \le y$, then the amount x is produced by the production class, and the amount $(y - x)$ by the back-up source with unit cost h. The costs in time unit g in this case are

$$g = cx + h(y - x) \tag{6.11}$$

If the demand is smaller than the capacity of the class, $y < x$, then the whole demand y is produced by the class, and the costs per time are

$$g = cy \tag{6.12}$$

Expected costs are

$$G(y) = \int_0^y [cx + h(y - x)] f(x) dx + cy \int_y^\infty f(x) dx$$
$$= (c - h) \int_0^y x f(x) dx + hy \int_0^y f(x) dx + cy \int_y^\infty f(x) dx \tag{6.13}$$

which gives, after some elaboration:

$$G(y) = cy + (h - c) \int_0^y \Phi(x) dx \tag{6.14}$$

The cost function G is convex as expected. The marginal costs F are given by the derivative, which again is an increasing (non-decreasing) function:

$$F(y) = G'(y) = c + (h - c)\Phi(y) \tag{6.15}$$

6.5.3　Expected costs for a system consisting of several production classes

Now, we analyse a production system consisting of several production classes $n = 1,...,N$. As in the treatment in Section 6.4, the classes are arranged in the order of decreasing production costs.

The problem is again to compute the expected production costs per time unit for any demand y, when this demand is covered optimally by the production system. The value functions are defined as in Section 6.4 with plants replaced by produc-

tion classes. Now, suppose that the expected costs (the value function) $G_{n-1}(x)$ have already been computed.

Then, we augment the system with the production class n, having a lower or at most equal unit production cost to that of the production classes already in the system. The demand is y, and suppose that the capacity of the class n, $X_n = x$. If $x \leq y$, then the amount x is produced by the class n, and the amount $y - x$ is produced optimally by the classes $1,...,n-1$. Optimal costs for producing the amount y when the available capacity of class n is x are

$$g_1(x, y) = c_n x + G_{n-1}(y - x) \qquad (6.16)$$

If $y < x$, then the whole demand y is produced by class n, and the costs are

$$g_2(y, x) = c_n y \qquad (6.17)$$

The expected optimal costs are

$$G_n(y) = \int_0^y [c_n x + G_{n-1}(y - x)] f_n(x) dx + \int_y^\infty c_n y f_n(x) dx$$

$$= c_n \left[y - \int_0^y \Phi_n(x) dx \right] + \int_0^y G_{n-1}(y - x) f_n(x) dx \qquad (6.18)$$

This gives the recursion for the solution of the expected minimal cost function G_n on the basis of the previous cost function G_{n-1}, and the distribution of the available capacity of the production class n, i.e. the functions f_n and Φ_n. Assuming that the first class has unlimited capacity and the unit production cost c_1, the iteration is started from

$$G_1(y) = c_1 y, \; \text{for all} \; y \qquad (6.19)$$

Taking the derivatives of (6.18) and (6.19) with respect to y, and recalling that $G'_n(x) = F_n(x)$, we obtain the recursion for the marginal cost functions F_n:

$$F_n(y) = c_n [1 - \Phi_n(y)] + \int_0^y F_{n-1}(y - x) f_n(x) dx \qquad (6.20)$$

and

$$F_1(y) = c_1, \; \text{for all} \; y \qquad (6.21)$$

The cost functions G_N and F_N give the results for the whole production system. The recursion results (6.18)-(6.21) for the computation of the expected costs are general, valid for any distributions $\Phi_n(x)$ and $f_n(x)$. In practice, the distribution can be a normal distribution, some modification of a normal distribution or a binomial distribution. In the notation, we have assumed that the distributions are continuous. However, it is easily seen that the results are not dependent on this assumption, and the distributions may be discontinuous. The recursive formulae (6.18)-(6.19) and (6.20)-(6.21) are generalisations for production classes of the corresponding results (6.7)-(6.8) and (6.9)-(6.10) respectively, for discrete power plants.

7. Wind power and co-production power

7.1 General

There are a large number of wind power plants in the Nord Pool market area, and they are situated in different locations throughout the geographical area. Mathematically, the production (power) of every wind power station in the system can be modelled as a stochastic process determined by the capacity and other technical characteristics of the plant, its availability for production and the local weather conditions (wind). The total production of wind power in the system is the sum of all the plants, a stochastic process completely determined by its components.

    Neither the geographic distribution of production and consumption of electricity nor the transmission capacities between different areas are represented in the market model. Geographically, it is a point model. It cannot be used to study any questions to which the answer depends essentially on the location of production and consumption in the net (market) area. The treatment of wind power in the market model is based on the total aggregated production from this source.

7.2 Models for wind power in the market model

7.2.1 Stochastic models of wind power production

We examine a time period of one year. The total production of wind power in the system

$$Q^1 = \left\{ Q^1(s), s = 1,...,S \right\} = \left\{ Q^1(t,h), t = 1,...,T, h = 1,...,H \right\} \qquad (7.1)$$

is a stochastic process where $Q^1(s) = Q^1(t,h)$ is production in hour s (hour (t,h)). The index s labels the hours of the year consecutively, $s = 1,...,8760 = S$. In the model, we apply two-level indexing, where t labels the weeks, $t = 1,...,52 = T$, and h gives the hours within the week, $h = 1,...,168 = H$. The statistical distribution of the stochastic process Q^1 is completely determined by the technical characteristics of the wind power system and by overall weather conditions. The distribution of Q^1 is given in principle.

An hourly stochastic time series is too detailed a representation of wind power. A practical alternative is to aggregate the production into weekly figures and to compute the detailed distribution of production over the weeks using index series. We define the sequence

$$Q^2 = \left\{ Q^2(t) = \sum_{h=1}^{H} Q^1(t,h) \right\}, t = 1,...,T \tag{7.2}$$

The weekly wind power production $Q^2(t)$ for $t = 1,...,T = 52$ is a random variable, and the time series Q^2 is a stochastic process, weekly production of wind power over the year. The process Q^2 is completely determined by the basic process Q^1. For any realisation $q^2(t)$ of the production $Q^2(t)$ of week t, the hourly wind energy production figures $q^1(t,h)$ are now computed as expected values:

$$q^1(t,h) = \beta(t,h)q^2(t) \tag{7.3}$$

for $h = 1,...,168$, where the production index series $\beta(t,h)$ is given for all t and h by

$$\beta(t,h) = \frac{E\{Q^1(t,h)\}}{E\{Q^2(t)\}} = \frac{E\{Q^1(t,h)\}}{\sum_{h=1}^{H} E\{Q^1(t,h)\}} \tag{7.4}$$

The production index series defines the average distribution of wind power production within the week in question. In the definition (7.4), we have written the week t of the year and the hour h of the week as arguments. In practice, the index series depend only on the season of the year and the hour of the day. These indices are, in principle, determined by wind power statistics. In practice, it may be difficult to obtain the necessary data.

Aggregation of data may be carried one step further. The yearly energy production Q^3 is given by

$$Q^3 = \sum_{t=1}^{T} Q^2(t) \tag{7.5}$$

The yearly production is a random variable. If we choose Q^3 as the basic variable for the representation of wind power, then the total energy production during week t is computed from

$$q^2(t) = \alpha(t)q^3 \tag{7.6}$$

for every week $t = 1,...,T$. In (7.6), q^3 is a realisation of the total yearly energy production Q^3, $q^2(t)$ is the corresponding realisation of production in week t, and the

weekly production index series, {α(t)}, completely defined by production statistics, is given by:

$$\alpha(t) = \frac{E\{Q^2(t)\}}{E\{Q^3\}} = \frac{E\{Q^2(t)\}}{\sum_{t=1}^{T} E\{Q^2(t)\}} \tag{7.7}$$

7.2.2   Use of stochastic wind power models in the market model

The weekly subtask (the weekly optimisation module) of the market model does not contain any explicitly stochastic elements. In the weekly subtask, all stochastic variables and processes are represented by realisations of the stochastic elements or by their expected values. The detailed hourly stochastic process Q^1 defined in (7.1) cannot be used to represent wind power in the market model.

The production of wind energy can be represented in the market model by the stochastic process Q^2 defined in (7.2). The production of wind power in week t is then defined as a stochastic control variable $S_w(t)$.

$$S_w(t) = Q^2(t), t = 1,...,T \tag{7.8}$$

Note that this gives an additional stochastic control variable and increases the computational work correspondingly. Successive elements $Q^2(t)$ of the stochastic process Q^2 are assumed to be statistically independent, and they are independent of the other stochastic elements (inflow of water and change in demand). The distribution of $Q^2(t)$ is determined by production statistics, and in the market model a discrete approximation of the distribution is applied. The hourly production of wind power is computed from (7.3) and subtracted from the hourly demand. The difference is covered by thermal power.

If we do not want to have an additional state variable in the model to represent the stochastic variations in wind production, we choose the yearly production Q^3 as the basic variable and compute the yearly production time series from

$$q^1(t,h) = \beta(t,h)\alpha(t)q^3, \quad t = 1,..,T \quad and \quad h = 1,..,H \tag{7.9}$$

This time series is then subtracted from the demand.

7.3   Modelling wind power as a production class in thermal power

The wind power models in Section 7.2.2 form an hourly production time series that is subtracted from the demand for electricity before the loading of thermal power plants. It is also possible to model wind power and other similar power production, e.g. co-production power, as a production class in thermal power; see Section

6.5.3. Wind power production in week t is represented by the probability distribution $\Phi[Q^2(t)]$, and the corresponding probability density function $f[Q^2(t)]$ of the production of wind power $Q^2(t)$ *in* week t. $Q^2(t)$ is a random variable and its probability distribution $\Phi[Q^2(t)]$ is completely determined by the distribution of the stochastic process $\{Q^1(t,h),\ h = 1, ... H\}$; see (7.1). The production cost of wind power is taken to be zero in the model. Note that in this case the same (average) distribution of wind production is applied for all the hours of the week. If the distribution of the hourly production is available, it is also possible to form and apply different probability distributions for wind power production in different time segments of the week.

7.4 Co-production power

Co-production power is produced in connection with industrial heat production and district heating. The production of electricity is determined uniquely by the production (demand) of heat, and all electricity produced is fed into the net. Thus, we can apply exactly the same models as for wind power to co-production of electricity. Industrial co-production is determined by the level of primary production, and the production of district heating power by the heating demand, which again is influenced by the weather.

8. Demand, export and import of electricity

8.1 Short- and long-term variations in the demand for electricity

The hourly demand for electricity is perhaps the single most important factor determining the market price for electricity. The demand shows regular daily, weekly and yearly variations, as well as short- and long-term stochastic variations. If the market model is used to forecast the expected development and random variations in the price in the short term (some weeks), then a corresponding stochastic model of short-term development of the demand is required. These short-term variations are correlated with weather conditions. The model has mainly been used in long-term (several months or longer) studies, and in these studies the main source of uncertainty is the long-term development of the whole economy determining the demand for electricity. The next chapter presents a simple stochastic model for the long-term development of demand.

8.2 A stochastic model of the long-term development of demand for electricity

The basic variables of the model are

$d(t,h)$ Demand for electricity in week t, hour h, and

D Total demand *for* electricity during the year.

In the case of stationary yearly demand D, the hourly demand is computed on the basis of the demand D by applying weekly and hourly indices:

$$d(t,h) = \eta_{2t}(h)\eta_1(t)D, \ \ t = 1,..,T \ \ \ and \ h = 1,..,H \tag{8.1}$$

and the index series are defined as follows:

$\{\eta_1(t), t = 1, ... , T\}$ A set of weekly indices giving the distribution of the yearly demand for the weeks t, satisfying $\Sigma_{t=1}^{T} \eta_1(t) = 1$.

$\{\eta_{2t}(h), h = 1, \dots ,H\}$ For each week t a set of hourly indices giving the distribution of the demand in week t to the hours h of the week and satisfying $\Sigma_{h=1}^{H} \eta_{2t}(t) = 1$.

In principle, every week t may have its own series of hourly indices. In practice, the same series applies for several weeks for the different seasons of the year.

Now, we redefine the yearly demand D in (8.1) by substituting for D a variable demand $D(t)$, defined as follows:

$\{D(t), t = 1, \dots ,T\}$ The level of yearly demand reached in week t. This is an operative definition and means that the development of the demand in week t is computed from (8.1) after substituting $D(t)$ for D in the formula.

Then, we assume that $\{D(t), t = 1, \dots ,T\}$ is a stochastic process (8.2):

$$D(t) = D_0\left[1 + \sum_{\tau=1}^{t} \delta(\tau)\right] \tag{8.2}$$

where D_0 is the (level of) yearly demand at $t = 0$, and for every t, $\delta(t)$ is a random variable: the relative change in demand at the beginning of week t. We usually choose the random variables $\delta(t)$ to be statistically independent and differ from 0 for some weeks t of the year only. Inserting (8.2) into (8.1) gives the model for the demand (8.3) applied in the market model.

$$d(t,h) = \eta_{2t}(h)\eta_1(t)D_0\left[1 + \sum_{\tau=1}^{t} \delta(\tau)\right], \quad t = 1,..,T \quad and \ h = 1,..,H \tag{8.3}$$

In the dynamic programming formulation, the relative level of demand $D(t)/D_0 = x_2(t)$ is a state variable, which, corresponding to (8.2), satisfies the dynamic equation (8.4) with the initial value (8.5).

$$x_2(t) = x_2(t-1) + \delta(t), \quad t = 1,..,T \tag{8.4}$$

$$x_2(0) = 1 \tag{8.5}$$

In dynamic programming applications, the relative additive terms $\delta(t)$ are assumed to be statistically independent and to have discrete distributions as follows:

$\delta(t) = 0$, for $t \epsilon Q_0$ (8.6)

$\delta(t) = \delta_i$, with probability π_i , $i = 1, \dots ,I$, for $t \epsilon Q_1$ (8.7)

In (8.6)-(8.7), the sets Q_0 and Q_1 form a partition of the index set $\{1, ... , T\}$, and the probabilities π_i satisfy $\Sigma^I_{i=1} \pi_i = 1$. The nonzero terms (8.7) are usually assumed to be identically distributed.

8.3 Estimation of the parameters of the demand model

The estimation of the weekly and hourly indices $\{\eta_1(t), t = 1, ... , T\}$ and $\{\eta_{2t}(h), h = 1, ... , H\}$ for $t = 1, ... , T$, and the initial level of the yearly demand D_0 are important tasks in validating the model. We assume that the required data are available.

