HELSINKI SCHOOL OF ECONOMICS (HSE)
Department of Accounting and Finance

OPTIMIZATION AS COMMODITY FORWARD CURVE CONSTRUCTION METHOD

Finance
Master´s thesis
Kari-Pekka Hakola
Fall 2005

Approved by the Council of the Department 15 / 12 2005 and awarded the grade good 70 points
KTT Vesa Puttonen KTT Matti Keloharju

Helsinki School of Economics
Master's thesis
Kari-Pekka Hakola

Abstract
November 24, 2005

Optimization as commodity forward curve construction method

Purpose of the study

Commodity forward curves have had much less attention on academic papers than interest yield curves. However, recently also commodity forward curves have begun to have more attention as interest on commodity derivatives has been increasing. The objective of this thesis is to test the suitability of optimization as commodity forward curve construction method in the case of oil and pulp. In more detail, we test if forward curves generated with optimization produce statistically significant pricing errors, i.e. do pricing errors differ statistically speaking from zero. In addition we study the dynamics of the generated commodity forward curves.

Data

The dataset contains par swap quotes for oil and pulp and all the data are denominated in USD. In addition, there are futures quotes of maturities ranging from 1 month to 12 months in the case of oil. Also interest rate data is used in the generation of forward curves and also in the valuation of swaps. We have used USD-Libor rates ranging from overnight to five year maturities to construct yield curves for valuation.

Methods

Pricing accuracy is tested with Wilcoxon matched – pairs signed – rank test. The accuracy is tested over number of maturities by testing the pricing difference between actual market par swap prices and the prices obtained from generated forward curves. The dynamics of the forward curves is analyzed with principal component analysis.

Results

The pricing accuracy of the generated forward curves is found to be statistically significantly different from zero in all the cases except in one. However, pricing errors are noticed to be on average quite but some large pricing errors critically reduce the overall pricing accuracy of the model. When studying the dynamics of the curves, we note that trend explains most of the variations for oil and pulp. However, also tilt and convexity have relatively large explanation power of the variations.

Keywords

Commodity, forward curve, optimization

List of figures

List of tables

List of appendices

1. Introduction

1.1 Background and motivation of the study

Commodities have been growing in importance because of the economic growth and rapid expansion of Chinese economy has boosted the demand for commodities such as steel, freight, oil and paper rapidly. In addition to producers and refineries of commodities, investors have also been more and more interested in commodity derivative instruments. Commodity derivatives markets offer good chances for e.g. speculators because of sometimes significant price fluctuations. A good example of commodity price behaviour is the significant price fluctuation in Nordpool[1], the Nordic electricity exchange. On November 14, 2005 the price of one MWh (Mega Watt Hour) fell from over 200 NOK to 20 NOK and rose back to the level of 200 NOK on the next day (http://www.nordpool.no/nordpool/spot/index.html). Nordpool electricity prices between January 1, 1998 and July 1, 2004 are plotted in figure 1-1. An example of another price spike occurred at the end of 2002 when the price of electricity rose from 120 NOK to 803 NOK just in four months and fell back to the level of 200-300 NOK in three months. Similar price fluctuations are quite normal in electricity markets because of the fact that electricity can't be stored.

[1] Norpool website: www.nordpool.no

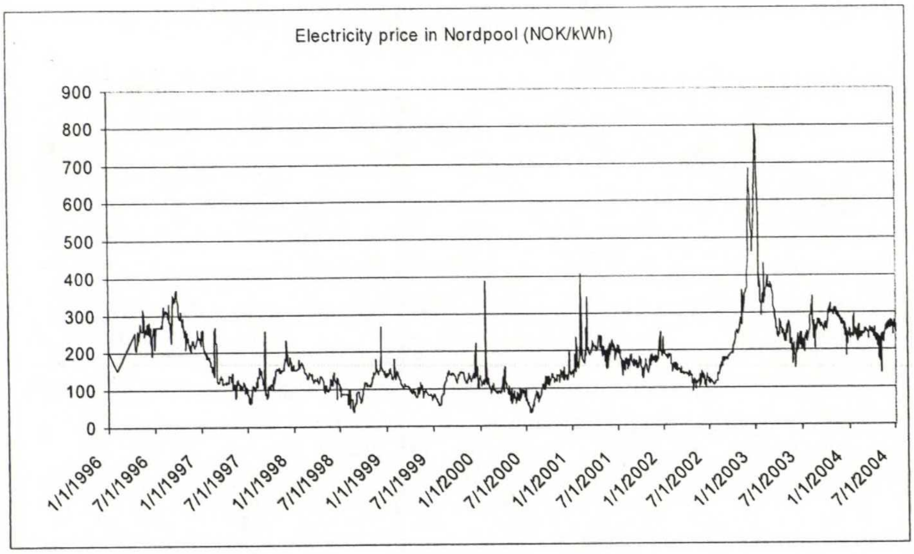


Figure 1-1: Norpool electricity price

In addition to sometimes volatile prices, according to Gorton and Rouwenhorst (2005) commodity derivatives returns are negatively correlated with other asset classes (such as equities) which also enhance investors' interest to commodity derivatives as they can more efficiently build portfolios with lower volatilities and higher returns.

Trading volumes in certain commodity derivatives markets have been growing rapidly during the recent years. Hull (2002) points out that the popularity of commodity swaps is increasing but nevertheless commodities have quite small exposure in his book. A good example of the growing importance of commodity derivatives trading is current development of freight derivative markets. For example, tanker futures transactions rose by 91% from December 2004 to January 2005 in Imarex[2]. In addition, new commodities to trade have been introduced to investors, for example the trading with emission rights has begun, as has been widely discussed in public media recently.

---

[2] Imarex website: www.imarex.no

China's rapidly growing economy has also an important effect on the demand for commodities. According to Roach (2004), China's demand for raw materials and other commodities has a major effect to commodity prices and to trade volumes of commodities, especially the prices of raw materials, steel, oil and freight have increased in the near past. In 2003, from the total world consumption China consumed 25% of aluminium, 27% of steel, 30% of iron ore, 31% of coal and 40% of the global cement consumption. Because of the economic growth in China, India and other developing countries, it can be said with quite high certainty that the activity in commodity markets will at least stay in the current level also in the near future.

Growing interest on commodity derivatives is one of the main motivations for the paper in hand. As the volumes in commodity derivatives increase, the valuation of commodity contingent claims becomes also more and more important. A vital tool for valuing commodity derivatives is a forward curve. Despite the growing importance of commodity derivatives, commodity forward curves have had only a little interest, when compared to interest rate, corporate finance and equity related issues, in academic papers. Papers which have studied forward curves, have mainly concentrated on yield curve construction for interest rates. Existing studies of commodity forward curves concentrate mainly on oil whereas e.g. pulp forward curves haven't had much interest without couple of exceptions (see e.g. Järvinen, 2004). However, this can of course be explained with the fact that practitioners' (industrial companies and traders) interest on pulp derivatives has been lower than that on e.g. oil and metals.

1.2 Objective and contribution of the study

The research problem of the thesis is the construction of forward curves for different commodities. In more detail, I examine the optimization forward curve generation method and test how accurately swaps can be priced using optimized forward curves. I also shortly present four other forward curve construction methods which are bootstrapping, cubic spline, regression and Heath-Jarrow-Morton (HJM) framework. These methods are discussed in more detail in the literature review part later on. The suitability of the optimization method is tested with two commodities which are oil and pulp.

Academics have developed very sophisticated forward curve construction methods which value commodity contingent claims correctly and take into account seasonality, Samuelson hypothesis, correlations etc. but many of these methods are too complicated and too slow to use for practitioners (e.g. traders). For example, Humphreys (2005) state that there are many statistically beneficial models but their costs, time consumption and complexity, outweigh their benefits. I have chosen the optimization method by keeping in mind both the theoretical validity but also the practitioners' point of view i.e. efficiency and time consumption.

In addition, when compared to earlier studies of optimization as forward curve construction method (mainly Järvinen, 2004), I have used both futures price data and par swap quotes. Because of the inclusion of the futures data, I have made some minor changes to the model proposed by Järvinen (2004) in order to take advantage of both futures and swap data.

1.3 Limitation of the study

I have limited the scope of the paper in hand to two commodities which are oil and pulp. These two commodities provide quite wide information of commodity markets because they are very different in nature when the data available is considered. Oil, which is one of the most traded commodities in the world, has liquid markets and the data is easily available. Whereas pulp markets are much smaller in terms of volumes and the number of transactions, thus the data doesn't always reflect current state of the world (demand and supply of pulp).

Time scale of the study is limited to the near past, i.e. data consist of quotes from the end of 1995 to the end of 2004 for oil, while the time window for pulp data is a little shorter as the first price quotes for pulp are from the end of 1997 and the last quotes date to the end of 2002.