In applying the model (8.2) to the development of the level of demand during the study period $[0, T]$, the starting point and a basic assumption are frequently an estimate of expected growth and standard deviation of growth during the whole study period. The relative growth during the period is $\Delta = \Sigma^T_{\tau=1} \delta(\tau)$ and it is assumed that Δ has mean M and standard deviation Σ. There is considerable freedom in the choice of the distributions of the additive nonzero terms $\delta(\tau)$. Let n be the number of such nonzero terms, statistically independent and all equally distributed with mean value m and standard deviation σ. Then these parameters always satisfy:

$$M = nm \tag{8.8}$$

$$\Sigma^2 = n\sigma^2 \tag{8.9}$$

Suppose that we have estimated M and Σ. Then, the number n of nonzero terms and their distribution, the parameters m and σ are chosen so that (8.8)-(8.9) are satisfied. If nothing specific is known about the distributions of the additive terms, then it is natural to assume they are normally distributed. In the dynamic programming model, a discrete approximation (8.2) of the normal distribution is applied. This approximation is chosen so that it also satisfies (8.8)-(8.9) and is as near as possible to the normal distribution in a chosen sense.

8.4 Modelling price elasticity of the demand

For each hour h of the week, the demand for electricity $d(t,h)$ is a given constant in the weekly optimisation sub-problem. The demand for each hour has to be satisfied with production during the actual hour so that weekly production costs are minimised. Modelling the price elasticity of the demand is technically straightforward. For a constant demand d, we substitute a price-sensitive demand and instead of minimising the weekly production costs, the sum of the producers' surplus plus the consumers' surplus is maximised. This is a standard modification of a cost minimisation model, and we do not go into detail here. The extended problem is also convex, and it is approximated with a piecewise linear problem as discussed in Section 4.2.2. Technically, the modification is easy, but it may be difficult

to obtain reliable quantitative data on the elasticity of the demand, as the elasticity is highly time dependent.

8.5   Export and import of electricity

The basic model covers production and demand of electricity within the Nord Pool market area. The Nordic transmission net is connected to the Russian, Baltic, Polish, German and Dutch nets with transmission lines for export and import of electric energy. Including trade in electric energy with surrounding areas into the model is technically easy. Imports can be included as an additional source of electricity directly into the cost minimisation model. Exports are modelled in complete analogy with a price-sensitive demand.

   As with price elasticity, the real problem in modelling the trade of electricity with the surrounding nets and markets is data. Trade based on long-term contracts with clear agreements presents no problems. Short-term trade depends on market conditions, i.e. on the supply and demand at different price levels in the areas neighbouring the Nord Pool area and vary from hour to hour. In the model, we have to be satisfied with expected average values.

9. Piecewise linear approximations to convex cost functions of a production system

9.1 Piecewise linear convex cost functions and computational efficiency in optimisation

The expected minimal production cost functions $G(y)$ for the thermal system are convex functions of production y, and the marginal costs $F(y)$ are increasing functions. These functions were computed on the basis of plant data in Section 6.4 and on the basis of production class data in Section 6.5. Total cost functions computed on the basis of plant data are piecewise linear and convex, and the corresponding marginal costs are non-decreasing step functions.

All the cost functions in Chapter 6 are computed recursively in numerical form. The cost functions enter the objective function of the optimisation subtask (4.1)-(4.11) in the dynamic programming iteration. In our model, the subtask is an optimisation problem with linear constraints and a convex objective function. During the iteration, this subtask is solved over and over again very many times. Minimising the solution time is essential.

One possibility is to develop and apply an optimisation algorithm for linearly constrained convex minimisation problems with numerically defined objective functions. Piecewise linear approximations of convex cost functions is another possibility. The marginal cost function in that case is an increasing step function. A piecewise linear convex approximation leads to a linear programming problem, and the effective algorithms and computer programs for LP problems can then be used.

If we approximate an increasing marginal cost function with a step function, and with tolerance δ, then the cost steps of the approximating function should be chosen to be at most the height $2\,\delta$. In order to cover a cost range Δ with this accuracy, $\Delta/2\delta$ steps are sufficient.

In the electricity market model, we apply this approximation. The exact total cost functions $G(y)$ are approximated with piecewise linear, convex functions with a smaller number of linear segments and correspondingly fewer steps in the marginal costs.

9.2 Piecewise linear approximations to convex cost functions: an example

For clarity, we first examine the simple cost model $G(y)$, see (6.14) in Section 6.5.2, based on the probability distribution $\Phi(x)$ of the available capacity, and write a piecewise linear approximation of the total optimal production costs $G(y)$ obtained by approximating the corresponding marginal cost function (6.15), $F(y) = c + (h - c)\Phi(y)$ by a step function. Let us examine Figure 9.1, which gives an example of the function $F(y)$, and its approximation with a three-step function $F_a(y)$.
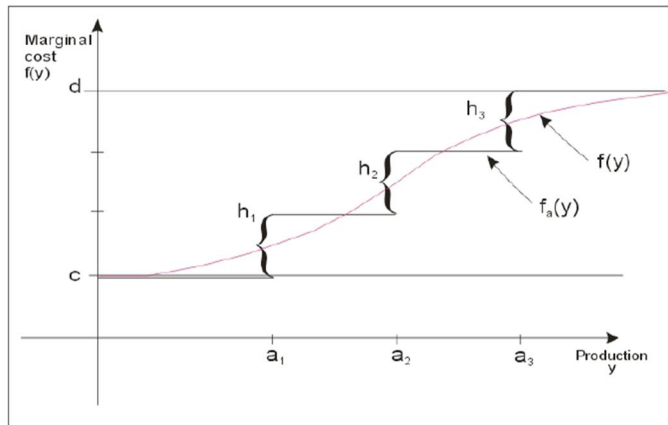

Figure 9.1. Marginal cost F(y) and its step function approximation Fa(y).

The approximation is determined by the heights h_i, and the locations a_i of the three steps i. These are to be chosen so as to minimise some measure of the closeness of the approximation to the true cost function. Let us choose $F_a(0) = F(0) = c$, $F_a(y) = F(y) = h$ for $y \rightarrow \infty$, and let us minimise the maximum absolute value of the difference, $max\{|F_a(y) - F(y)|: y \in (0, \infty)\}$ under these constraints.

It is easily seen that for every continuous function F, the best approximation has three equal steps of height $(1/3)(h - c)$ located at points a_i determined from the equalities $\Phi(a_1) = 1/6$, $\Phi(a_2) = 1/2$ and $\Phi(a_3) = 5/6$. The maximum absolute difference between the value of the exact marginal cost function $F(y)$ and its approximation $F_a(y)$ is $(1/6)(h - c)$. In a cost minimisation model, the approximation leads to the following representation of the approximate cost function $G_a(y)$:

$$G_a(y) = \min \sum_{i=1}^{4} c_i y_i \qquad (9.1)$$

$$y \leq \sum_{i=1}^{4} y_i \qquad\qquad (9.2)$$

$$0 \leq y_i \leq b_i \,, \, i = 1, \dots , 4 \qquad\qquad (9.3)$$

The cost coefficients are

$$c_i = c + [(i - 1)/3](h - c), \, i = 1, \dots , 4 \qquad\qquad (9.4)$$

and the upper limits $b_1 = a_1$, $b_2 = a_2 - a_1$, $b_3 = a_3 - a_2$ and $b_4 = \infty$, where the points a_i were solved above. It is also easy to see that if we have a cost function with marginal costs starting from cost level c and rising to h, then we can approximate the marginal cost with a step function with r steps, chosen suitably, to give a maximum difference *(1/2r)(h - c)* between the true marginal cost and the approximation.

9.3 An algorithm for the generation of piecewise linear approximations to convex functions

Let us generalise the previous result. The production is denoted by *y*. Any convex (cost) function *G(y)* with derivative (marginal cost) *G'(y) = F(y)* can be approximated on the interval *(0, K)* with a piecewise linear function $G_a(y)$. The derivative $G'_a(y) = F_a$ is a step function. The approximation is generated with an algorithm with the following steps:

1. Choose a tolerance *δ* and require that the maximum absolute difference between the true marginal costs *F(y)* and marginal costs $F_a(y)$ according to the approximation is *δ*.

2. Compute   *G'(0) = F(0) = c*   and   *G'(K) = F(K) = h*.   Here, we choose $F_a(0) = c$ and $F_a(K) = h$ (other choices are possible).

3. Choose the number of steps *r* to be the smallest integer satisfying *r > (h − c)/2δ*.

4. Compute the cost coefficients

$$c_i = c + \big[(i-1)/r\big](h-c), \quad i = 1,...,r+1 \qquad\qquad (9.5)$$

5. Determine the division points a_i from the equalities

$$F(a_i) = c + (1/2r)(h-c)(2i-1), \quad i = 1,...,r+1 \qquad\qquad (9.6)$$

6. Compute the upper bounds

$$b_1 = a_1, \, b_i = a_i - a_{i-1}, \quad for \; i = 2,...,r, \quad and \quad b_{r+1} = K - a_r \quad (9.7)$$

7. The approximation is given by

$$G_a(y) = \min \sum_{i=1}^{r+1} c_i y_i \qquad\qquad (9.8)$$

$$y \leq \sum_{i=1}^{r+1} y_i \qquad\qquad (9.9)$$

$$0 \leq y_i \leq b_i, \quad i = 1,...,r+1 \qquad\qquad (9.10)$$

A step function approximation with maximum error 5 (units)/MWh to the exact marginal cost function of the system defined in Table 9.1 is given in Figure 9.2.

Table 9.1. A thermal production system.

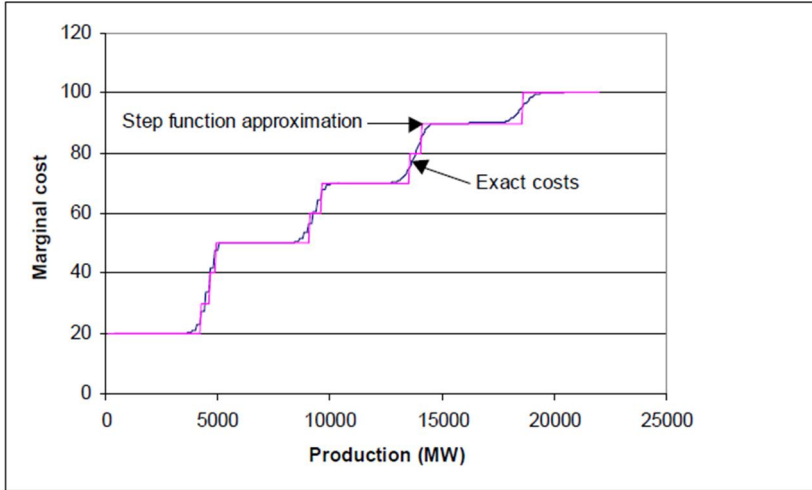| Plants				Production class		
Number of units in the class	Nominal capacity of a unit (MW)	Availability	Cost Coeffi-cient (€/MWh)	Expected capacity M (MW)	Std.dev. of availa-ble ca-pacity σ (MW)	σ/M %
25	200	0.9	20	4500	300	6.7
25	200	0.95	50	4750	218	4.6
100	50	0.9	70	4500	150	3.3
100	50	0.95	90	4750	109	2.3
1	∞	1.0	100	∞		

Figure 9.2. Expected marginal costs of the system in Table 9.1 and their step function approximation with max error 5 €/MWh.

Finally, let us note that the approximations used in applications need not be uniform with respect to the production. We can choose step approximations with great detail where this is required and even generate such approximations dynamically during the course of the iterations.

10. Price forecasts

10.1 Market price for electricity

Now, suppose that the overall problem has been solved, i.e. the complete optimal policy and the value functions have been computed, then, for every week *t,* and for every possible market condition, the price of electricity is computed as the derivative of the optimal expected costs from the beginning of week *t* to the end of time horizon *T,* with respect to the demand for electricity in week *t*. By definition, a perfect market realises this price.

Let the initial state at week *t,* i.e. the state reached at the end of week (*t* - *1*), be *y,* and let the realisation of the stochastic control element *S(t)* be *s(t)*. Then, we define *p[y,s(t),t]* as the expected average market price of electricity during week t, if the state of the process at the end of the week (t - 1), x(t - 1) = y, and the realisation of the stochastic control variable S(t) = s(t).

Here, the average denotes the weighted average over the hours of the week with the hourly demand as weight for each hour. The expectation is taken with respect to the distribution of the stochastic process $S(\tau)$ for $\tau > t$. The statistical expectation *E{p[y,s(t),t]}* of the time average market price in week *t* over the distribution of the stochastic element *s(t)* gives a prognosis for the market price in week *t* for initial state *y*.

The optimal costs from the beginning of week *t* to the end of the horizon for *x(t - 1) = y* and *S(t) = s(t)* are obtained as the optimal value $Z^*_t[y,s(t)]$ of the appropriate time-point-wise optimisation problem, (3.29)-(3.30) in general notation or (3.22)-(3.26) in more specific notation. Denote the demand for the week as a variable in this problem by *d* and let the optimal demand be *d**. Then

$$p[y,s(t),t] = \partial Z^*_t[y,s(t)]/\partial d, d = d^*$$ (10.1)

In the basic formulation (3.22)-(3.26) of the optimisation problem giving the price for electricity (10.1), the demand *d* for the week is given as a constant lower bound, denoted by $D_0\eta_{1t}[y_2 + s_2(t,i)]$ in the constraint (3.25), and the price (10.1) is the so-called shadow price of this bound. The problem (3.22)-(3.26) is a convex

optimisation problem with linear constraints. It follows from the duality theory of convex optimisation that the derivative (10.1) is equal to the component of the optimal dual solution that corresponds to the inequality (3.25).

If we examine the structure of the optimisation problem (3.22)-(3.26) in detail, it can also be seen that in this problem

$$p[y, s(t), t] = dc_t(k)/dk, \quad at\ k = u_2^*(t) \tag{10.2}$$

where $c_t(k)$ gives the thermal production costs in week t as a function of the thermal production k, and $u*_2(t)$ is the optimal thermal production in the problem.

In the actual detailed model, the hourly variations in the demand and production within the week are also represented. The realisations of the hourly spot prices are obtained as derivatives of the optimal costs with respect to the hourly demands of electricity.

10.2 Generating price forecasts with the model

Realisations of the stochastic price process are generated step by step as follows:

1. The computations start from the known, or assumed, initial state $x(0) = x^0$.

2. A realisation $s^r = \{s^r(1),...,s^r(t),...,s^r(T)\}$ of the stochastic control S is generated.

3. The price of electricity during the first week $p^r(1)$, corresponding to the realisation s^r, is computed from (10.1) by choosing $y = x^0$, $s(1) = s^r(1)$ and $t = 1$.

4. The state at the end of time step 1 is determined from the dynamic equation of the system by inserting the stochastic control variable $s(1) = s^r(1)$ into it and the optimal active control variable $u*(1)$ determined from the appropriate optimisation problem, problem (3.22)-(3.26) for the basic model.

5. The procedure is repeated for all time steps $t = 2,...,T$, and the realisation of the   price $p^r = [p^r(1),..., p^r(t),..., p^r(T)]$ is generated.

In creating price prognoses, a representative sample $\{s^r, r = 1,...,R\}$ of realisations of the stochastic control process is generated and a corresponding set $\{p^r, r = 1,...,R\}$ of price prognoses is computed. On the basis of the set, expected values and standard deviations etc. can be formed. We note again that in this simulation, we can use real, measured and scaled water inflow processes (realisations) in the sample. Such realisations show a proper correlation between the successive weekly inflows.

11. Computer system, basic data and the use of the model

11.1 Computer system

The methods presented in earlier chapters have been programmed as C language functions using text files for input and output of data. An APL language interface is used to manage the functions and input and output files. An Excel macro-based model is used as a help in creating the input files. These functions and interfaces form a set of computer programs called VTT-EMM (VTT Energy Market Model). With VTT-EMM, the user can set up model installations for different electricity markets. The market of greatest interest to us is the Nordic electricity market covering Finland, Sweden, Norway and Denmark. We have also modelled markets consisting of Finland only, Germany and the Baltic countries. VTT-EMM can flexibly be used to simulate different market scenarios according to the interests of the user.

With the user's interface, it is straightforward to set up and run different computational cases for a market. The organisation of data files is shown in Figure 11.1. There are a number of parameters to fine-tune the properties of the models and the market conditions and to select the time horizon for a model run. The input data on the market are presented separately from the program control and steering parameters.
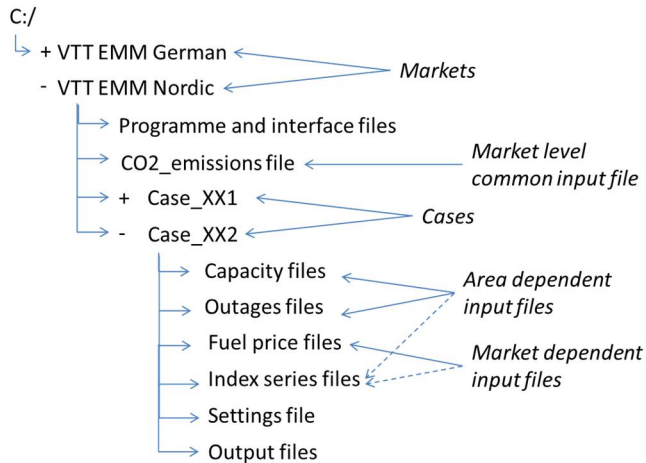
```
C:/
 └→ + VTT EMM German  ←
      - VTT EMM Nordic  ←──────────────  Markets
          ├→ Programme and interface files
          ├→ CO2_emissions file  ←──────  Market level
          ├→ + Case_XX1  ←                common input file
          └→ - Case_XX2  ←──────────────  Cases
                ├→ Capacity files  ←       Area dependent
                ├→ Outages files  ←        input files
                ├→ Fuel price files  ←      Market dependent
                ├→ Index series files ←---  input files
                ├→ Settings file
                └→ Output files
```

Figure 11.1. Schematic structure of VTT-EMM installations, cases and input files.

The running time of the model is determined by the solution time of the recurring optimisation sub-problem in the dynamic programming iteration. In the model, this problem is formulated as a linear programming problem and solved by the extremely fast MIPKIT library for LP problems developed at VTT. The approximations and accuracy of the results and the solution time can be controlled by the analyst.

11.2 Input data

11.2.1 General

The main part of the input data is organised in separate text files for all given areas (e.g. countries) of a market. The boundaries of a market are drawn by the user. For the Nordic market model, the areas are countries, as this corresponds to the availability of data.

Every production class, except hydropower, is included in the category of thermal power in the model and defined using the same input data format, e.g. wind power is thus also defined as a production class within thermal power. The input data for thermal power are organised in two data sets, one for the production classes and the other for the fuels. Data are stored in text files and, in practice, these data are semi-automatically derived from data stored in Excel files. The time resolution of input data is a year or a week and is chosen by the user according to the nature of the data. The weekly resolution is applied to the hydro inflow, planned outages, CHP heat demand and fuel prices.

11.2.2 Thermal power production classes

A thermal power production class in the model is identified by its market area and class name. A production class has a given total capacity and class-specific power plant characteristics. In the data set for thermal power production classes, see Figure 11.2, the input data are given for each area, production class and year.


```
FIN2020.CAP - Notepad

File  Edit  Format  View  Help

!FIN   2020
!.\FIN2020.CAP
!Tiedostosta C:\VTT EMM Nordic\case_xx\Kapasiteetit.xls 31.12.2011
10:01:46

!NIMI TY POL TEHO  LKM  p       mmk     Hyötysuhteet

YDIN NUC NUC 2601   4   1,00    0,0     1,00
GAST GTU LOI 805   11   0,95    0,0     0,20
LHII CON COA 2019  14   0,95    3,6     0,35
LTUR CON PEA 275    3   0,92   -1,0     0,35
LMKA CON GAS 12     1   0,95    0,5     0,50
KLHI CHP COA 598    8   0,85    3,6     0,88 0,29 0,29
KLÖL CHP HOI 16     1   0,90    0,0     0,90 0,30 0,30
KLMK CHP GAS 1141   9   0,95    0,5     0,91 0,47 0,50
KLTU CHP PEA 576   14   0,80    4,0     0,85 0,25 0,25
KLJä CHP JÄT 154    5   0,70    4,0     0,70 0,20 0,20
KLBI CHP BI1 838   28   0,85    4,0     0,85 0,25 0,25
THII CIP COA 127    8   0,50    3,6     0,88 0,22 0,22
TMKA CIP GAS 533    9   0,85    0,0     0,90 0,30 0,35
TTUR CIP PEA 216    9   0,50    4,0     0,88 0,20 0,20
TBIO CIP BI1 969   32   0,65    1,0     0,88 0,20 0,20
TSOO CIP LIQ 1072  24   0,60    0,0     0,85 0,20 0,20
Tmuu CIP JÄT 131    9   0,70    2,0     0,70 0,15 0,15
WIND WIN WIN 2047  20   0,30    0,0     1,00

VESI HYD 3224 5,6 14,2
RUS   EXP RUIM 1560 RUI RUEX 300 RUE
EST   EXP ESIM 1000 BAI ESEX 1000 BAE
KULU DEM 93
```

Figure 11.2. Input file for the thermal production classes of an area, containing the capacity, number of plants, availability, variable operation and maintenance costs, and efficiencies for each class.

The identification data are the area and the name of the class. The capacity data give the total net capacity of the class at the beginning of the year and the average or typical size of the plants in the class or number of plants. Note that the model treats the class as being formed of identical plants with the defined average or typical size.

The power plant characteristics are production type, fuel, availability, variable cost other than fuel and overall efficiency (fuel rate) (for CHP plants the efficiency in CHP mode and in condensing power mode separately). Production types include condensing power (CON), district heating CHP (CHP), industrial CHP (CIP), nuclear power (NUC), gas turbines (GTU) and wind power (WIN). The production types (CHP and CIP) are given special treatment in the model. The output genera-

tor of the model computes the distribution of production by the defined production types. Fuels available in the model are given by the user and can also be very specific. Data for the fuels are defined in the fuel data set.

Availability is measured during planned use, in other words it equals the complement of the forced outage rate. Other variable costs include the costs for operation and maintenance and all running costs proportional to production other than fuel costs. Other variable costs can also include a profit margin. The total efficiency is the share of net power (and net heat, for CHP plants) output of fuel input. For the CHP and CIP plant types: electricity efficiency in combined production mode and in condensing production mode are defined separately.

11.2.3 Special treatment of the combined heat and power classes

The CHP production classes (production types CHP and CIP) are mainly run in the economic combined heat and power mode. The available electric power capacity of these classes in combined production is determined by the heat load, which varies over the year. In the model, we define a weekly utilisation factor (≤1.0), and the nominal weekly combined power capacity is given by the utilisation factor times the total nominal capacity of the class. The weekly utilisation factors are determined by the variations in the expected district heating and industrial heat loads. The yearly variations for these two are markedly different. The remaining capacity of the CHP and CIP type production classes (the capacity that cannot run in combined mode) is available for condensing power production with a fuel consumption rate higher than in combined mode production. Forced outages are treated in the same way as in other thermal classes.

11.2.4 Availability of thermal power capacity

The total nominal net capacity for each thermal class is defined in the thermal power data set. For each week of the year, only a share of this capacity is manned and planned as available for production. The rest is scheduled to be closed down for planned maintenance, lack of demand or other reasons. The planned share (≤1,0) of the net capacity of a class, scheduled to be available during each week of the year, is given as input to the model.

The random availability (≤1,0) of the class gives the actually available share of the planned available capacity after random outages of plants.

11.2.5 Fuel data

Each thermal power production class (even wind formally) has an associated fuel. The data for fuels are defined in the fuel data set. There are no restrictions on the definition of fuels and, in principle, each production type could have its own fuel. In practice, the fuels are more generic, e.g. coal or peat, but these fuels may be

diversified by area, due to differing fuel prices. Gas prices in Norway, for example, are often lower than elsewhere.

The data set for fuels contains the annual average price for each fuel and the annual average price for CO_2 emissions. The annual prices defined are overridden by weekly fuel and emissions prices, if the analyst defines weekly fuel and emission prices.

Emissions of carbon dioxide per unit of fuel consumed are defined as a model level parameter for each fuel causing emissions. Carbon capture and storage-(CCS) based production classes are modelled using an appropriate fuel definition. For biomass-CCS, this leads to negative carbon dioxide emissions for the fuel.

11.2.6   Hydropower

The input data for the aggregated hydro system consist of capacity data and water flow data. Capacity data and total yearly inflow are given by area and year. The stochastic distribution of the flow of water during the year is defined by giving a representative sample (usually 100) of equally probable realisations of the weekly inflow process. The probability of each realisation is 1/N, if N is the number in the sample. These realisations are normed as weekly index series that give the distribution of inflow for the aggregated Nordic system as a whole. The analyst can model different inflow scenarios by choosing the sample of index series and the total average inflow accordingly. The physical inflow is measured in electric energy produced. The logical structure of the processing of hydropower data for the model is shown in Figure 11.3.
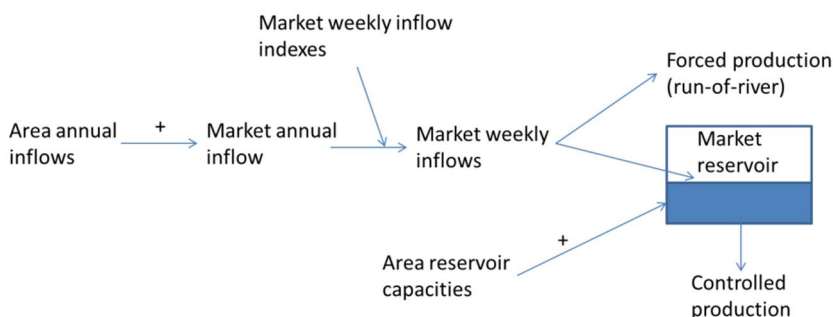

Figure 11.3. Logical structure of the preparation of hydropower data for the model.

For each area and year, the hydropower data contain the following items:

1.   Total hydro production capacity (in power units, MW),
2.   Total hydro reservoir capacity (in energy units, TWh), and
3.   Annual inflow (in energy units, TWh).

In addition, the user may optionally give the share of run-of-river capacity of the total hydro production capacity and the share of run-of-river inflow of the total water inflow. If the user does not define these shares, default shares defined in the model will be used. The difference between the total and run-of-the-river is storable and controllable. Excess water, either in the rivers or in the reservoirs, leads to an overflow past the turbines in the model, and the feasibility of the solution is always guaranteed.

The numerical parameters of the aggregated model of the Nordic hydro system represent the integrated behaviour and properties of a large number of individual and partially interconnected hydropower systems and components: power plants, water reservoirs and rivers. The water levels in the reservoirs and water flows in the rivers are regulated. The behaviour of the aggregated system is determined by its components, but the system is so large and complicated that the parameters are estimated on the bases of production statistics.

11.2.7   Foreign trade (trade in electricity with neighbouring areas)

Trade in electricity between the market and the neighbours is currently represented in the model in the simplest possible form. For every interconnection, transmission power capacity (MW) and energy price are given both for export and import of electricity. The capacities are constant over the year and the prices may vary weekly.

It would be interesting and, from the point of view of modelling, straightforward to represent price-elastic trade with the neighbours in the model, if the price curves of supply and demand of the neighbours at different hours and weeks of the year were available. As they are not, the model uses the, to a certain degree, simplistic representation of foreign trade. A user may control the trade by setting artificial bounds on the power transmission. This also corresponds to the increasingly common practice of using the transmission capacities for short-term system balancing of, e.g., wind and solar power, and not for trade per se.

11.2.8   Demand for electricity

The annual demand for each area (TWh) is given as input. In addition, each area has its own index series for the computation of the weekly and hourly demand. The total demand for each hour has to be covered by the trade-adjusted supply in the model.

Instead of dividing the week into 168 hours, a coarser subdivision of time within the week is usually used, corresponding to a step function approximation of the duration function of the demand chosen by the user. We have often used a very simple subdivision into three demand steps: peak, medium and low demand with chosen durations.

11.3 Control parameters

The time span of a studied case can be one or several years, starting from a selected week. In addition, the user may give the initial contents of the aggregated hydro reservoir. Otherwise, an iterative procedure is used for the estimation of the initial contents.

The user can also select the year to be normal, wet or dry. This parameter changes the total average annual water inflow values with a user-defined scaling factor, for example 0.8 for a dry year, 1.0 for a normal year and 1.2 for a wet year. Weekly inflow indices can also be modified by specific snow and precipitation correction parameters by the user.

The development of the annual demand over the study period is given as the forecasted expected growth during the year and the variation in the growth. The model generates a stochastic process with independent increments with a discrete approximation of the normal distribution, which realises the forecasted growth. Weekly demand indices can also be modified by temperature-dependent correction parameters by the user.