1.4 Structure of the study

The structure of the thesis is as follows: section 2 presents existing studies and papers about commodity and interest rate forward curve construction methods as well as studies about the dynamics and characteristics of commodity prices. In this part, I also present other forward curve construction methods and in addition compare the optimization method to those methods. Section 3 proposes the main hypotheses tested in the thesis. The following section 4 describes the data and methods used in the empirical part of the paper. In section 5, the objective is to conduct the analysis of the suitability of the optimization method and to discuss the results obtained from the analysis. Section 6 gives conclusions and summarises the thesis.

2. Literature review

2.1 The definition of commodity

This thesis is about commodity forward curves, so it is useful first to determine what is commodity. As Toivonen (2005) mentions, "*Any intermediate good useful for production with a constant and standard quality may be called commodity, from the Latin word of commodus meaning convenient*". Important words in the quotation are "*useful for production*" and "*constant and standard quality*". The first ensures some value for a commodity, because it can be used in the production of some end product which has value. The second, this is required for a commodity derivatives market to exists, ensures that there are certain norms for commodities. Without the standard quality, there wouldn't be any commodity derivatives markets, because it would be quite hard to determine the underlying product for a commodity derivate instrument. For instance, Chicago Board of Trade[3] (CBOT) has the following specification for deliverable grades of corn futures, "*No. 2 Yellow at par, No. 1 yellow at 1 1/2 cents per bushel over contract price, No. 3 yellow at 1 1/2 cents per bushel under contract price*" (http://www.cbot.com/cbot/pub/cont_detail/1,3206,1213+14389,00.html). This way CBOT can guarantee that both buyers and sellers of corn futures know what they are buying and on the other hand what they are expected to offer.

Even though, the definition of commodity discussed above seems to be quite straightforward in nature, there are some commodities which are not necessarily at the first glance thought to be commodities. For example, the commodity discussed by Koekebakker & Ollmar (2001), i.e. electricity, has some characteristics why it is not easily thought as a commodity. First of all, electricity is not something that one can "drop to their toes" and it can't be stored (this has many interesting implications to the valuation of electricity; see e.g. Kellerhals (2001) or Geman & Roncoroni (2004)). Despite the first "feeling", electricity fulfils both requirements mentioned by Toivonen (2005), electricity is useful in production of other products and it is constant and standard

---

[3] Chicago Board of Trade website: www.cbot.com

quality. Other objects which are not obviously considered to be commodities but are traded in commodity derivatives markets (and obviously also in spot markets as well) are for example freight, emission rights and telephone time. Pulp, oil, gold, silver, corn, orange juice and pork bellies are examples of more obvious commodities.

2.2 Forward price of a commodity

Before we define what a commodity forward curve is, we have to clarify the term forward price. As Hull (2002) states the forward price is "*the delivery price in a forward contract that causes the contract to be worth zero*" when the contract is initiated, but how can the forward price be determined. Introductory derivatives literature (e.g. Willmot (1998) or Hull (2002)) present the following formula for the determination of forward price.

$$F = S(t)e^{r(T-t)}$$

Equation 2-1 Forward price without convenience yield

Equation 2-1 gives simple way to determine forward price if there is no convenience yield. F is forward price, S(t) is spot price at time t and T is the maturity of the contract. In the equation, r is interest rate and finally e is the base number of natural logarithm.

Convenience yield describes the benefit received by the owner of the commodity (Brennan (1991)). Convenience yield will be discussed in more detail below. However, as it can be thought, commodities have convenience yield, because there is apparently some benefit of owning a commodity. For example, a producer of steel clearly gets some benefit if he/she has iron ore, which is used in the production of steel, in stock. On the other hand, there are also some costs incurred when owning a commodity because of the inventory costs. Thus, the factors which are important in valuing commodity forward contracts are storage costs and convenience yield. Widely used formula (e.g. Hull (2002)) for valuing forward (and futures) prices is given below.

$$F = Se^{(r+u-y)T}$$

Equation 2-2 Forward price with convenience yield

F, S and r are similar to those in equation 2-1. u and y describe storage costs and convenience yield, respectively. With equation 2-2 we can value commodity forwards correctly if we know storage costs and convenience yield. Storage costs can fairly easily be measured but convenience yield is much more difficult to measure. In the later chapters I present shortly some papers where methods for measuring convenience yield are discussed.

2.3 Commodity forward curve

As James & Webber (2000) argue, yield curve is a term structure for interest rates i.e. one can look from the curve the interest rate which can be locked today for future borrowing or lending. In the same manner, commodity forward curve is a plot of forward prices which can be locked today for future delivery. Commodity forward curves consist of pure or direct price quotes i.e. there should not be any averaging, discounts or other such factors calculated in the prices. Thus a forward curve is a curve of forward prices.

Forward curve should be separated from other price curves such as par swap curves or simple price curves, which just simply plots the prices of certain instrument (e.g. oil swaps, forwards and futures) with different maturities in the same curve. These curves can't be used in valuation purposes in the same manner as forward curves.

(Commodity) forward curves are used to value commodity contingent claims (such as swaps, swaptions, futures, forwards, caps floors etc). According to Järvinen (2004), forward curves are especially important for OTC[4]-markets where long term commodity derivative contracts are actively traded and public information of the price of certain contract can't be obtained because contracts are customized to different situations and similar contracts can't necessarily be found.

[4] Over-the-counter

2.4 *Average pricing*

Average pricing in commodity markets is very common because it reflects more accurately the usage of a commodity because commodities are usually used continuously during a month, a quarter or a year (or whatever the time frame of the contract is). Thus average prices are more suitable for hedging purposes than point prices. For example NBSK RISI (northern bleached softwood kraft pulp) swaps can be settled against monthly average of the spot price index.

The calculation of average price for certain time period is not always as simple as it seems. For instance in the dry bulk markets (freight) a monthly average consist only the average of the 7 last trading days of a month (see contract specifications in Imarex website). Yearly (annual) dry bulk contracts have also an interesting specification when calculating the settlement price as the settlement price is the average of the monthly averages of the first month of each quarter i.e. the average of January, April, July and October. Average price has a major effect on commodity forward curve construction methods when compared to yield curve generation where average pricing does not exist. This will be discussed in more detail in chapter 4.3.

2.5 *Forward curve generation methods*

The aim of the forward curve construction methods is to produce clear (or direct) forward price curve i.e. forward curve without averaging as discussed above. Studies such as Järvinen (2004), Hagan and West (2004), Fleten and Lemming (2001), Koekebakker and Ollmar (2001), Linton, Mammen, Nielsen and Tanggaard (2001), De Rossi (2004), Diebold and Li (2003), Bhar and Chiarella (1995) and Routledge, Seppi and Spatt (2000) discuss several forward curve construction methods but the vast majority of the papers mentioned concentrate on interest rate curves i.e. yield curves. The following section discusses in more detail the construction methods, especially optimization method, which is used in the analysis.

Forward curve construction methods can be divided roughly into two different categories; dynamic methods and static (fitting) methods. Dynamic models include e.g. HJM[5] framework modified for commodity context (Koekebakker & Ollmar, 2001). Bootstrap, spline methods (cubic spline discussed here) and optimization are examples of static methods.

Dynamic models contain one or more state variables which are estimated from historical data for example with Kalman Filter (Jazwinski, 1970; Maybeck, 1982) and then current forward curve is constructed using these historical state variables and current information about the market prices.

Static methods use usually only current market data (commodity prices and interest rate data) and the forward curve is fitted to the prevailing market prices. When comparing the categories (dynamic and static models), the main advantage of dynamic models is that they offer more information (e.g. term structure of volatilities) about commodity markets. Whereas static models provide less information about the market variables but they are usually much simpler to use and understand.

As Järvinen (2004) states, the main tool used in the following curve generation methods is the pricing formula for swap contracts. Swaps are financial contracts where fixed payments are exchanged to floating payments during the period when the swap is effective (see e.g. Hull (2002) or Willmot (1998)). Thus, a swap contains usually two legs, a floating leg and a fixed leg. The value of the swap contract can be calculated as the net present value of both legs keeping in mind which leg is received and which one is paid. Pricing formula for interest rate swaps is presented in Hull (2002) and it can be used also in the case of commodities because of the similar type of cash flow structure in interest rate swaps and commodity swaps. The main difference between the two is the fact that usually, as mentioned above, settlement price is calculated differently but this doesn't affect the valuation formula itself. As discussed above, settlement price in the

[5] Heath-Jarrow-Morton

case of interest rates is a point price whereas in the case of commodities it is usually calculated as an average from the settlement period.

According to Hull (2002) the value of a swap, when paying fixed and receiving floating leg, can be calculated as follows.

$$V_{swap} = V_{fl} - V_{fx}$$

Equation 2-3 Value of a swap

Where V_{fl} and V_{fx} are the values of the floating leg and the fixed leg, respectively. Calculations of the values of floating and fixed leg are presented in the following equations.

$$V_{fl} = \sum_{i=1}^{N} F_i P_i$$

Equation 2-4 Value of the floating leg

$$V_{fx} = \sum_{i=1}^{N} G_i P_i = G_i \sum_{i=1}^{N} P_i$$

Equation 2-5 Value of the fixed leg

In the floating and fixed leg formulas, F_i is the floating price in payment i and G_i is the fixed price for the swap contract. P_i is the discount factor associated to the payment i.

$$P_i = e^{-r_i t_i}$$

Equation 2-6 Discount factor P

Where r_i is the discount rate for payment i and t_i is time to payment from the valuation date. Thus, it can be seen that the valuation of a swap is fairly straight forward if the prices (future prices) can somehow be estimated. This is where forward curve construction methods step in. Different methods are discussed in more detail in the following chapters.

2.5.1 Bootstrap method

The first forward curve construction method discussed is bootstrap method. This method is an iterative process to construct a forward curve. Hull (2002) presents how to use bootstrapping in interest rate context to generate yield curves from zero and coupon bonds. However, bootstrapping can also be used to generate commodity forward curves as discussed by Järvinen (2004).

According to Järvinen (2004) spot price is the starting point in bootstrap method since it is the first point in the resulting forward curve and spot price is used when calculating following price points in the iterative process. In addition to spot price, the author mentions that also par swap prices are needed as data. The first operation in bootstrapping, is to interpolate par swap prices for the settlement dates of the swaps (i.e. linear interpolation), which are usually month end dates or last business dates in a month.

Next step in the procedure is to solve implied average forward prices iteratively by using the swap valuation formula presented above. But as Järvinen (2004) mentions, implied average forward prices can't be effectively used in the valuation of other commodity derivatives, direct forward prices need to be extracted from average prices. Järvinen (2004) presents alpha factor which has to be used in the extraction. Alpha factor is discussed in section 4.3.1. in more detail. With the help of alpha factor, monthly average prices can be estimated even though only month end prices are known.

The main advantage of bootstrapping is that the method is very straightforward and easy to implement. In addition, the resulting forward curve we can value e.g. swaps correctly with no pricing errors. However, as Järvinen (2004) points out, the method has one major drawback; the resulting forward curves are saw tooth shaped and thus are not really useful in the real world. The following picture presents an example of a forward curve generated with bootstrap method.

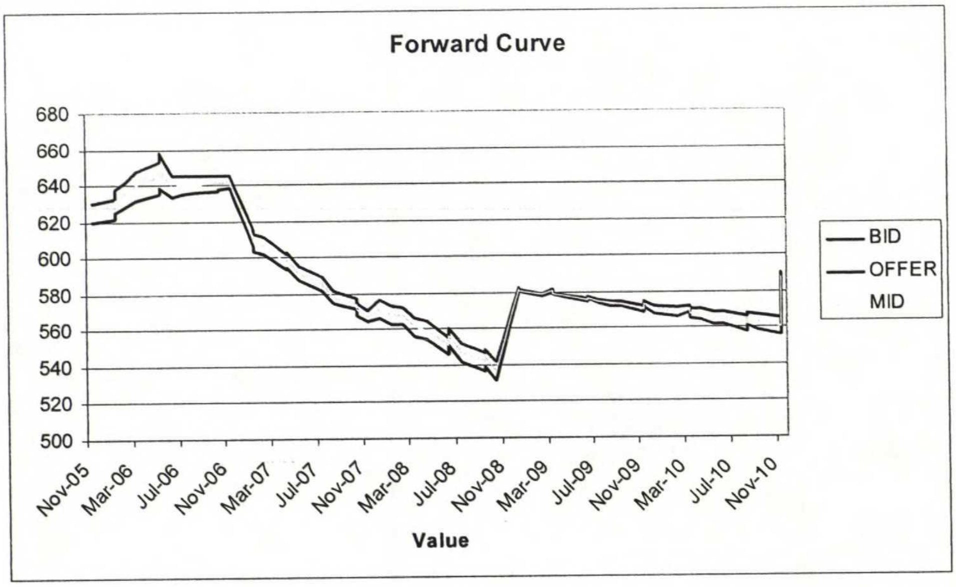


Figure 2-1 Forward curve generated with bootstrap method

The forward curve in figure 2-1 is constructed from the following dataset.

	BID	OFFER
SPOT	620	630
SWAP6M	630	645
SWAP1Y	633	645
SWAP2Y	610	620
SWAP3Y	592	602
SWAP5Y	585	590

Table 2-1 Dataset for a bootstrap forward curve (par swap quotes)

As it can be seen in the figure 2-1, the forward curve is quite saw toothed and has large price fluctuations. This is in line with the findings of Järvinen (2004). As discussed below, forward curves should be smooth shaped in order that they can be used also in the valuation of other derivatives than just swaps.

2.5.2 Cubic spline

Spline methods can also be used to generate commodity forward curves. Here I present Cubic spline method which is used in commodity context e.g. by Järvinen (2004).The method is based on estimating functions of third degree between knot points (James & Webber (2000)) so that instruments used as data for the curve can be valued as accurately as possible. Forward curve is fitted to the underlying data by linear regression.

In cubic spline method, functions between the knot points are twice differentiable because they are third degree polynomials. In order to create a smooth and continuous forward curve there are two requirements which must be satisfied at each knot point (James & Webber (2000)). First, slopes on both sides of the knot points must match i.e. first order derivatives of the functions on both sides of the knot point have to be equal. This ensures that the resulting forward curve is continuous (Pitkäranta, 2005) everywhere, also in the knot points. In addition, same first order derivatives on both sides of a knot point ensure that there aren't sharp edges (as can be found when using bootstrap) on the resulting forward curve. Second, curvatures on both sides of the knot points have to match. Curvature is measured by the second order derivative and thus also second order derivatives have to be equal on both sides of the knot points. In other words, this finally ensures that the forward curve is smooth.

2.5.3 Heath-Jarrow-Morton (HJM) framework

Heath-Jarrow-Morton (HJM) framework is the only dynamic forward curve generation method discussed in this paper. Initially HJM framework was designed for constructing interest rate term structure curves (see e.g. Heath et al. (1992)). A good example how to use HJM in interest rate term structure modelling is given e.g. by Hull (2002). He also discusses how one factor HJM-model can be extended to several factor model and in addition gives an example how to use it to value interest rate derivatives. Even though HJM framework was initially designed for interest rate context, it can also be used to generate commodity forward curves (see e.g. Miltersen & Schwartz (1998) or Cortazar & Schwartz (1994)).

HJM framework is more complex method to use than e.g. optimization but at the same time, it (HJM) provides more information about the term structure of a commodity (e.g. information about volatility). Next we will describe shortly how Miltersen & Schwartz (1998) use HJM framework in commodity context. The authors use three different price processes in the paper: the spot price process of the underlying commodity, the term structure of forward interest rates and the term structure of future convenience yield. The term structure of forward interest rates is given as follows:

$$f(t,s) = f(0,s) + \int_0^t \mu_f(u,s)du + \int_0^t \sigma_f(u,s) \cdot dW_u$$

Equation 2-7 Forward interest rate process

where f is forward interest rate, W is d-dimensional Wiener process, μ is the drift component and σ is the diffusion term. Similarly future convenience yield and the spot price of the underlying commodity can be modelled:

$$\varepsilon(t,s) = \varepsilon(0,s) + \int_0^t \mu_\varepsilon(u,s)du + \int_o^t \sigma_\varepsilon(u,s) \cdot dW_u$$

Equation 2-8 Future convenience yield process

$$S_t = S_0 + \int_0^t S_u \mu_S(u)du + \int_o^t S_u \sigma_S(u) \cdot dW_u$$

Equation 2-9 Spot price process

Parameters to the three processes (i.e. diffusion terms) have to estimated from historical data. Miltersen & Schwartz (1998) also give explicit equations for the drift terms which are derived with the help of no arbitrage assumption.

$$\mu_S(t) = f(t,t) - \varepsilon(t,t),$$

Equation 2-10 Drift term of spot price

$$\mu_f(t,s) = \sigma_f(t,s) \cdot (\int_t^s \sigma_f(t,v)dv)$$

Equation 2-11 Drift term of forward interest rate

$$\mu_\varepsilon(t,T) = \sigma_f(t,T) \cdot (\int_t^T \sigma_f(t,s)ds) + (\sigma_f(t,T) - \sigma_\varepsilon(t,T)) \cdot (\sigma_S(t) + \int_t^T (\sigma_f(t,s) - \sigma_\varepsilon(t,s))ds)$$

Equation 2-12 Drift term of future convenience yield

Derivation of the drift equations is out of scope of the paper in hand but they are presented in detailed manner in Miltersen & Schwartz (1998). Now we have all the components needed to build up futures term structure by using expected values from the processes. We will use *G(t,T)* as futures price and *F(t,T)* as forward prices, the notation is similar than in Miltersen & Schwartz (1998). Futures prices are

$$G(t,T) = S_t e^{\int_t^T (f(t,s) - \varepsilon(t,s))ds}$$

Equation 2-13 Futures prices

According to the authors, forward prices can be derived from futures prices as follows:

$$F(t,T) = G(t,T) + \frac{S_t}{P(t,T)} Cov(e^{-\int_t^T r_s ds}, \frac{S_T}{S_t} | F_t)$$

Equation 2-14 Forward prices

where $P(t,T)$ is the price of a zero coupon bond and r is the spot interest rate. Thus, HJM allows us to construct forward curves for commodities and the only inputs which need to be estimated from the historical data are diffusion terms i.e. volatilities. The estimation can be done for example with principal component analysis (PCA) (see Blanco et al. (2002) or Koekebakker & Ollmar (2001)). In the next section we discuss in more detail the optimization method which is used to construct the forward curves in the paper in hand.