11.4 Output from the model

A complete study has two phases. First, the value function for the study year is computed under the given assumptions. Then, as the main result, a number (e.g. 100) of realisations of the marginal production price for each week and each hour (sub-period) is generated. On the basis of the sample, a number of results of the study can be computed and printed out, e.g. the value function, the development of the contents of the water reservoirs, production of hydro and thermal power, water overflow, imports and exports of electricity and total costs (as defined in the model). The value function is common for the whole sample, other results depend on the realisation and can be printed as such or summarised as a statistical distribution computed on the basis of the sample.

Further results can be derived and computed using a procedure that is a reversal of the computation of the expected total and incremental costs for the whole thermal system, that is, the reversed EIC method. For each simulated case, the distribution of the aggregated thermal production among the different areas and production classes can be computed on the basis of the simulated spot price and the detailed class data (capacity, cost and availability data). For the whole sample of realisations, the amount of such output data is very large and, usually, only aggregated results such as mean values are read out.

12. Results and applications

12.1 Model output

A complete solution to the central production optimisation problem of the model consists of the dynamic programming value functions for all weeks of the study period. During the computation of the value functions, the decisions forming the complete optimal policy and the associated spot prices of electricity are also available. Owing to the large amount of solution data, only the value functions are usually stored as output for an optimisation run.

The DP value functions also determine the illustrative water value function. The value of water as a function of the time of the year and the contents of the reservoir is shown in Figure 12.1 for the stationary solution of a problem with a planning horizon consisting of several identical years in succession.
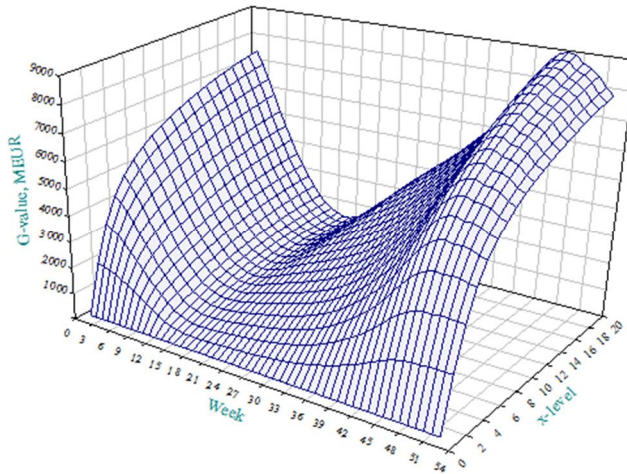
Figure 12.1. The water value function. G-value is the total value of water in the reservoir in million euros as a function of time (week) and content of the reservoir (x-level) from empty (x=0) to full (x=20).

Using the value functions, the model is used in simulation mode. The output of a simulation run gives an optimal production plan and the associated spot price, i.e. a realisation of the stochastic price process, for the whole period. The expectation and distribution of the spot price can be computed on the basis of a representative sample of simulated realisations of it.

   A simulation run gives the optimal production of the aggregated hydro and of the thermal plant of the model over the whole study period in the simulated case. The allocation of the optimal thermal production among the production classes that form the aggregated thermal plant is uniquely determined by the composition and properties of the thermal capacity. This distribution is computed by a separate program for every time step of the study period and forms an important part of the results of the model.

12.2 A price forecast generated with the model

A forecast for the price of electricity on the Nord Pool electricity market, computed with the model for the years 2007 to 2008, is presented in Figure 12.2. Both years were wet, although relative reservoir levels started to drop towards the end of 2008. The forecast (the thick blue line in the figure) is the statistical expectation of the system price computed on the basis of a representative sample of simulated

realisations of the self-same price. The solid red line shows the actual develop-ment of the Nord Pool Spot price and the dashed red line gives the Nord Pool forward quotations as they were on 10 January 2007.
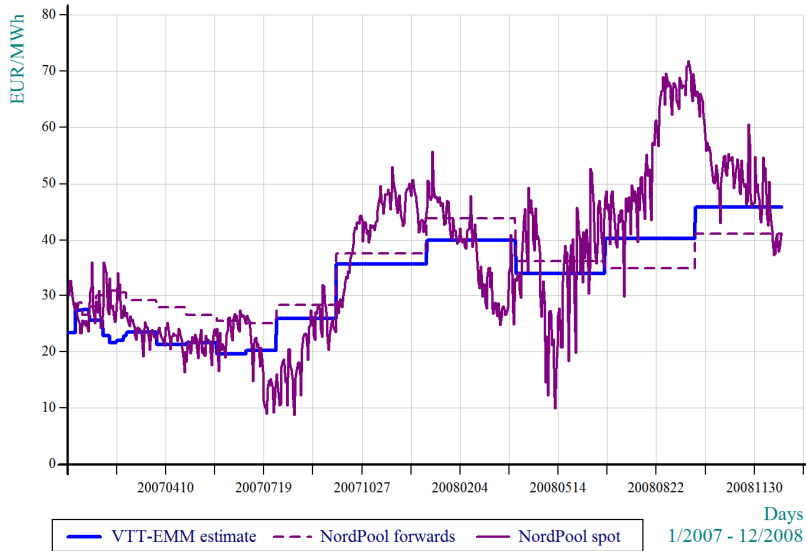

Figure 12.2. A price forecast computed with the VTT-EMM for the time span of 1/2007 to 12/2008 compared with the realised Nord Pool Spot prices and forward quotations. The model estimates are aggregated to match the time steps of the quotations for the electricity futures making it easier to compare them.

The forecast is computed for two years starting from the beginning of January 2007. The reported value is used for the contents of the aggregated water reser-voir in the beginning of 2007. The probability distribution of the inflow of water in the statistical sample of the model run is the same for both years, but the total annual inflows are set to match the realised values. Realised fuel prices, including the price of CO_2 emission rightsare used in the computation of the forecast. Fuel prices rose sharply during 2007 and in the beginning of 2008. The price of CO_2 emission rights for 2007 was very low; clearly less than 1 €/t_{CO2} most of the year. The price for 2008 emission rights was much higher, which could be seen in the forward prices already in January 2007, with an average of about 21 €/t_{CO2} during 2007 and 2008.

The forecasted price of electricity in the figure is the time average of the statisti-cal mean of the individual realisations in the sample. The time averages are taken over periods of one week, four weeks and three months, corresponding to the periods of the futures traded on the Nord Pool market. It can be seen that the

forecast worsens the farther it goes, but also that it is on par with the best market guess: the Nord Pool forward prices.

Figure 12.3 gives an idea of the spread of the simulated price distribution compared with one realisation, the real observed system price. Typically, the observed weekly average is a bit higher than the computed average. Sometimes the observed peak prices pierce the high end of the simulated distribution due to, e.g., extreme cold weather
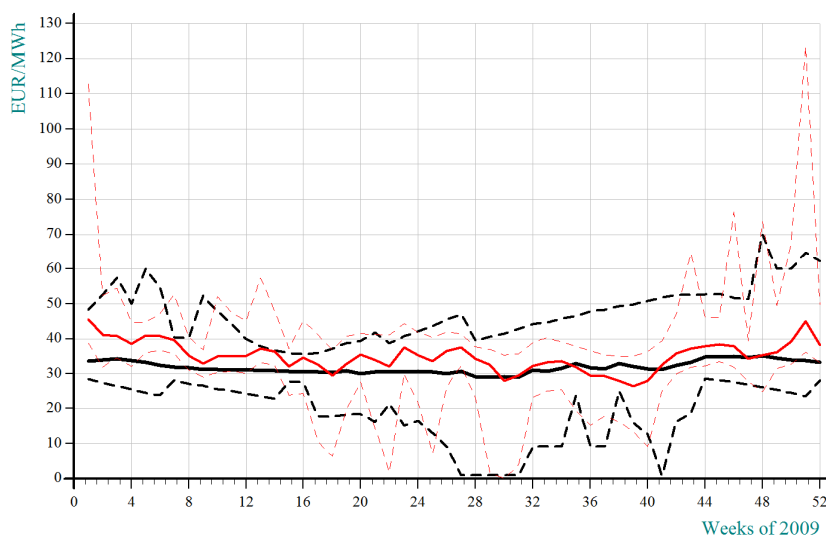

Figure 12.3. Modelled 2009 Nordic system price distribution (black lines) compared with the real observed system price (red lines). The bold black line is the weekly average of 100 simulations, and the thinner black dashed lines are the highest and lowest simulated one-hour values within a week. The middle red solid line is the weekly average of the observed system price and the finer dashed red lines are the highest and lowest hourly values for each week.

12.3 Applications of the model in market and policy analyses

VTT-EMM has been used in several market and policy studies and what-if analyses. As a fast model, it is well suited to scenario analyses and ad hoc assessments. Here is a brief non-inclusive list of analyses carried out with VTT-EMM:

- Comparing VTT-EMM results with other electricity market models' (Balmorel, ECON-Classic, PoMo) results to gain a better understanding of how the Nordic market works and for which market issues a specific model feature is a benefit and for which it is not. This analysis was part of

the Nordic Energy Perspectives (NEP) project (Unger et al. 2010, Rydén 2006, NEP 2010).

- Effects of alternative Finnish nuclear power construction plans on the Nordic electricity market. The study was conducted in 2009 and 2010 for the Energy Department of the Ministry of Employment and the Economy of Finland and carried out by VTT in close co-operation with the Finnish Government Institute for Economic Research (VATT) (Forsström et al. 2010). In this study the model was used to analyse the effects of a potential expansion of the Finnish nuclear production capacity on the Nordic electricity market. The market price and structure of the supply of electricity in the Nordic market area were computed for alternative Finnish nuclear power capacity expansion programmes for the years 2015, 2020, 2025, 2030 and 2040, and for the reference year 2007 for several alternative scenarios for the expansion of nuclear capacity in Finland.

- Assessing the change in the overall CO_2 emissions when converting selected power plants to carbon capture and storage (CCS). Conversion to CCS lowers the efficiencies and the power output of the units. This, together with the cost of emitting CO_2 and the cost of fuel, affects how the units are operated in a market environment and what kind of power plants are used to cover the power capacity reduction. (Rydén 2010)

- Assessing how fuel and/or CO_2 emission prices affect the operation of the power system and the system price. The impact of the emission trade has been an important research task since the planning of the EU emissions market started. (Koljonen and Savolainen 2004, Koljonen and Kekkonen 2005, Rydén 2010)

- Assessing the energy sector development and the electricity market development by linking together, iteratively, VTT-EMM and TIMES Finland (Kara et al. 2008, Loulou et al. 2005, Loulou and Labriet 2007). TIMES Finland is a long-term, multi-period partial equilibrium model that covers the whole energy production and consumption system of the national economy. It determines a solution that minimises total costs, including investments, over the study horizon. Finnish power capacity and demand for electricity are input data to VTT-EMM, which gives the market price. The price is fed back to TIMES Finland, which gives a new capacity and a new demand for VTT-EMM. This process converges after a few iterations. The results of the TIMES model can further be used as input data in estimating the effects on the Finnish national economy with a general equilibrium model for the Finnish economy, VATTAGE (Honkatukia 2009). For results computed with EMM and the Finnish Times and VATTAGE models, we refer to Forsström et al. (2010).

- Assessing the reaction of the market price to changes in demand and capacity (both power plants and cross-border transmission lines). The studies have looked at, for example, what happens if nine nuclear power plants in the Nordic system are shut down for one year, i.e. a loss of 66 TWh or 17% of the electricity production (Rydén 2010, NEP 2010); how

the market power of actors changes with the introduction of new nuclear power plants by different producers (Ruska and Koreneff 2009); the impact of a new transmission line between Finland and Russia on the Finnish and Nordic electricity markets (Kekkonen et al. 2006); the effects of changes in demand on the Nordic market price (Koreneff et al. 2009); what the EU target for renewable energy sources by 2020 might mean for the Nordic electricity market (Rydén 2006, Rydén 2010, Unger 2010); and how IEA fuel price scenarios reflect on Nordic power price estimates (Koljonen et al. 2012).

13. Discussions and conclusions

VTT-EMM estimates the power production and the market price of electricity using the production structure, electricity demand, hydro inflow and fuel, and EU EUA prices as inputs. We can, thus, simulate the market behaviour under different conditions, which makes the model very usable, especially for scenario work.

The level of detail versus aggregation is a central question in modelling. Accuracy has to be weighed against the effort of data acquisition, computation time and manageability of the model. VTT-EMM is a fast and easy-to-use power market model. VTT-EMM has only one hydro reservoir with a hydropower plant and one run-of-river hydropower plant, and other generation capacity is also aggregated. Are the results good enough? This question is quite complicated as it leads to the dilemma of how it should be measured.

First of all, a model of electricity production and the power market should reproduce, using realised inputs, the observed production mix by plant type and fuel within an acceptable tolerance. Experiences of the VTT-EMM model show that once calibrated, this requirement is satisfied even for new years on an annual basis.

Secondly, the estimated market price levels should be close to those observed. So, how close is good enough and is it a straightforward task to estimate the goodness? For example, the results of a one-year model run and the corresponding realised spot price are actually two different things. The analyses with the model in retrospect are based on only one (and simplified as such) representation of the decision processes and external factors determining the development of the price, whereas the realised spot price for one year is based on at least 365 different settings of external factors, as they change each day, and decision processes. Of the actors on the Nordic market, hydro producers with water reservoirs are the real price setters. Hydro producers make decisions based on their own estimates of the value of water in storage according to their views of the future at the decision moment. The expectations concerning the development in time of the demand, inflow of water and prices of fuels and emission rights change from day to day. As a result of this discrepancy, the model gives a relatively stable price estimate over the year, whereas the actual spot prices fluctuate more from week to week. Even if we ran the model every day for a year, would the results match the spot price? In principle, for the estimation of the hourly spot price, a model with an

hourly time resolution would be required. In the model, the decision process proceeds with weekly steps, and a week is divided into a small number of load segments. The examples in Chapter 12 were computed with a version of the model with only three intraweek time segments. Qualitative experience with this and other corresponding models points to five to seven time segments maybe being a suitable level of aggregation for the representation of intraweek power variations.

The best estimate of the quality of the model results is attained by comparing them with the electricity forward prices. The forward prices are based on the best knowledge of the market actors, which in turn use estimates from the best and most detailed market models available (for example, EMPS for the Nordic market) for reference. A statistically valid quantitative comparison between the model results and the forward prices would require weekly run estimates continuously for several years to represent all kinds of hydro years and forecast situations. It would still be difficult to say which are better: the model results or the forward prices, as they can only be compared with the realised spot prices, which in turn, as mentioned already, are based on a different set of external factors and on results from the models used.