2.5.4 Optimization method

Järvinen (2004) finds out that the optimization method is the best fitting method to construct forward curves. Other fitting methods in his study were bootstrap and spline methods. Optimization, first introduced as forward curve construction method by Järvinen (2004), is based on the minimization of the squared pricing errors. Thus, this method doesn't provide 100% accurate pricing of market instruments, the reasons will be discussed below. Another pitfall when compared to some more sophisticated methods such as HJM framework is that optimization doesn't provide any information about price process dynamics such as term structure of volatility etc, which would be useful to know when valuing commodity options.

Usually commodity swaps are used as data for optimization but also futures can be included into the dataset. As mentioned above, optimization is based on the minimization of squared pricing errors. In addition to the minimization of the pricing errors there is also a smoothing factor involved. This ensures that the resulting forward curve is smooth and that there aren't very large price fluctuations in the curve. The actual equation for the optimization is as follows (see Järvinen (2004).

$$\min\left[\sum_{i=1}^{N}(V_{fl}(t,T_i)-V_{fx}(t,T_i))^2+\lambda\sum_{i=1}^{N}(F(t,T_i)-F(t,T_{i-1}))^2\right]$$

Equation 2-15 Optimization model for the construction of commodity forward curves

In the equation 2-15 V_{fl} is the price of the floating leg of a swap and V_{fx} is the price of the fixed leg of a swap. Fixed leg prices are observed in the markets and floating leg values are calculated from the resulting forward curve. Both fixed and floating leg prices are calculated by discounting the expected cash flows and summing them up. For example, fixed leg values are calculated as follows:

$$V_{fx} = \sum_{t}^{T} \frac{P_{fx,t}}{(1+r_t)^t}$$

Equation 2-16 Value of the fixed leg in a swap

$P_{fx,t}$ is the fixed price of the swap in time t and usually fixed price is constant over the life of the swap contract. r_t is risk-free interest rate at time t and the rate can be obtained from the interest rate curve. Construction of interest rate curves are out of the scope of this thesis (for examples of construction methods, see James & Webber (2000)). Yield curves used in the paper in hand are generated with bootstrap method. Floating leg values are calculated in the same way with the exception that fixed price is replaced by the floating price associated to time t.

The second part of the equation 2-15 is the smoothing part. λ is the weight or scaling factor which can be determined to reflect the relative importance of the smoothing factor compared to pricing part of the equation. It can also be used to ensure that the first part and the second part in equation 2-15 are relatively on the same level, thus for scaling. The smoothing part minimizes the difference of two consecutive forward prices and hence ensures that the resulting forward curve is smooth, which is one of the requirements for a good forward curve. Criteria for good forward curves are discussed in more detail in sections below.

2.5.5 How to handle futures' prices in forward curve generation

Cox, Ingersoll and Ross (1981) show that theoretically futures and forward prices are equal if interest rates are constant. The difference in prices arises from different timing of cash flows. As Jarrow & Turnbull (1996) or Hull (2002) describe futures and forwards, in the case futures counterparties change margin payments on daily basis whereas in the case of forwards the total payment is made at the end of the contract but nominal payments are the same in both cases. Thus one reason for the differences in forwards and futures prices is the time value of money.

In this study I have used actual interest rate data which leads to the fact that as well theoretically and practically futures and forward prices differ. However, in the forward curve generation I have assumed that futures and forward prices are equal in order to maintain the simplicity of the model. The difference between futures and forward prices is minimal because futures contracts used here have maturities less than one year and the interest rate effect is not material. Theoretically accurate fitting of futures prices into the forward curve model is out of scope of the study and is left to future studies.

2.6 Quality criteria for forward curves

According to Hagan and West (2004) there is no single way to determine which forward curve is the most suitable for a certain situation, but the assessment has to be based on multiple criteria. I.e. there is no single test which tells which forward curve is the best.

Järvinen (2004), Hagan and West (2004) use several criteria for assessing the quality of a forward curve. First and the most important is that a forward curve should satisfy the finance exactly i.e. in the case of commodities, the forward curve should value swaps and other instruments correctly or if not exactly (e.g. in the case of optimization) at least as well as possible. In more detail, this means that when using the curve for valuing the instruments used in constructing the curve, the price obtained from the valuation should be the same as actual market price for the instruments. In many cases when commodities are involved this is not true because of the averaging in derivatives settlement prices introduce difficulties which lead to pricing errors.

Another criterion is the looks of the curve. The curve should look smooth and there shouldn't be too big variations in short time spans. Järvinen (2004) criticises bootstrap method because it generates saw tooth like curves, which we saw in section 2.5.1. Concerning the looks of a curve, Hagan and West (2004) emphasize also that the curve should be positive and continuous. In addition to continuity, authors underline stability. Forward curves should be stable i.e. one unit of change in some input shouldn't affect a material change in the resulting forward curve.

Stability of a forward curve can be statistically tested with in-of-sample and out-of-sample tests. Examples of in-of-sample and out-of-sample concept (however not all are in commodity context) can be found in De Jong et al. (1997), Bollen (1997) or Girma & Paulson (1998).

2.7 *Characteristics of commodity derivatives*

In this section we discuss special characteristics of commodity derivatives. The topics discussed below are important when concerning commodity derivative valuation and also commodity forward curves and their dynamics.

2.7.1 Derivative markets

Clark et al. (2001) give a good definition for (commodity) derivative markets. They argue that "*a derivative market is a financial market on which purely financial instruments representing some underlying physical commodities available for delivery at some future date are traded under various conditions*". They also distinguish two types of markets, namely organised markets and OTC markets. Nordpool, NYMEX[6], CBOT[7] and LME[8] are examples of organised commodity derivative markets. OTC trading is conducted between banks and corporations, thus public price data is not always available of the

[6] New York Mercantile Exchange www.nymex.com
[7] Chicago Board of Trade www.cbot.com
[8] London Metal Exchange www.lme.com

trades in OTC markets. However, some OTC prices are also published, for example OTC pulp prices can be found from FOEX[9]-website.

The products traded in different kind of exchanges discussed vary because of the nature of the market place. As Clark et al. (2001) continue, futures and options are mostly traded in organised exchanges and the instruments traded in OTC markets are forwards, swaps, caps and floors. However, even though Clark et al. (2001) don't particularly mention, options are also traded on OTC markets and it can be said that the distinction of the market places of particular instruments is not very strict. In the OTC markets, the participants can always choose whatever instruments they want to trade as long as they find counterparty for the transaction. The main differences between OTC markets and organised exchanges are not discussed in more detail here, but discussion of the differences can be found from e.g. Jarrow & Turnbull (1996).

2.7.2 Historical returns and correlations

Historically commodity derivatives (futures) have had same kind of average returns than stocks (Gorton & Rouwenhorst (2005)). The authors report that even though the returns for commodity futures and stocks have been at the same level the volatility of commodity futures returns is lower and thus historical average sharpe ratio is better for commoditites than stocks. Both asset classes have return distributions which have fat-tails relative to normal distributions but the distribution of commodity futures returns has more weight in the positive tail when compared to stocks' return distribution. Thus, the authors highlight that commodity futures have less downside risk than equities.

In addition to lower volatility for commodities relative to equities, Gorton & Rouwenhorst (2005) argue that commodity returns have negative correlation with equity returns. In addition, they study the conditional average return of commodity futures when equities have the lowest performance (i.e. 1% of the worst performance). They conclude that when equities average return in the 1% negative tail is -13.87%, average return for

[9] www.foex.fi FOEX Indexes Ltd. is a private company, based in Finland, which provides audited, trade-market pulp and paper price indexes.

commodity futures is 2.38%. Hence, according to authors, commodity futures are a good alternative for portfolio diversification.

Gorton & Rouwenhorst (2005) argue also that commodities have historically offered better inflation hedge than equities or bonds. Equities and bonds have negative correlation coefficients with inflation whereas commodities have positive correlations with inflation in their study. Thus commodities are the only asset class (in the study) whose price movements offset inflation and preserve wealth most efficiently.