Overall, the differences between the forward market prices and the VTT-EMM results are not big empirically, with the model results usually being more stable and the forward prices varying more. The main reason for this is the structural aggregation of both the decision process and the production capacities, especially the fact that all hydro reservoirs are combined into a single storage, the use of which is optimised as a whole. All thermal production is represented by a single expected incremental cost function. This greatly simplifies the model and its data requirements. Aggregated production class data are used instead of individual production plant data. The lost features include start-up costs and behaviour of the plants at partial load, internal fuel switching and the influence of the local weather. However, these factors are more important for the daily market behaviour than the long-term studies.

To sum up, structural simplifications (one reservoir, time aggregation) and data input generalisations (index series for power and CHP heat demand, modelled hydro inflow and capacity aggregation) in the VTT-EMM model both affect the results. The shorter the time span under study, the more it affects the accuracy. The data input generalisations can be replaced by accurate time series for each desired model run, but this would require a huge additional effort by the user. As it is, simplification and aggregation cut the computing time and data acquisition effort required and improve the flexibility of the model. The results are considered good enough empirically, especially on an annual level. Short-term spot price fluctuations are not well captured, especially not price spikes. However, with additional effort with the inputs, the short-term spot price forecasts could be better and more accurate.

Capacity development is exogenous in VTT-EMM, which is why it models the power system behaviour, not its development. As investment models are more complex and time dependent, the addition of capacity investment calculations would make the model more rigid and less agile, but it could still retain its energy-

only character. However, most investments in the Nordic countries are restricted: nuclear needs a permit, hydropower needs a usable hydro source such as a river, CHP needs a local heat load, and most renewables need feed-in tariffs or other subsidies, which make endogenous investment decisions mostly unusable and not worth the bother. Further model development could include methods to integrate the short-term intermittency of wind power into the model and modelling of intra-market price areas. There are several discussed solutions to stochastic wind power, but all solutions demand big changes to the programme. Intra-market price areas are created due to transmission bottlenecks in the system, which means that the simulation phase part of the solution would need to be rewritten.

Acknowledgements

References

Archibald, T.W., Buchanan, C.S., McKinnon, K.I.M., Thomas, L.C. 1999. Nested Benders decomposition and dynamic programming for reservoir optimisation. Journal of the Operational Research Society 50(5), pp. 468–479.

Amundsen, E., Bergman, L. 2006. Why has the Nordic electricity market worked so well? Utilities Policy 14 (2006), pp. 148-157.

Benth, F.E., Koekebakker, S. 2008. Stochastic modelling of financial electricity contracts. Energy Economics 30 (3), pp. 1116–1157.

Bunn, W.D., Karakatsani, N. 2003. Forecasting Electricity Prices. London Business School, EMG. Working Paper.

Deng, S.-J. Wenjiang, J. 2005. Levy process-driven mean-reverting electricity price model: the marginal distribution analysis. Decision Support Systems 40 (2005) 483–494.

Fleten, S.E., Lemming, J. 2003. Constructing forward price curves in electricity markets. Energy Economics 25, pp. 409–424.

Fleten, S.-E., Wallace, S.W. 1998. Power Scheduling with Forward Contracts. Proceedings of the Nordic MPS. Molde, Norway, May 9–10, 1998.

Forsström, J., Pursiheimo, E., Kekkonen, V., Honkatukia, J. 2010. Energy system, electricity market and economic studies on increasing nuclear power capacity. [In Finnish: Ydinvoimahankkeiden periaatepäätökseen liittyvät energia- ja kansantaloudelliset selvitykset]. Espoo, VTT. VTT Working Papers 141. ISBN 978-951-38-7482-7. http://www.vtt.fi/inf/pdf/workingpapers/2010/W141.pdf

Gjelsvik, A., Mo, B., Haugstad, A. 2010. Long- and Medium-term Operations Planning and Stochastic Modelling in Hydro-dominated Power Systems Based on Stochastic Dual Dynamic Programming. In: Rebennack S. et al. (eds.), Handbook of Power Systems I, Energy Systems, Springer-Verlag, Berlin, Heidelberg, Germany. Pp. 33–55.

Honkatukia, J. 2009. VATTAGE – A dynamic, applied general equilibrium model of the Finnish economy. VATT Tutkimuksia 150.

Kara, M., Syri, S., Lehtilä, A., Helynen, S., Kekkonen, V., Ruska, M., Forsström, J. 2008. The impacts of EU CO2 emissions trading on electricity markets and electricity consumers in Finland. Energy Economics. Vol. 30 (2008), No. 2, 193–211.

Kekkonen, V., Koreneff, G., Pursiheimo, E. 2006. Report on the effects of the Kotka-Kernovo sea cable on the Finnish and Nordic electricity markets [In Finnish: Selvitys Kernovo – Kotka merikaapelin vaikutuksista Suomen ja Pohjoismaiden sähkömarkkinoihin]. Ministry of trade and industry, Energy Department.
http://www.tem.fi/files/16175/VTT_Kernovo_Kotka_2006_04.pdf

Koljonen, T., Koreneff, G., Similä, L., Forsström, J., Ekholm, T., Lehtilä, A., Ruska, M., Pahkala, K., Hakala, K., Lötjönen, T., Niemeläinen, O., Rintamäki, H., Aro-Heinilä, E. 2012. Future of Finnish energy business – scenarios and strategies. Summary report of the SALKKU research project [In Finnish: Suomalainen tulevaisuuden energialiiketoiminta – skenaariot ja strategiat. SALKKU-hankkeen yhteenvetoraportti]. VTT Technology 25.
http://www.vtt.fi/inf/pdf/technology/2012/T25.pdf

Koljonen, T., Kekkonen, V. 2005. The Impact of Emission Trading on the Nordic Energy Systems. Proceedings Riso International Energy Conference. Technologies for Sustainable Energy Development in the Long Term. Riso National Laboratory, 23–25 May 2005. (http://www.risoe.dk/rispubl/SYS/ris-r-1517.htm)

Koljonen, T., Savolainen, I. 2004. Impact of the EU Emission Trading Directive on the Energy and Steel Industries in Finland. Greenhouse Gas Control Technologies. Proceedings of the 7th International Conference on Greenhouse Gas Control Technologies. 5-9 September 2004, Vancouver, Canada. ELSEVIER 2005. ISBN:0-080-44881-X.

Koreneff, G., Ruska, M., Kiviluoma, J., Shemeikka, J., Lemström, B., Alanen, R., Koljonen, T. 2009. Future development trends in electricity demand. VTT Research Notes 2470. Espoo 2009. ISBN 978-951-38-7273-1.
http://www.vtt.fi/inf/pdf/tiedotteet/2009/T2470.pdf

Loulou, R., Remme, U., Kanudia, A., Lehtilä, A., Goldstein, G. 2005. Documentation for the TIMES Model. Energy Technology Systems Analysis Programme (ETSAP).

Loulou, R., Labriet, M. 2007. ETSAP-TIAM: the TIMES integrated assessment model. Part I: Model structure. Computational Management Science special issue on Energy and Environment 5 (1–2), 7–40.

Lucia, J.J., Schwartz, E.S. 2002. Electricity prices and power derivatives: evidence from the Nordic power exchange. Review of Derivatives Research 5 (1), pp. 5–50.

NEP. 2010. Nordic energy perspectives project website http://nordicenergyperspectives.org/. Last visited 27.3.2013.

Nord Pool. 2013. Nord Pool Spot market online data. http://www.nordpoolspot.com. Last visited 27.3.2013.

Pereira, M.V.F., Pinto, I.M.V.G. 1991. Multi-stage stochastic optimization applied to energy planning. Mathematical Programming 52 (2), pp. 359–375.

Ruska, M., Koreneff, G. 2009. New nuclear power plants and the electricity market competition. [In Finnish: Ydinvoimalaitoshankkeiden vaikutukset kilpailuun sähkömarkkinoilla]. VTT Research Notes 2509. Espoo 2009. ISBN 978-951-38-7531-2. http://www.vtt.fi/inf/pdf/tiedotteet/2009/T2509.pdf

Rydén, B. (ed.) 2010. Towards a sustainable Nordic Energy system. 20 perspectives on Nordic energy, 10 opportunities and challenges. ISBN 978-91-978585-8-8.

Rydén, B. (ed.) 2006. 10 perspectives on Nordic energy. Stockholm, Elforsk. ISBN 91-631-9259-4. See: http://nordicenergyperspectives.org/boksep06.asp

SINTEF. 2014. The Norwegian Research Organisation SINTEF, the EMPS power planning model. Last visited 17.11.2014. http://www.sintef.no

Skantze, P., Gubina, A., Ilic, M. 2000a. Bid-based Stochastic Model for Electricity Prices: The Impact of Fundamental Drivers on Market Dynamics. Cambridge: Massachusetts Institute of Technology. Energy Laboratory Publication # MIT_EL 00-004. 61 p.

Skantze, P., Ilic, M., Chapman, J. 2000b. Stochastic Modeling of Electric Power Prices in a Multi-market Environment. Transactions of IEEE PES Winter Meeting, Singapore, January 2000.

Stedinger, J.R. 1998. Stochastic multi-reservoir hydroelectric scheduling, in: Proceedings, 28. Internationales Wasserbau-Symposium. Rheinisch-Westfälische Technische Hochschule Aachen, Aachen, Deutschland, pp. 89–117.

Tamminen, E., Kekkonen, V. 2001a. A dynamic programming model for forecasting the prices on an electricity market with stochastic demand and water inflow I, Theory. Research Report ENE6/38/01. VTT Technical Research Centre of Finland, Espoo, Finland, 2001.

Tamminen, E., Kekkonen, V. 2001b. A dynamic programming model for forecasting the prices on an electricity market with stochastic demand and water inflow II, Detailed structure of the optimisation sub-model. Research Report ENE6/43/01. VTT Technical Research Centre of Finland, Espoo, Finland 2001.

Unger, T. (ed.) 2010. Coordinated use of energy system models in energy and climate policy analysis – lessons learned from the Nordic energy perspectives project. Stockholm, Elforsk. ISBN 978-91-978585-9-5.: http://nordicenergyperspectives.org/modellrapport.pdf

Vehviläinen, I., Pyykkönen, T. 2005. Stochastic factor model for the electricity spot price – the case of the Nordic market. Energy Economics 27, pp. 351–367.

Weron, R. 2005. Heavy tails and electricity prices. The Deutche Bundesbank's 2005 Annual Fall Conference. Heavy tails and stable Paretian distributions in finance and macroeconomics. Eltville, 10–12 November 2005.

Yakowitz, S. 1982. Dynamic programming applications in water resources. Water Resources Research 18(4), 673–696.

Yeh, W.W.G.1985. Reservoir management and operation models, a state of the art review. Water Resources Research 21(12), pp. 1797–1818.

Appendix A: Computer model inputs and outputs

The mathematical model presented in this report has been programmed in the C language. An APL-language-based user interface is used to control the programme and manage input data and results, both in text files. This computer programme is called VTT-EMM (Electricity Market Model, MH-malli in Finnish). For different markets, we use different installations of the model, for example one for the Nordic market and another for Germany. Input data are defined separately for each market.

1 Input data

1.1 Thermal production classes[1]

For each class in each area[2] and for every year
- Class name (ascii4)
- Production type (ascii3)
- Fuel type (ascii3)
- Installed capacity (MW)
- Average unit size (MW)
- Availability (%)
- O&M-variable cost (€/MWh)
- Fuel efficiency ($MWh_e/MWh_{fuel,el}$) [3] for the production of electricity.
- Additional parameters for the calculation of conventional (condensing) production in cogeneration plants:
  - electric efficiency of the plant in cogeneration mode (MWh_e/MWh_{fuel})[4]
  - electric efficiency of the plant in condensing mode (MWh_e/MWh_{fuel})

---

[1] A thermal production class consists of identical thermal power plants. Thermal production classes include nuclear power and wind power in the model. Hydro is not included.

[2] An area (e.g. Finland) is a sub-area of a market (here the Nordic market). Areas are used in the management of input and output data. The model calculations are based on aggregated market data.

[3] Fuels are allocated to power and heat production by applying the energy method, see, e.g., Energy Statistics Manual by IEA/OECD/EUStats OECD/IEA, 2005, thus fuel efficiency is defined as electricity efficiency for conventional production and total efficiency for cogeneration production.

[4] The CHP power capacity is not the same as the power capacity of the plant in condensing mode. The electricity efficiency of the cogeneration mode is only used to calculate the fuel capacity of the plant, which is then used with the electricity efficiency in condensing mode to calculate the power capacity of the plant in condensing mode.

For each cogeneration production type[5] in each area
- Weekly capacity index (%)

For each wind production type[6] in the market (to be added in the future)
- Weekly capacity index (%) (outer index)
- Capacity index for time segments within each week (inner index)

For class, area, year and week
- Planned outages (unavailable capacity) (MW or %)

Per market
- Market power (positive or negative excess capacity) (±%)

1.2 Hydropower

Per market
- Inflow statistics (indexes for 100 years x 52 weeks each)
- Type of hydrological year (e.g. factor for wet, normal or dry conditions)
- Share of river inflow (%)
- Share of river production capacity (%)

For each area and year
- Annual inflow energy (TWh)
- Total installed hydro production capacity (MW)
- Reservoir capacity (TWh)

For year and week
- Deviation from normal snow conditions (±%), default 0%
- Deviation from normal precipitation conditions (±TWh), default 0 TWh

1.3 Fuel prices

Per market, for each fuel causing CO2 emissions
- Fuel-related emissions of CO2 (t_{CO2}/MWh_{fuel})

Per market for each fuel, year and week
- Price of fuel ($€/MWh_{fuel}$)
- Price of CO_2 allowance ($€/t_{CO2}$)

---

[5] Cogeneration production types are district heating CHP (CHP) and industrial CHP (CIP)
[6] WIN is a special production type reserved for wind power

1.4 Demand

Per market
- Deviation from normal outdoor temperature per week (± °C)
- Annual growth (±TWh)
- Range of stochastic weekly variation given as estimated deviation of annual growth (±TWh)
- Price elasticity given as price-elasticity steps

For area and year
- Annual demand including transmission and distribution losses (TWh)
- Weekly index (outer index)
- Index for time segments within each week (inner index)

1.5 Export and import of electricity

For each area, year, adjoining market and direction
- Capacity (MW)
- Price (€/MWh) (can be assigned to each year, week and time segment within each week)

1.6 Other parameters

Per market
- Other parameters for approximations, tolerances, accuracy and control

Per market and calculation[7]
- Starting week and year
- Number of years
- Parameters for calculation control

1.7 Verification and comparison data, not obligatory

Per market
- Historic spot prices (system price in Nord Pool)
- Historic forward prices

For area, year and week
- Historical level of reservoir [8](TWh)

---

[7] A calculation set is a single model run performed via the user interface.
[8] It is possible to fix the weekly hydro reservoir levels for a calculation. This is especially helpful in the calibration and assessment of the model using historical years as a comparison.
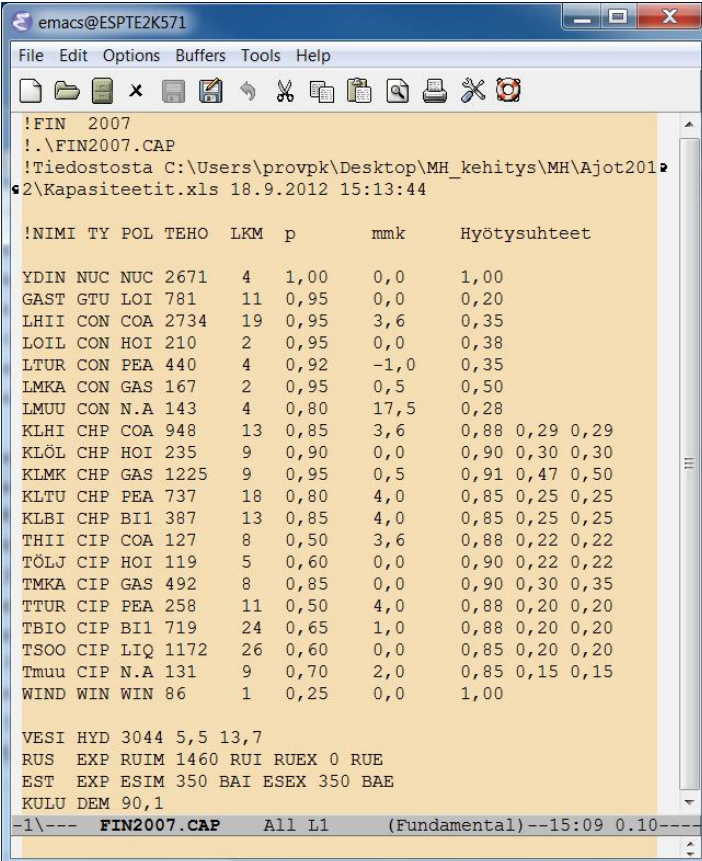
1.8 Input data management

Power generation capacity is defined for production classes consisting of similar (identical in the model) plants. Input text files, see Figure A1, are not easy to manage in a concentrated and synchronous fashion. Thus, an additional input data management file has been introduced. This file is based on Excel with multiple input data sheets and macros able to write the data to text files; see Figure A2.

Studies with the model usually comprise several cases. A case is an instance of the market and it differs from other cases, for example, in respect of capacities or fuel prices. Each case has its own input files including the Excel-based input management file.

Figure A1. An example of an input text file for a given area and year.

Figure A2. An example of production class capacity data in the Excel file, showing also two macro buttons, one for writing the data to annual text input files and the other to open these. The capacity volume data are in the xxxCap-sheets, while the capacity efficiencies, availabilities, operating costs etc. are given in the xxxPar-sheets, where xxx stands for the selected area.

The Cap Excel sheet shown in Figure A2 is also a handy place to store annual demand (TWh), total installed hydro production capacity (MW), annual hydro in-flow energy (TWh) and the total reservoir capacity (TWh) for an area. It also contains all the interconnection capacities between the area and countries outside the market; see Figure A3.

Figure A3. Input data based on area and year for hydropower and non-production class-related items such as demand, and import and export capacities in a Cap Excel sheet.

The FuelPri sheet in the Excel file contains annual prices for fuels, including CO_2 emission rights, and for imports/exports. The annual fuel prices are written into input text files using a macro. Annual fuel price input files are overridden by eventual weekly fuel price files. These files, in turn, are either managed by VTT-EMM's user interface or they may, for example, be automatically delivered from an outside source.

1.9 Input examples

Weekly capacity index of cogeneration

The CHP capacity utilisation depends on the week of the year and is different for district heating CHP and industrial CHP; see Figure A4.
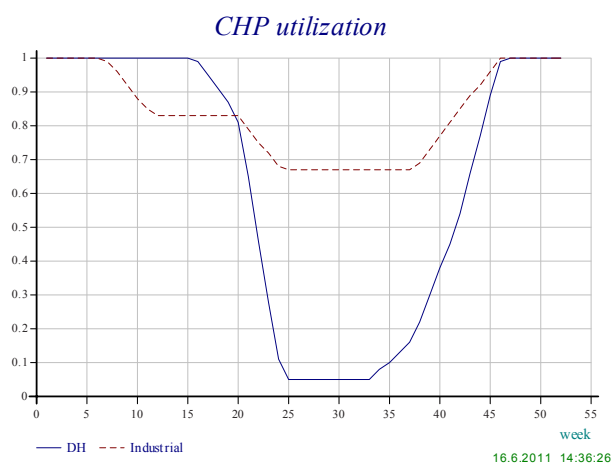

Figure A4. The share of CHP capacity that can be used each week at most in cogeneration production (weekly capacity indices for cogeneration). 'DH' stands for municipal district heating power plants and 'Industrial' for industrial CHP power plants.

The demand for heat is low, especially in the summer, and correspondingly only a relatively small part of the installed CHP capacity can be run in cogeneration mode. The rest of the installed CHP capacity can be used in condensing mode[9] with considerably lower efficiency. The total efficiency is typically 80%…90% for Scandinavian CHP plants in pure cogeneration mode[10], whereas electricity-only efficiencies are, at best, still below 50%...55%.

---

[9] For back-pressure CHP, this requires additional condensing ends or auxiliary coolers, which in reality is not always the case.
[10] Annual total efficiencies for many central Danish CHP power plants are below 70%, which means that they are more or less frequently run in condensing mode.

Hydro inflow indices

The stochastic water inflow is represented by a set of 100 index series, each for a different year. The average of the index series corresponds to 100% of the annual inflow, with some years above it and some below. However, when using the model, the user can select the hydrological year from three predefined types: normal, dry or wet. A wet year is defined as a year with a total inflow of 115% of a normal year and a dry year as having 85% of a normal inflow. A given percentage of the inflow can be stored in the reservoirs to wait for the optimal time to be used, while the rest has to be used immediately in so-called run-of-river generation. Abnormal snow and precipitation situations can be entered per week. Figure A5 shows the average, maximum and minimum inflow for each week of the year and two realisations of the 100 index series sample, as well as the maximum hydro production potential, taking into account the run-of-river potential's dependency on the inflow.
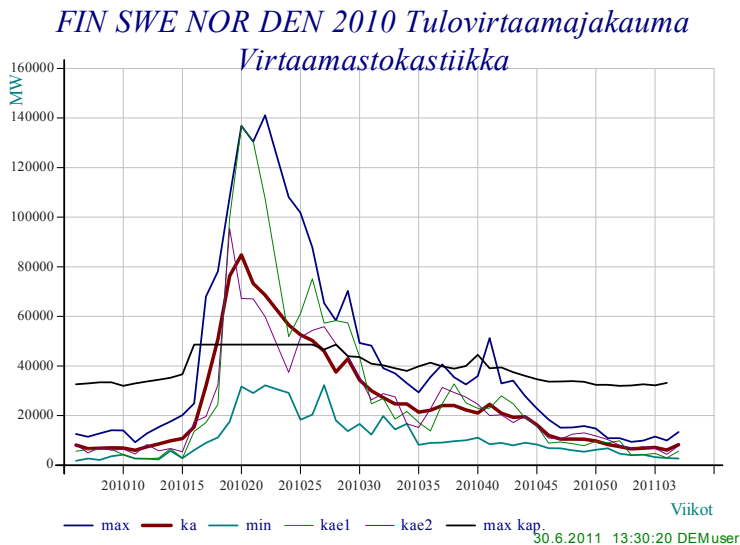

Figure A5. Total inflow scaled to 2010 Nordic level. 'max', 'ka' and 'min' are the weekly maximum, average and minimum inflows, respectively, in the sample. 'kae1' and 'kae2' are two realisations, and 'max kap' gives the maximal weekly generation potential.

Area demand for electricity

The annual demand is user given for each area. The variation in the demand over the year is given as a weekly index series. The variation in the demand within the

weeks is given by another set of indices for the time segments within the weeks; see Figure A6.
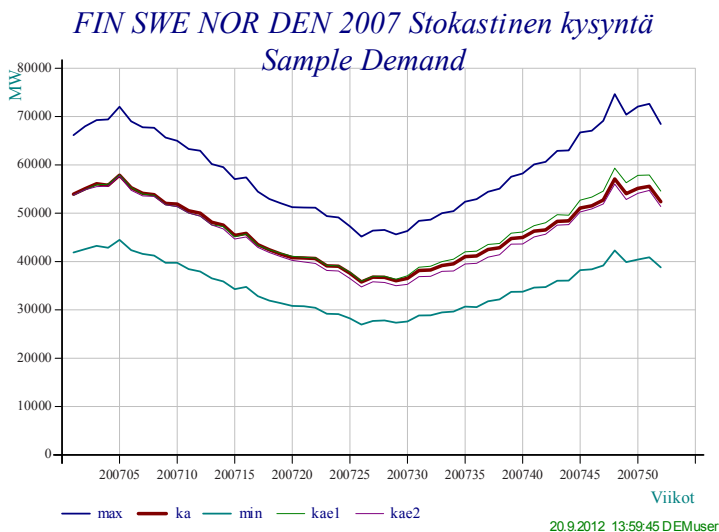

Figure A6. Stochastic demand. The weekly average ('ka'), maximum and minimum demand, and two random realisations 'kae1' and 'kae2' for the stochastic average demand in each week. The figure is based on a significant annual demand growth with a high deviation.

The dynamics of the weekly demand include a positive or negative growth trend. In addition, each week can experience a stochastic, normally distributed deviation. It is also possible for the user to give a weekly outdoor temperature deviation compared with the normal, which will affect winter heating and summer cooling loads.

With suitable time segment selection within the week, the calculation time for a year can be kept at 30 seconds while retaining a reasonably accurate overall result level.

2. Results

The output of a market model run includes the following data:
- Market price distributions
- Development of the hydro reservoirs
- Hydropower production, overflows and thermal production
- Exports and imports to and from the market
- Costs, and value of water in the reservoirs
- Ex post calculations

Logs and debugging information are also available for the user.

2.1 Market price distributions

The key result is the market price and its distribution. It is derived from the set of realisations of the price computed with the model. For every realisation of the stochastic input process, the corresponding realisation of the market price is computed. Figure A7 (left) shows the average of the weekly average as well as the maximum and minimum hourly prices occurring within the week in the whole sample. Figure A7 (right) shows the quartiles of the sample distribution of the highest hourly market price within a week for each week of the year. These and the following figures are from a model run for the Nordic market in 2010.
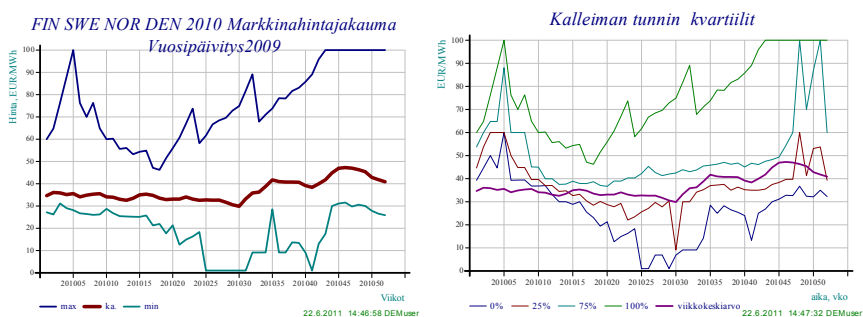

Figure A7. On the left: distribution of the weekly market price. 'ka' is the statistical mean of the prices within a week, and 'max' and 'min' are maximum and minimum hourly prices occurring in the sample during the week. (A model run for the Nordic market in 2010.) On the right: for each week of the year, the quartiles of the highest hourly market price. 'viikkokeskiarvo' = the average weekly price.

The price distributions can also be studied in more detail for a selected time frame. Figure A8 shows the statistical distribution of the lowest hourly market price occurring in a selected week: here week one in year 2010.

Figure A8. Probability distribution of the lowest hourly price of electricity occurring during a chosen week (first week of 2010), based on a sample of 100 simulated cases. The price of electricity is in euro/MWh on the horizontal axis, with the cumulative distribution on the vertical axis.

The distribution of the annual price is also part of the computed results. Figure A9 shows the histogram of the annual price.
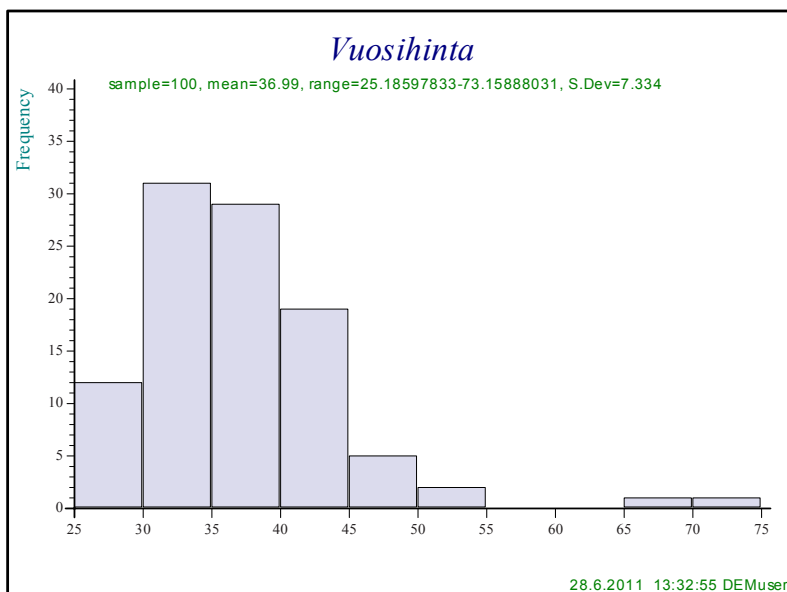
Figure A9. Histogram of the average annual price (€/MWh) of electricity based on a sample of 100 simulations

2.2 Development of the hydro reservoirs

The content of the aggregated hydro reservoir is the single most important variable of the model. Figure A10 illustrates the development of the water reserves in a set of simulations.