2.7.3 Samuelson hypothesis

Samuelson (1965) presented the proposition that the volatility of futures and forwards prices decreases with time to maturity. I.e. futures and forwards whose maturity date is closer to current date have higher volatility than the contracts with longer maturities. The phenomenon is called Samuelson hypothesis or maturity effect. Allen & Cruickshank (2002) discuss the reasons for the Samuelson hypothesis and they underline the amount of information available. I.e. there is more information available concerning the contracts which are closer to expiry than for longer maturity contracts. Increased amount of information reflects better the fundamentals behind spot asset price and futures prices.

Samuelson hypothesis is concerned to be present in commodity futures prices and there are many studies which support the maturity effect (see Dusak-Miller (1979), Castelino & Francis (1982), Milonas (1986)). Also Allen & Cruickshank (2002) founded the evidence to support that maturity effect is evident for most of the commodities they tested. However, there are also studies which have different conclusions i.e. that the increased volatility for nearby futures doesn't arise because of the increased information on the closest contracts but for some other reasons or that the volatility is not larger for shorter contracts at all (see Leistikow (1989), Grauer (1977)). According to Herbert (1995), it is not the maturity which explains the increased volatility of natural gas futures but the trading volume of shorter contracts.

2.7.4 Mean reversion

Mean reversion is a phenomenon where prices of an asset revert back to a certain price or vary around that price (see. e.g Hull (2002)). Bessembinder et al. (1995) studied the existence of mean reversion in various asset classes, including bonds, equities and commodities. They report that mean reversion is not significant with financial assets (bonds and equities) whereas they found strong evidence of mean reversion with commodities. Bessembinder et al. (1995) highlight also the importance of including mean reversion into derivatives pricing models in order to get accurate price estimates. Thus, when constructing forward curves with dynamic forward curve construction methods, we need to include mean reversion factor into the forward price process. In academic papers there are multiple examples how to do this, see for example Nielsen (1999) or Gourieroux & Jasiak (2001). However, it is good to remember that when using fitting methods for the generation of forward curves, we don't have to have any special tricks for mean reversion because mean reversion should already be noted in market prices. The level to which mean reverting prices revert is not necessarily constant over time. This can be seen from the price processes, which have drift terms, proposed by Nielsen (1999) or Gourieroux & Jasiak (2001). In other words, the long run average price level varies over time for some assets.

2.7.5 Convenience yield

Convenience yield is the advantage of owning a commodity now rather than buying the commodity in the future. I.e. according to Brennan (1991), convenience yield describes the benefit received by the owner of the commodity. The owner of a futures or forwards contract on the commodity doesn't receive the benefit because he/she doesn't yet own the commodity but will buy it in the future. Convenience yield is sometimes regarded as dividend yield for commodities (Cortazar & Naranjo (2003)).

Ribeiro and Hodges (2004) estimate convenience yield with two consecutive futures. They present the following equation to calculate annualized convenience yield:

$$\delta_{T-1,T} = (r + c) - 12\ln\frac{F(S,T)}{F(S,T-1)}$$

Equation 2-17 Convevnience yield by Ribeiro and Hodges (2004)

In the equation δ denotes convenience yield, r is risk free interest rate and c storage costs. Authors argue that it is more convenient to calculate convenience yield from the two closest futures rather than from the spot price and the closest future, as it should be done in theory, because spot price data is difficult to obtain, in the case of most commodities. They also state that as well as commodity spot prices have long run averages (see mean reversion) also the convenience yield reverts to its long run average (at least in the case of light crude oil, the commodity which they studied).

According to Casassus & Collin-Dufresne (2002) convenience yield is higher for commodities which are used as inputs in production (e.g. oil) than for commodities which can also be used as wealth storage (e.g. gold). There is a clear logic behind this statement because investors can choose from various ways to store their wealth. On the contrary, a producer of steel must have iron ore in his/her production process and he/she can't choose the input from numerous other commodities.

Seasonal factors can also affect the level of convenience yield as discussed by Dong & Liu (2003). They find that electricity forward prices in PJM[10] have positive convenience yields during winter, spring and fall, thus electricity forward prices are in backwardation (backwardation and contango are discussed in the following section). Whereas in the summer time electricity forwards have low convenience yields and thus forward prices are in contango according to the authors.

2.7.6 Contango and backwardation

According to Keynes (1930, 1937) normal backwardation is a result of a fee paid by a seller to the buyer of the commodity derivative for the privilege of deferring the delivery.

[10] PJM = Pennsylvania, New Jersey and Maryland market for electricity

It implies that the futures price is lower than the expected spot price. It is important to note the term expected spot price when discussing the normal backwardation. Normal backwardation and backwardation differ in the sense of expected spot price and current spot price. Normal backwardation was first presented by Keynes (1930) and it is a theoretical model. According to Keynes and Higgs, one important reason for the normal backwardation is the behaviour of the hedgers and speculators. In normal backwardation case, hedgers hold mostly short positions and they must pay a return to speculators for the risk the speculators bear as speculators hold mostly long positions. Thus, as mentioned, in normal backwardation the futures price is lower than the expected spot price. According to Allen, Cruickshank, Morkel-Kingsbury & Souness (1999), Keynes interpreted normal backwardation as a positive risk premium.

Backwardation and contango compare current spot price to the futures prices. Therefore, forward curve is in backwardation when current spot price is higher than futures prices and in contango in the opposite case. It is worthwhile to notice that a forward curve can be in normal backwardation at the same time when it is in contango (Gorton & Rouwenhorst (2005)). The existence of backwardation and contango in commodity markets has been widely studied. See for example Kolb (1992), Allen et al. (1999), Deaves and Krinsky (1995) or Litzenburger and Rabinowitz (1995). The conclusions of the papers concerning backwardation are quite varying. For example Litzenburger and Rabinowitz (1995) find that oil markets are 77% of the time in backwardation. At the same time Allen et al (1999) conclude, that there doesn't seem to exist a positive risk premium in commodity markets thus normal backwardation is not a characteristic of commodity futures market. Kolb (1992) founds same kind of results.

2.7.7 Theory of storage

Theory of storage is one of the main theories concerning commodity prices. Originally the theory of storage was developed by Working (1949) and Kaldor (1939). Theory of storage states that there is negative correlation (inverse relationship) between convenience yield and the level of inventories. I.e. when inventory levels increase convenience yield should decrease at same time. Thus, as it can be intuitively thought, commodity prices should decrease when inventory levels increase. According to Järvinen (2004) positive correlation between spot price changes and convenience yield changes is induced by the level of inventories.

Because of the importance of the theory of storage, it has been widely studied. One of the most present studies is made by Toivonen (2005). He studied the theory of storage with data on NBSK pulp and Brent crude oil markets. The author found out that there seems to be an inverse relationship between the convenience yield and the level of inventories.

Ribeiro and Hodges (2004) highlight the importance of storage levels in the price of commodities and also in the commodity derivatives pricing. They state that when storage levels are low, convenience yield increases and, naturally, spot price of the commodity will increase faster than futures prices and thus the forward curve is in backwardation, because futures prices are less elastic to the variations on storage levels. The opposite is true for high storage levels and the forward curve will be in contango. The authors emphasize also the relation between volatility of both price and convenience yield to storage levels, both volatilities increase when storage levels decrease.

3. Hypotheses

In this section we will formulate the hypothesis for the empirical testing. The main hypothesis of the thesis is concerned with the pricing accuracy of the forward curves generated with optimization. Formally stated as follows:

> H0: Forward curves generated by optimization value par swaps with no pricing errors.
>
> H1: Pricing errors are statistically significant.

Testing of the hypothesis is conducted in section 5.

4. Data and methodology

In this chapter we discuss the data and the methods used in testing the hypothesis. First we will present the data and at the end of the section, we discuss some aspects of forward curve generation and the testing framework for hypothesis examination.

4.1 Discount curves

Some forward curve constructions methods require also interest rate data in addition to commodity swap, future and forward market prices. Optimization needs interest rate yield curves if the dataset used contains swap contracts since swap prices are calculated by discounting future cash flows to present time. I use USD Libor zero yields as interest rate data in the case that commodity dataset contains swaps. USD Libor rates can be used because all the prices are quoted in US dollars in this thesis. Discount curves used for commodity forward curve generation are constructed by bootstrapping as described by Hull (2002). Interest rate curve bootstrapping is not discussed in more detail here as it is out of scope of the paper.

Interest rate dataset contains daily quotes for 17 different maturity LIBOR rates for the same periods as commodity data. LIBOR maturities are presented in the following table.