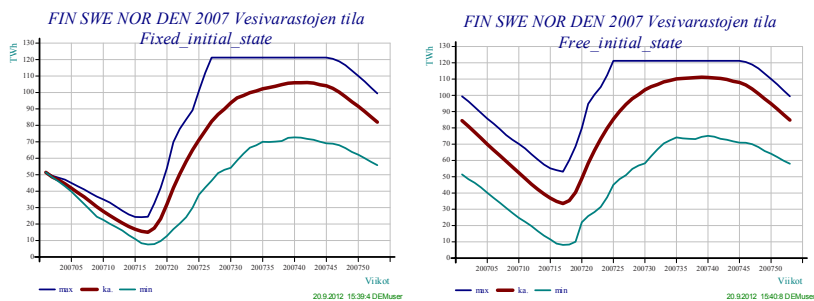Figure A10. Development of the aggregated hydro reservoir in a set of simulations. The bold red line ('ka') is the average storage, and the thin blue lines show the biggest and smallest contents of the reservoir occurring in the sample. In the left figure, the initial contents are user given and, thus, the same in all simulated cases. The right hand figure shows a case in which the initial state has been left

open, and its probability distribution is found by simulation. The left-hand case corresponds to the situation of estimating the near future starting from today, whereas the right-hand case corresponds to the simulation problem for a general year, e.g. in the future, when the starting point is not known.

2.3 Hydropower production, overflows and thermal production

Power production is aggregated into two production categories in the model: hydropower and thermal power. The latter category includes all production other than hydropower. Production of hydropower and thermal power are decision variables in the model; see Figure A11 and Figure A12.

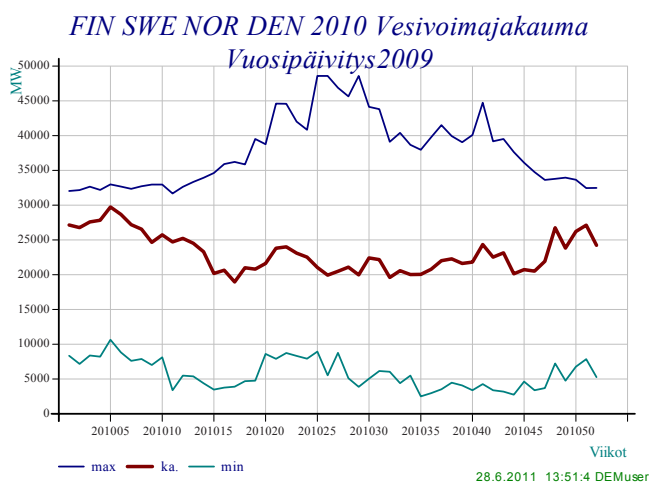Figure A11. Average hydropower production during each week of the year. The figure shows the results of a set of simulations. The bold red line is the statistical mean of the weekly (time) average power, and the thin blue lines give the minimum and maximum average weekly power levels in the sample.

Overflow, that is bypass of water past the turbines, affects the production potential of hydropower; see Figure A13.

FIN SWE NOR DEN 2010 Lämpövoimajakauma
Vuosipäivitys2009
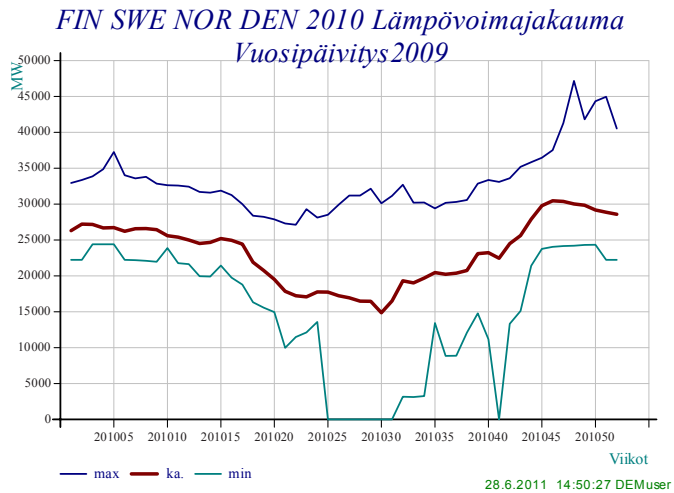
Figure A12. Weekly average thermal power production over a year. The bold red line is the statistical mean of the weekly averages and the thin blue lines show the maximum and minimum weekly production levels in a set of simulations.

FIN SWE NOR DEN 2010 Ohijuoksutus
Vuosipäivitys2009

Figure A13. Weekly overflows – bypass of water past the turbines – over a year, measured as the corresponding average weekly power. The black lines show the water bypassing run-of-river power plants and the red lines water bypassing power plants with reservoirs. The bold lines are the statistical means, and the thin lines are the maximum and minimum values, respectively, in a set of simulations.

2.4 Exports and imports to and from the market

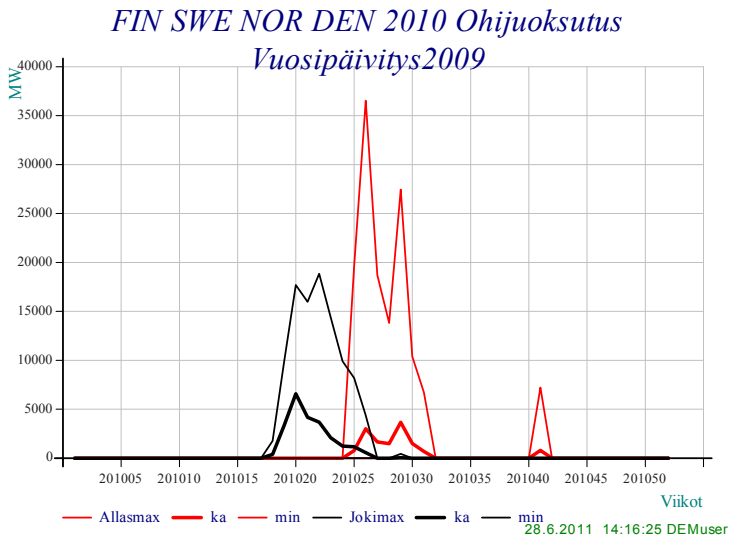The balance between supply and demand of power also includes exports and imports of electricity; see Figure A14. Detailed results of exports and imports between specific areas inside and outside the market are also available.
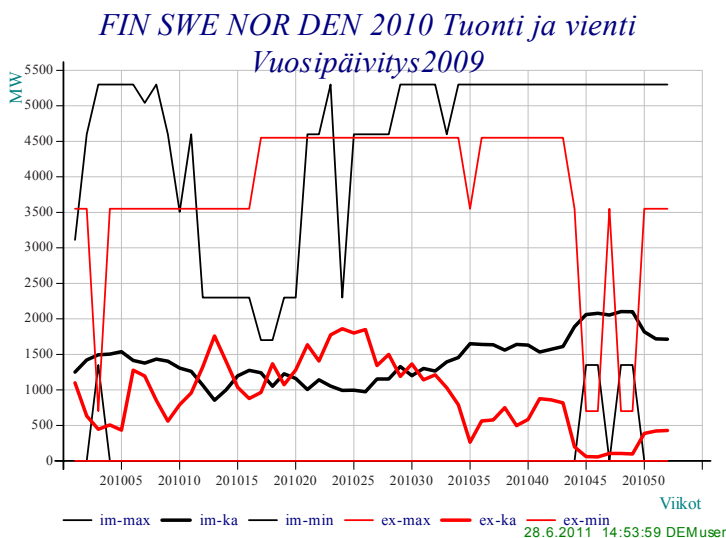
Figure A14. Imports and exports of electricity presented as average weekly power. The red lines represent exports and the black lines imports. The bold lines give the mean values and the thin lines show the maximum and minimum values in a sample.

2.5 Costs and the value of water in the reservoirs

The cost function includes thermal production costs, costs for imports and revenues from exports, and terms to represent the price elasticity of the demand. Figure A15 gives an example of the development of weekly costs.

FIN SWE NOR DEN 2010 Lämpövoiman viikkokustannus Vuosipäivitys2009
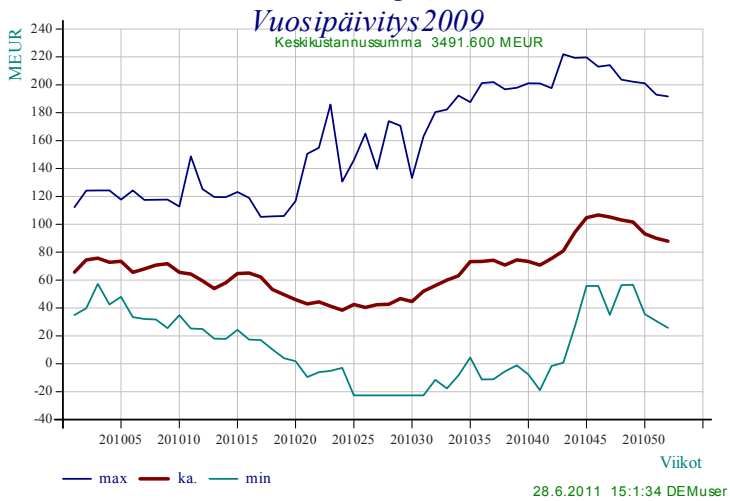
Figure A15. Weekly thermal production costs including import costs and export revenues. 'ka' is the statistical average, and 'max' and 'min' the highest and lowest value, respectively, in the set of simulations. At times of low demand and high hydro potential, export revenues can be bigger than the low thermal (mostly industrial CHP) production costs.

Figure A16 (left) shows the total cost function G as a function of time and hydro reservoir level. Note that the reservoir level on the y-axis starts at value 1 from a full reservoir and goes to an empty reservoir at value 20. The value of the water function V on the left is uniquely determined by the dynamic programming cost function.

FIN SWE NOR DEN 2007 G(x,t)

27.11.2012 16:6:57 DEMuser

Figure A16. The dynamic programming value function G on the left, and the value of water in the reservoir V (right). Time on the x-axis and contents of the reservoir on the y-axis. N.B.: The reservoir level scale on the left (1=full) is the reverse of the one on the right (1=empty).

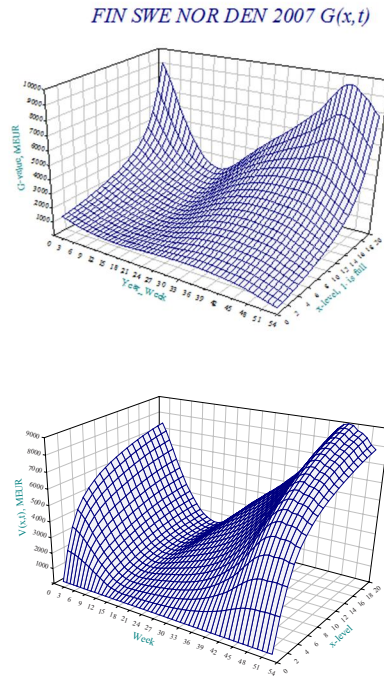The value function G in Figure A16 represents the cost that will cumulate for the production system from any selected point to the end of the horizon using an optimal route. The costs of imports and the revenues from exports and the value of water remaining in the reservoir at the end of the horizon are included in the value function.

2.6 Ex post calculations

In order to obtain more useful information, the model results can be processed further. As the model uses an aggregated thermal cost function, the results have to be reverse engineered to estimate the production by individual production classes. For any time segment, the expected production of the different thermal production classes is reconstructed by reversing the construction of the expected incremental (marginal) cost function. Figure A17 shows how the expected incremental cost function is formed for the thermal system defined in Table 9.1 and

Figure 9.2. The system has five production classes with costs of 20, 50, 75, 90 and 100 €/MWh, respectively, and infinite backup capacity at a price level of 120 €/MWh.

UTO 2000 Tuotantokapasiteetti (eivesivoima) viikolle 1

29.6.2011 14:40:24 DEMuser
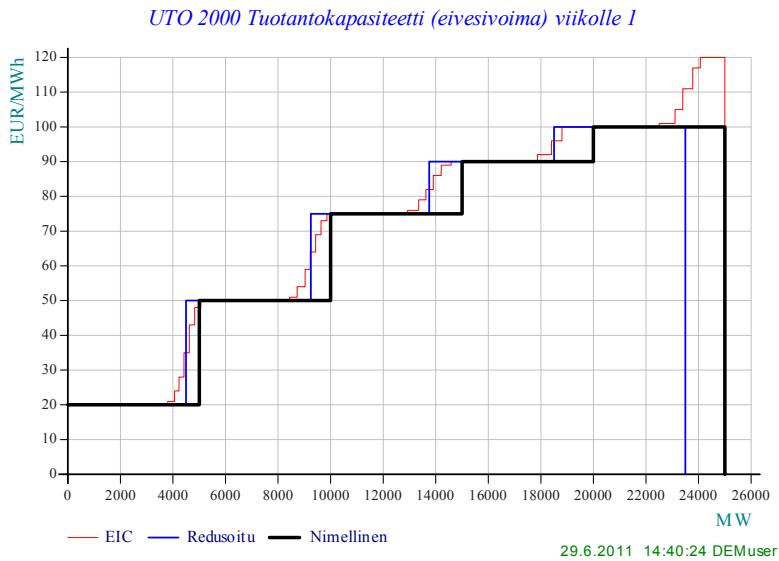
Figure A17. Expected incremental (marginal) cost (EIC) function of a thermal system. The black line gives the nominal capacity and corresponding marginal production costs of the system. The blue line gives the same data, but the total capacity of the classes is reduced to the expected available capacity. The red line is a step function approximation of the expected incremental cost for the system when the loading is done in merit order and the unit size of the plants is taken into account.