Name	Maturity	Currency
LIBORONUSD	Over night	USD
LIBOR1MUSD	1 month	USD
LIBOR2MUSD	2 months	USD
LIBOR3MUSD	3 months	USD
LIBOR4MUSD	4 months	USD
LIBOR5MUSD	5 months	USD
LIBOR6MUSD	6 months	USD
LIBOR7MUSD	7 months	USD
LIBOR8MUSD	8 months	USD
LIBOR9MUSD	9 months	USD
LIBOR10MUSD	10 months	USD
LIBOR11MUSD	11 months	USD
LIBOR12MUSD	12 months	USD
LIBOR2YUSD	2 years	USD
LIBOR3YUSD	3 years	USD
LIBOR4YUSD	4 years	USD
LIBOR5YUSD	5 years	USD

Table 4-1 LIBOR maturities used in discount curve generation

4.2 Commodity data

Commodity data is not as easily available as e.g. equity data. In addition, especially in the case of pulp, because of low liquidity, there might be days when no quote is available for certain derivative product. In this kind of cases I have used the previous available price to complete the dataset for the contract which lacks prices.

Spot prices for some commodities are difficult to obtain and also the spot prices may contain for example discounts. That's why I have used the closest future price as spot price whenever reliable spot price hasn't been available. In the following sections I describe in more detail the construction of datasets for forward curve generation.

Commodity markets differ from other financial markets, namely equity and interest rate markets, in some ways. For example, the number of a company's shares is always known precisely whereas the quantity of a commodity available to trading is never known exactly. In addition, storage levels play an important factor in the commodity pricing, as discussed above. Most of the commodities can be stored with couple of exceptions, e.g. freight, electricity and telephone time. In this chapter I will present shortly the commodities analysed (i.e. oil and pulp) in the paper in hand and also the market places where oil and pulp are traded.

4.2.1 Oil

Oil is one of the most important commodities traded in the modern economic world. For example IMF[11] (2000) has estimated that USD 5 increase in oil barrel price will decrease economic growth by 0.3% in the following year in global scale. Oil is used in many different ways and one of the most important is energy production. According to Statistics Finland[12], in 2001 7.4% of world electricity was produced by burning oil. In addition when all the energy production is concerned, i.e. also other than electricity production, oil is the most used energy source in the world. Statistics Finland states that 50% more Joules were produced in 2000 by oil than by nuclear power and about 30% more when compared to wood based materials. In the same year, other major energy sources represented less than half of the energy amount that was produced by using oil.

The activity of oil derivatives trading can be understood when it is compared to the actual quantity of oil available. It is interesting to note that the level of oil derivatives trading exceeds the annual production of oil (Toivonen (2005)), which means that if all oil derivatives were delivered physically, there wouldn't be enough oil in the world to execute all the deliveries. Oil products are traded in numerous derivatives exchanges and main market places are NYMEX, ICE[13] and TOCOM[14], to mention few.

Oil data consists of weekly brent swap quotes from January 1996 to December 2004. Swap quotes for 6-month, 1, 2, 3, 4 and 5 years are used in the study. In addition to swap quotes, I use also daily brent futures quotes for maturities up to 12 months. Following table presents brent products used here and descriptive statistics for the dataset.

---

[11] International Monetary Fund
[12] Statistics Finland website: www.stat.fi
[13] Intercontinental Exchange website: www.theice.com. Formerly ICE was known as IPE (International Petroleum Exchange)
[14] Tokyo Commodity Exchange website: www.tocom.or.jp.

Instrument	Count	Average price	Standard deviation	Kurtosis	Skewness
BRENT_SPOT	478	23.99	7.708	0.716	0.661
BRENT_1M_FUTURE	478	23.92	7.627	0.756	0.661
BRENT_2M_FUTURE	468	23.65	7.561	0.847	0.745
BRENT_3M_FUTURE	468	23.25	7.440	1.024	0.848
BRENT_4M_FUTURE	478	23.21	7.154	1.250	0.892
BRENT_5M_FUTURE	478	23.02	7.015	1.343	0.941
BRENT_6M_FUTURE	478	22.80	6.862	1.492	1.001
BRENT_7M_FUTURE	478	22.66	6.753	1.515	1.027
BRENT_8M_FUTURE	478	22.41	6.580	1.780	1.112
BRENT_9M_FUTURE	478	22.24	6.452	1.915	1.162
BRENT_10M_FUTURE	478	22.08	6.332	2.047	1.210
BRENT_11M_FUTURE	478	21.92	6.220	2.174	1.256
BRENT_12M_FUTURE	478	21.77	6.101	2.289	1.294
BRENT_1M_SWAP	388	21.47	5.568	-0.844	-0.019
BRENT_2M_SWAP	388	21.26	5.355	-0.862	-0.004
BRENT_3M_SWAP	388	21.05	5.132	-0.882	0.010
BRENT_4M_SWAP	388	20.84	4.924	-0.902	0.023
BRENT_5M_SWAP	272	21.59	5.280	-1.036	-0.427
BRENT_6M_SWAP	389	20.69	4.512	-0.887	-0.038
BRENT_1Y_SWAP	478	21.91	6.096	2.821	1.398
BRENT_2Y_SWAP	478	20.75	5.210	3.543	1.687
BRENT_3Y_SWAP	478	20.35	4.816	3.356	1.730
BRENT_4Y_SWAP	478	20.29	4.741	3.193	1.791
BRENT_5Y_SWAP	478	20.29	4.722	3.278	1.864

Table 4-2 Descriptive statistics for oil dataset

4.2.2 Pulp

If oil markets are one of the most liquid ones, pulp markets have different kind of liquidity features. In the case of pulp most of the trades are completed in OTC markets because exchange traded products are not as numerous as in the case of oil. However, NYBOT[15] has futures and options on futures with NBSK pulp as underlying commodity.

Pulp data contain weekly quotes for NBSKRISI pulp quality spot and swap contracts. The time frame for data is from October 13, 1997 to December 31, 2002 and the data contain spot prices and 6 month, 1 year, 2 year, 3 year and 5 year swap prices for the

[15] New York Board Of Trade website: www.nybot.com

period under review. The following table describes the pulp dataset used in forward curve generation.

Instrument	Count	Average price	Standard deviation	Kurtosis	Skewness
NBSKRISI	273	562.99	78.426	-0.922	0.596
SWP6MNBSKRISI	101	527.48	50.931	2.079	1.637
SWP1YNBSKRISI	273	583.88	72.501	-0.546	0.798
SWP2YNBSKRISI	113	561.48	36.511	0.455	1.234
SWP3YNBSKRISI	273	593.17	46.786	-0.119	0.853
SWP5YNBSKRISI	273	589.16	29.697	-0.766	0.675

Table 4-3 Descriptive statistics for pulp dataset

In the table, NBSKRISI is the spot price, SWP6MNBSKRISI is six month swap contract and other short names are 1, 2, 3 and 5 year swap contracts. It is interesting to note that long term contracts have higher average prices and also lower standard deviations than spot and shorter term contracts.

4.3 Methods

4.3.1 Forward curve generation

As mentioned in the literature review, usually commodity swap settlement prices are monthly averages of the spot prices i.e. calculated from daily or weekly prices. In the optimal situation when generating forward curves, we would obtain daily forward prices from the optimization. However, the calculation capacity of computers and also the data don't support this kind of procedure. For example, when generating 5 year forward curve, we would have 5 * 365 = 1825 variables in the optimization problem and it would take quite a long time to solve the problem. Instead, we solve month end prices in the optimization procedure and have 5 * 12 = 60 variables in the optimization. This reduces considerably the time needed to solve the optimization (minimisation) problem. In addition to the capacity of computers also the data has its limitations because the contracts used have maturities ranging from 1 month to 5 years and thus we don't have contract maturity for every date. For the valuation purposes, after the generation of the forward curve, daily prices can be obtained by interpolating from month end prices.

Month end prices introduce also another problem into the optimization but the problem can in some extend be solved. The problem is that we get only one price quote per month from the optimization even though we would need more in order to calculate the monthly average. Järvinen (2004) presents a solution to the problem as he introduces a weighting factor (here named as alpha) which enables us to calculate a monthly average from the preceding end month price and the current end month price. The symbols used here differ slightly from Järvinen's paper. The first and only parameter needed here is k, which is the number of observations for the calculation of the settlement price. As the author mentions, k is usually the number of trading days in the settlement period. Alpha factor can be calculated as follows:

$$alpha = \frac{k - \sum_{i=1}^{k} \frac{i}{k}}{k} = \sum_{i=1}^{k} \frac{i}{k^2}$$

Equation 4-1 Alpha factor

When k increases, alpha closes to 0.5 but is always greater or equal to 0.5. The limit can be calculated by noticing that

$$\sum_{i=1}^{k} i = \frac{k(k+1)}{2}.$$

Equation 4-2 Sum equation solved

Thus when k goes to infinity,

$$\lim \frac{k - \sum_{i=1}^{k} \frac{i}{k}}{k} = \lim \frac{k - \frac{k(k+1)}{2k}}{k} = \lim(1 - \frac{k+1}{2k}) = \lim(1 - \frac{k}{2k} - \frac{1}{2k}) = 1 - \frac{1}{2} = \frac{1}{2}$$

Equation 4-3 Limit of alpha

For example when k=25, alpha equals to 0.52 and 0.525 when k=20. The value of k is usually between 20 and 25 depending on the month, but here we use constant k (25) in the optimization.

According to Järvinen (2004) average price for month i can now be calculated with the following formula:

$$F_{A,i} = (1 - alpha) * F_{i-1} + alpha * F_i$$

Equation 4-4 Monthly average price calculated using alpha

4.3.2 Method for testing pricing accuracy

The hypothesis in the thesis, as mentioned in section 3, is that there aren't any pricing errors when using forward curves generated by optimization. Pricing errors are calculated by comparing market par swap prices and respective swap prices which are obtained from forward curves. However, we don't make any assumptions of the distribution of pricing errors and that's why we use non-parametric tests to examine the statistical significance of the hypothesis. Wilcoxon matched – pairs signed – rank test is chosen here because it is independent of the distributions of the variable which is tested. Wilcoxon matched – pairs signed – rank test is discussed in more detail below.

4.3.3 Wilcoxon matched – pairs signed – rank test

Wilcoxon matched – pairs signed – rank test (Wilcoxon test) examines the medians of two populations (Easton & McCall website). More information of Wilcoxon test can be found in Lehtonen (1998), Bluman (1997), Cohen & Lea (2004) or Sheskin (1997). Wilcoxon test is non parametric test which doesn't require that variables are normally distributed. However, with large samples, test statistics (marked as Z) can be approximated from normal distribution and we can easily determine the significance of the hypothesis.

In Wilcoxon test we must first calculate the difference of swap prices obtained from forward curve and market prices for par swaps. Next, absolute differences are ranked so that the smallest absolute rank has rank 1 and the largest difference has the biggest rank. Zero differences are excluded from the analysis and if there are equal differences, they all get the same rank number which is the average of the ranks that they would have otherwise got. The last procedure before calculating test statistics is to calculate the sum of negative differences S- and the sum of positive differences S+. S is chosen to be the minimum of S- and S+. Test statistic Z is calculated as follows:

$$Z = \frac{S - (\frac{N(N+1)}{4})}{\sqrt{\frac{N(N+1)(2N+1)}{24}}}$$

Equation 4-5 Test statistic Wilcoxon matched – pairs signed – rank test

N in equation 4-5 is the total number of pairs i.e. price differences in this case. Z is distributed normally, as mentioned above, with mean 0 and variance of 1. The significance of the hypothesis is then determined by comparing Z to normal distribution critical values with the chosen confidence level.

4.3.4 Principal component analysis

The dynamics of generated forward curves are studied with the help of principal component analysis (later PCA). One important aim of PCA is to reduce dimensions, this way we can analyze the most important sources of information (Alexander (2001)). In the case of forward curves we can analyze the causes of the changes in the curves. As the result of the PCA we get coefficients for trend, tilt and convexity (Alexander (2001)) and we can analyze these coefficients to determine which factors explain the movements of the curve over time.

According to Alexander (2001), the data needed as input into PCA must be stationary and because prices are usually non-stationary but returns aren't, we must calculate the returns form forward prices. Thus, we calculate the return of different maturities over time, e.g. the returns of 1 year forward price with the following equation:

$$r = \ln(\frac{P_t}{P_{t-1}})$$

Equation 4-6 Return over time

In the equation, r is return and Ps are prices for times t and t-1. In this case, we calculate daily returns over time. The returns must next be normalized so that their mean is 0 and standard deviation equals to 1. If we don't normalize the returns, the first principal component will be dominated by the input variable which has the biggest volatility. Actual calculation of eigenvalues is out of scope of the paper in hand but detailed information can be found from Alexander (2001).

However, the results of the PCA are discussed shortly in this chapter. The most interesting principal components are three first ones (trend, tilt and convexity) marked as P1, P2 and P3, respectively. Trend can also be thought to represent parallel shifts in curves. The following table is an example of the results from a PCA analysis which has 14 variables (example is discussed in Alexander (2001)).

Component	Eigenvalue	Cumulative R^2
P1	11.01	0.786
P2	1.632	0.903
P3	0.4963	0.938

Table 4-4 Example of PCA results

As shown in previous table, trend explains 78.6%, tilt explains 11.7% and convexity explains 3.5% of the total variation of the yield curve analyzed in the example. This way we can explain over 93% of the total variations of the curve with just three principal components.

5. Analysis and results

In this chapter we discuss the results of the study in hand. As discussed above, pricing accuracy is tested with Wilcoxon matched – pairs signed – rank test. Because we don't have assumption of the direction of possible pricing errors we will conduct two-tailed test. Testing is conducted with 95% confidence level which leads to critical value of +/- 1.96. Generated oil forward curves are displayed in the following figure.

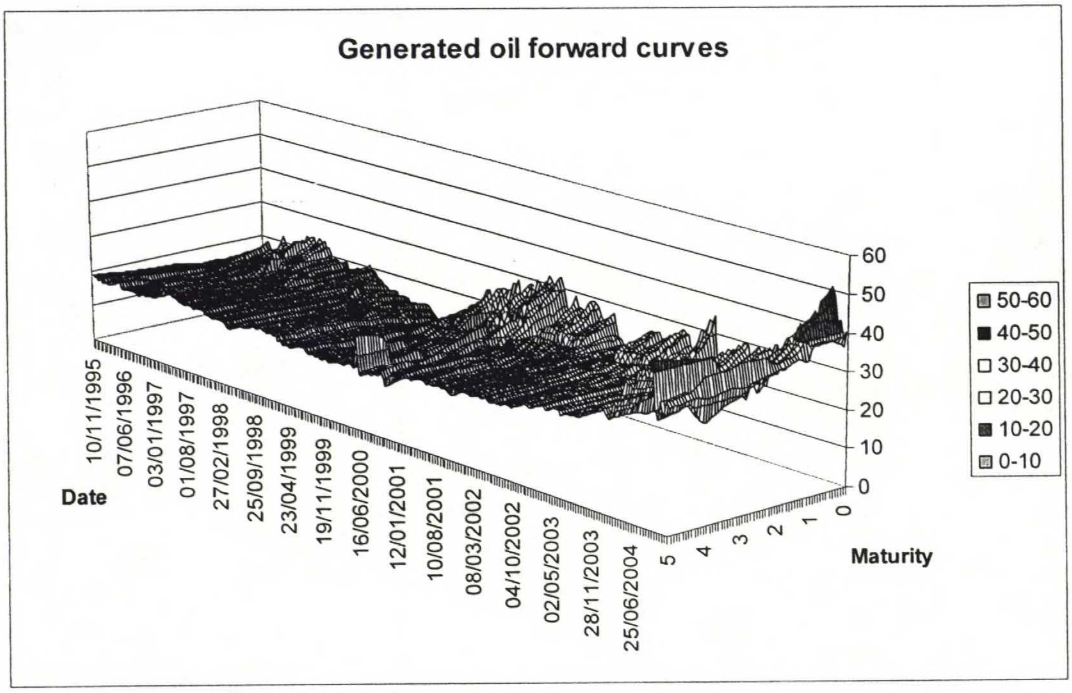


Figure 5-1 Generated oil forward curves

5.1 Oil pricing accuracy

The following table presents test statistics for different maturities and their respective P-values in the case of oil. Maturities are displayed as months in the following table.

Maturity	6	12	24	36	48	60
	Wilcoxon Matched-Pairs Signed-Rank Test					
Z	-9.99627	-0.24937	-8.19062	-9.34797	-10.61618	-12.57742
P-values	7.9127E-24	0.4015	1.2994E-16	4.4674E-21	1.2530E-26	1.4053E-36

Table 5-1 Oil pricing accuracy test statistics

As can be seen from the previous table, in five out of six cases the null hypothesis of no pricing errors is rejected which means that optimization doesn't produce accurate price in statistical view point. However, 1 year (12 month in the picture) par swaps are priced correctly, in statistical terms, when optimized forward curves are used.

Naturally the next question which arises from the results above is that how big are pricing errors. The following figure illustrates absolute pricing errors over time and the table following the figure presents mean absolute errors for all the maturities.