Figure A18 shows an example of the results for a production class.
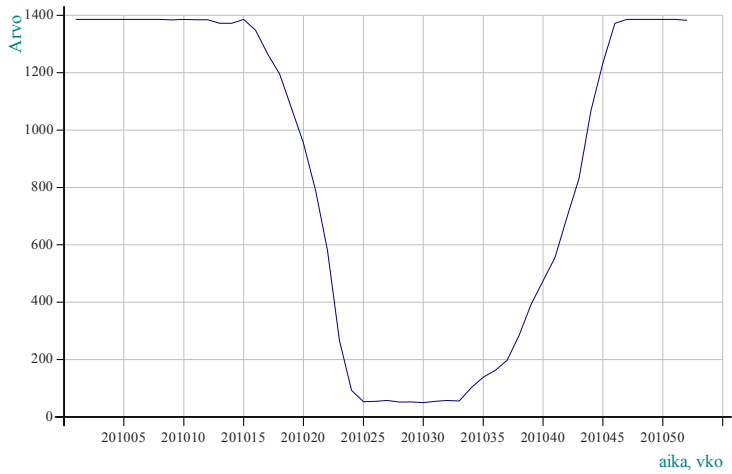
Figure A18. Expected average weekly power production (MW) in gas-fired combined district heat and power plants in Finland

The report generator of the model computes summaries of the results. Figure A19 shows a yearly summary report that gives the production of electricity by plant type and fuel for each area. It also reports the imports and exports of electricity, hydro reservoir levels, bypass of hydro energy, CO_2 emissions and, as a central result, the annual market price.

```
MH-malli VTT
29.6.2011  15:6:11

MH-vuosiraportti
C:\Program Files\Adato\Markkinahintamalli\
.\Adato2009\
Vuosipäivitys2009

Sähkön hankinta, TWh
 2010
      FIN  SWE  NOR  DEN  YHT
HYD    14   67  122    0  202
NUC    22   71    0    0   93
GTU     0    0    0    0    0
CON    10    1    0   14   25
CHP    18   10    0   13   41
CHC     1    0    0    2    3
CIP    14    6    5    2   27
CIC     1    0    1    0    2
WIN     0    3    2    7   13
XXX     0    0    0    0    0
IMP    11   -1   -2   -3    4
YHT    90  158  128   34  410
DEM    89  152  132   37  410
JOU     0    0    0    0    0
BAL     1    7   -5   -3    0

      FIN-RUS  FIN-EST  SWE-GER  SWE-POL  SWE-LIT  NOR-RUS  NOR-NLD  DEN-GER  YHT
IMP         9        2        0        0        0        0        0        0   12
EXP         0        0        1        0        0        0        3        4    8
YHT         9        2       -1        0        0        0       -3       -3    4

Sähkön tuotanto polttoaineittain TWh ja CO2 t:

      FIN  SWE  NOR  DEN  YHT
NUC    22   71    0    0   93
LOI     0    0    0    0    0
COA    12    2    0   19   33
HOI     1    1    0    1    2
PEA     7    0    0    0    7
GAS    11    4    0    8   22
N.A     1    0    0    0    2
JÄT     0    1    0    2    4
BI1     6    7    0    1   14
LIQ     6    3    0    0    9
WIN     0    3    2    7   13
MAS     0    1    0    0    1
NGA     0    0    5    0    5
XXX     0    0    0    0    0
HYD    14   67  122    0  202
YHT    79  159  130   38  406
CO2    16    3    0   17   36
REN    26   79  125    9  238

0 TWh tasepoikkeama (36.57594615  MWh)
2.29 TWh allasohijuoksutus
3.91 TWh jokiohijuoksutus
92.69 -0.17 TWh alkuvarasto ja lisäys
36.99 EUR/MWh keskihinta
```
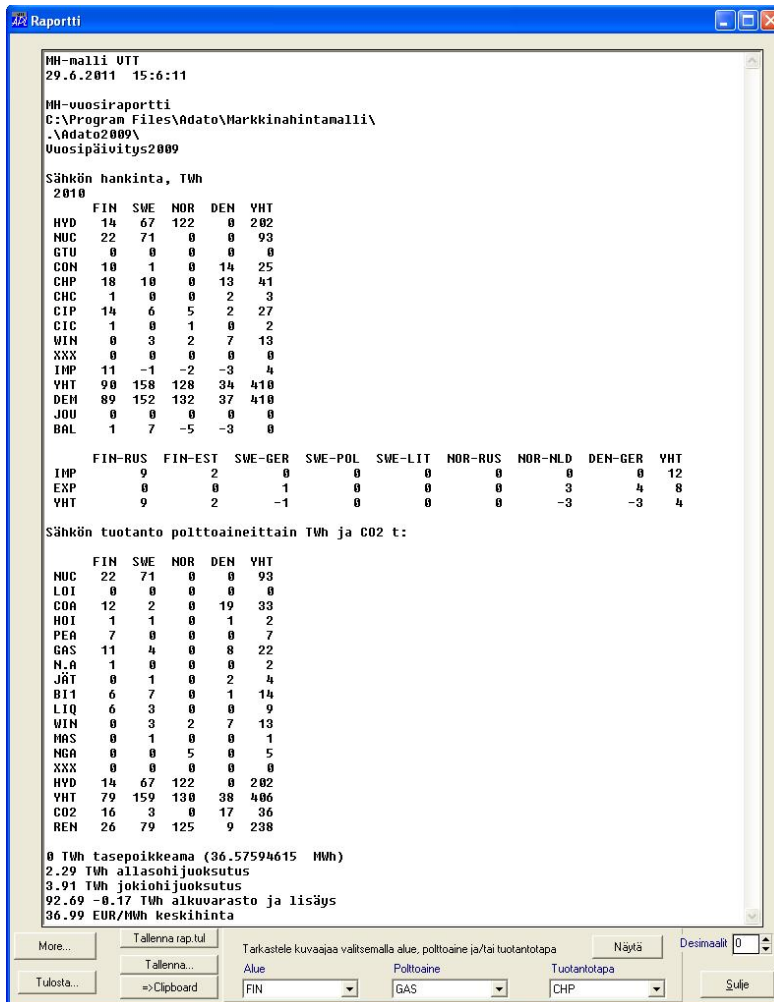
Figure A19. Summary report for a year: production of electricity by production class and by fuel, exports and imports by connections, change in hydro reservoir level, overflows and average market price.

Emissions of carbon dioxide are computed for each production class on the basis of the following data: fuel consumption rate in the production of electricity and emissions of CO_2 per unit of fuel consumed (for each fuel). The share of renewables in each fuel is also given so that the share of electricity production based on renewables can be calculated and reported.

In the following figure (Figure A20), the merit order of the production classes for a given week illustrates the overall balance between supply (blue line) and demand (red line) of electricity. The supply includes all production classes and import sources in order of rising costs. The demand includes the given demand in the

market and export possibilities. Price elasticity of the domestic (e.g. Nordic) demand is not included in the figure.

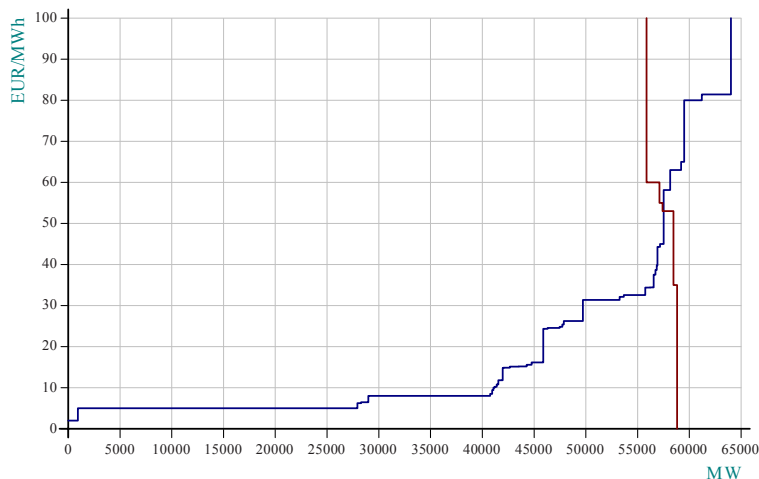*FIN SWE NOR DEN 2005 Kysyntä/Tarjonta viikolla 1*

Figure A20. Merit order of production classes for a given week. Power on the x-axis, marginal production cost on the y-axis. The red line represents the market demand and export possibilities and the blue line the market supply and import possibilities (weekly average values).

3. Execution and control of the computations

The user interface (Figure A21) is used for case management, management of control and input data, execution of computations with the model and to present the results. The APL interpreter can also be automated to perform recurrent calculations with varying user-defined input scenarios, thus enabling mass calculations of scenarios.

The basic computational case covers one year or, more precisely, 52 weeks. The starting point can be chosen freely. Several consecutive years can also be combined into one computational task. An analysis with the model has two steps. The first is the optimisation step, in which the value function G is solved by dynamic programming. In the second step, the simulation, the model is used to generate realisations of the stochastic output processes, such as production and price. These simulated realisations form the basic output.

Within the dynamic programming procedure, the sub-problem of optimal weekly load dispatching is solved with different parameters over and over again, very many times. In the model, this sub-problem is approximated with a linear pro-

gramming (LP) problem and solved with extremely efficient LP code. The structure of the coefficient matrix of this LP problem is shown in Figure A22.

The VTT-EMM model uses the MIPKIT library for the solution of the weekly LP sub-problems. The solver is extremely fast, especially in this kind of repetitive computation.
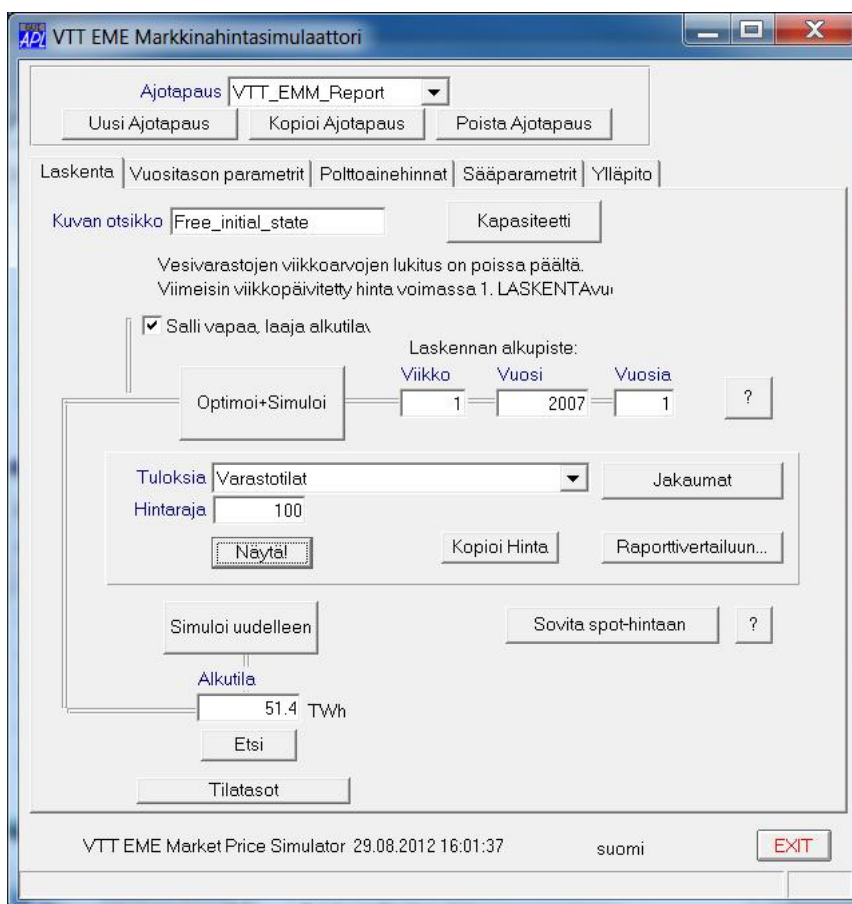

Figure A21. The main user interface of VTT-EMM (Finnish version) showing the Calculations tab ('Laskenta'). It is used for the management of different input cases for a market, for the execution of computational tasks and for looking at the results. There are separate tabs for year level parameters ('Vuositason parametrit'), fuel prices on a weekly scale ('Polttoainehinnat'), weather parameters ('Sääparametrit') and maintenance ('Ylläpito').

Figure A22. The linear programming structure of the weekly load-dispatching sub-problem.

Title	**A fast and flexible stochastic dynamic programming model of the electricity market** VTT-EMM – structure and use
Author(s)	Eero Tamminen, Veikko Kekkonen, Göran Koreneff & Tiina Koljonen
Abstract	The Nordic power system and electricity market consists of Finland, Sweden, Norway and Denmark, which have a common market place. Most of the electricity produced and consumed goes through the Nord Pool Spot exchange. Our main objective was to develop a fast, agile, robust and transparent, medium-to long-term mathematical model of the stochastic development of the spot price of electricity on the Nord Pool market. In this report we present the Electricity Market Model of VTT Technical Research Centre of Finland, VTT-EMM. The model is a mathematical representation, based on stochastic dynamic programming, of the dynamics of the demand, production and market price of electricity in the power system of the Nordic countries. The power system is versatile and dominated by hydropower, and modelling the market is a demanding task. There are structural simplifications (one hydro reservoir, time aggregation) as well as data input generalisations (index series for power and CHP heat demand, modelled hydro inflow, and capacity aggregation) in the VTT-EMM model, which affect the results. Simplification and aggregation cut the computing time and data acquisition effort required and improve the flexibility of the model. Can we obtain reasonable market price and production estimates using a greatly simplified representation of the hydro system and the other power plants? Test results are compared with both spot prices and forward prices on the market. The results are considered to be good enough empirically, especially on an annual level. One weakness is that short-term spot price fluctuations are not captured very well. Further model development could, thus, include methods to integrate the short-term intermittency of wind power into the model and modelling of intra-market price areas. One strength can be found in the assessment of a more long-term future using multiple scenarios, even up to several hundred.
ISBN, ISSN	
Date	December 2014
Language	English
Pages	84 p. + app. 22 p.
Name of the project	
Commissioned by	
Keywords	electricity market model, stochastic dynamic programming, Nordic electricity market, price forecast
Publisher	VTT Technical Research Centre of Finland P.O. Box 1000, FI-02044 VTT, Finland, Tel. 020 722 111

A fast and flexible stochastic dynamic programming model of the electricity market
VTT-EMM – structure and use

VTT-EMM is a fast, agile, robust and transparent electricity market model developed at VTT Technical Research Centre of Finland for medium to long term electricity price estimation, market analyses and energy policy studies.

The model is based on dynamic programming. It is a mathematical representation of the dynamics of the demand, production and market price of electricity in the Nordic power system, co-ordinated by the Nord Pool electricity market. All aspects of VTT-EMM are discussed: the mathematical approach, the overall structure, the detailed representation of the different sectors of the production system, approximations applied, generation of price forecasts and other results, the computer system, and applications of the model.