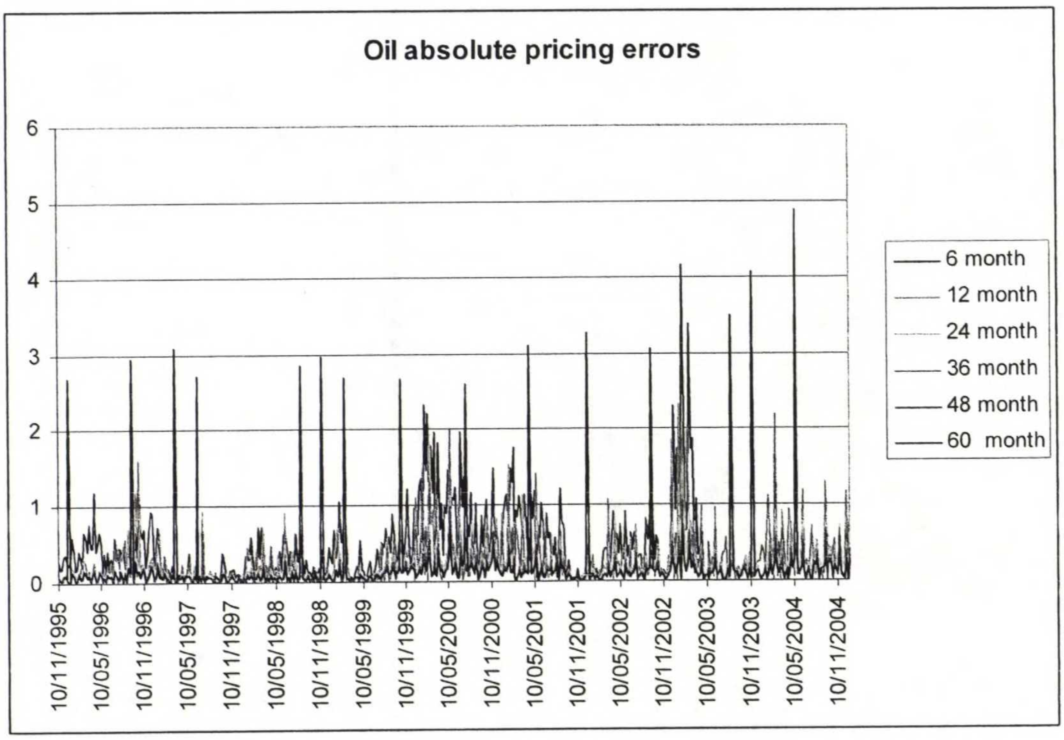


Figure 5-2 Oil absolute pricing errors

Mean absolute error						
Maturity	6	12	24	36	48	60
MAE	0.483	0.307	0.180	0.139	0.119	0.200
% of average market price	2.322%	1.403%	0.867%	0.682%	0.586%	0.986%

Table 5-2 Mean absolute pricing errors for oil

Mean absolute pricing errors are relatively small, except 6 month maturity, when compared to market par swap prices. Even though analysis so far seems to illustrate that the optimization isn't a good forward curve construction method, closer analysis reveals some interesting aspects of pricing errors. The distribution of pricing errors is shown in the following table.

Percentage distribution of absolute pricing errors						
Scale / Maturity	6	12	24	36	48	60
0.00	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
0.25	39.69%	62.34%	76.99%	85.77%	90.38%	90.38%
0.50	24.74%	19.67%	16.74%	11.92%	9.00%	6.07%
0.75	15.46%	7.11%	4.81%	1.88%	0.21%	0.42%
1.00	8.51%	4.81%	1.05%	0.21%	0.42%	0.00%
1.25	5.41%	2.72%	0.21%	0.21%	0.00%	0.00%
1.50	2.32%	0.84%	0.00%	0.00%	0.00%	0.00%
1.75	1.03%	0.84%	0.21%	0.00%	0.00%	0.00%
2.00	1.55%	0.63%	0.00%	0.00%	0.00%	0.00%
2.25	0.26%	0.84%	0.00%	0.00%	0.00%	0.00%
2.50	0.52%	0.21%	0.00%	0.00%	0.00%	0.00%
2.75	0.00%	0.00%	0.00%	0.00%	0.00%	1.05%
3.00	0.00%	0.00%	0.00%	0.00%	0.00%	0.63%
3.25	0.00%	0.00%	0.00%	0.00%	0.00%	0.63%
3.50	0.26%	0.00%	0.00%	0.00%	0.00%	0.42%
3.75	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
4.00	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
4.25	0.26%	0.00%	0.00%	0.00%	0.00%	0.21%
4.50	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
4.75	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
5.00	0.00%	0.00%	0.00%	0.00%	0.00%	0.21%

Table 5-3 Distribution of absolute pricing errors (oil)

Table 5-3 can be read as follows: left panel defines upper limit of the classification, i.e. 1.75 means errors between 1.50 and 1.75. Thus, the figure in cell which is referenced with row 1.75 and column 24 means that 0.21% of 24 month par swap pricing errors are between 1.50 and 1.75. As can be seen from the table 5-3, most of the pricing errors in all maturities are smaller in size than 0.25. For example 36 month par swaps have nearly 86% of pricing errors smaller than 0.25 even though pricing errors for that maturity were considered to be statistically significantly different from zero. Another important aspect to be noted from table 5-3 is that the model seems to be working better for longer maturities as the percentage amount of small errors increase when maturity increases. It is also good to notice, that even though the amount of small pricing errors increase when maturity increases also large pricing errors have the highest frequency in the case of five year (60 months) par swaps.

5.2 Oil curve dynamics

Detailed results of PCA for oil are displayed in appendices 1 and 2. The most important results are summarized in the following table.

Principal component	Trend	Tilt	Convexity
Value	0.004745844	0.002486592	0.001632251
Explanatory power	50.445%	26.431%	17.350%

Table 5-4 Summary of PCA results for oil

Thus, trend (or parallel shifts, as discussed in section 4) explains slightly over 50% of forward curves changes in the case of oil. Tilt and convexity explain 26.4% and 17.4%, respectively. Hence, three of the principal components explain 94.2% of the curve movements while all other principal components explain the rest.

5.3 Pulp pricing accuracy

Pulp curves generated with optimization are presented in the next figure.

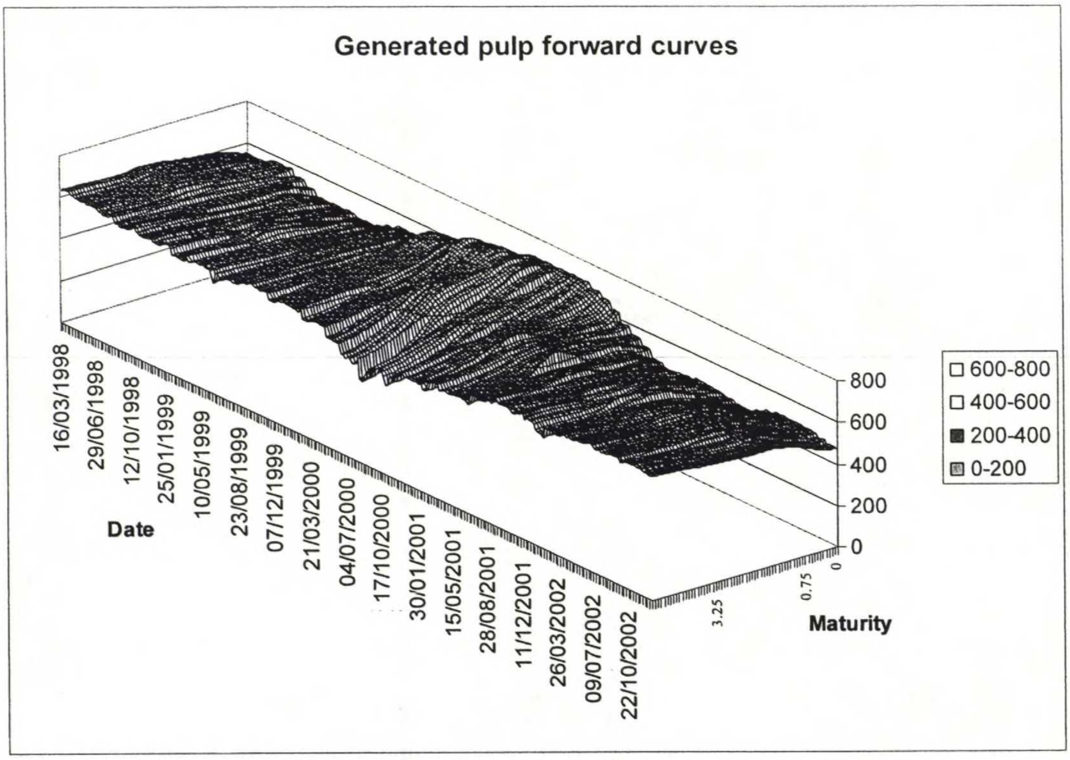


Figure 5-3 Generated pulp forward curves

The following table illustrates test statistics for pulp.

Wilcoxon Matched-Pairs Signed-Rank Test

Maturity	6	12	24	36	60
Z	-6.49579337	-4.47689612	-5.18680086	-2.6043427	-2.24974465
P-value	4.1298E-11	3.7868E-06	1.0697E-07	0.0046	0.0122

Table 5-5 Pulp pricing accuracy test statistics

As was the case earlier when oil was analyzed, the model gives also statistically significant pricing errors with 95% confidence level in the case of pulp. Yet, the pricing error of five year (60 month) par swaps would be statistically insignificant if confidence level would have been chosen to 99%.

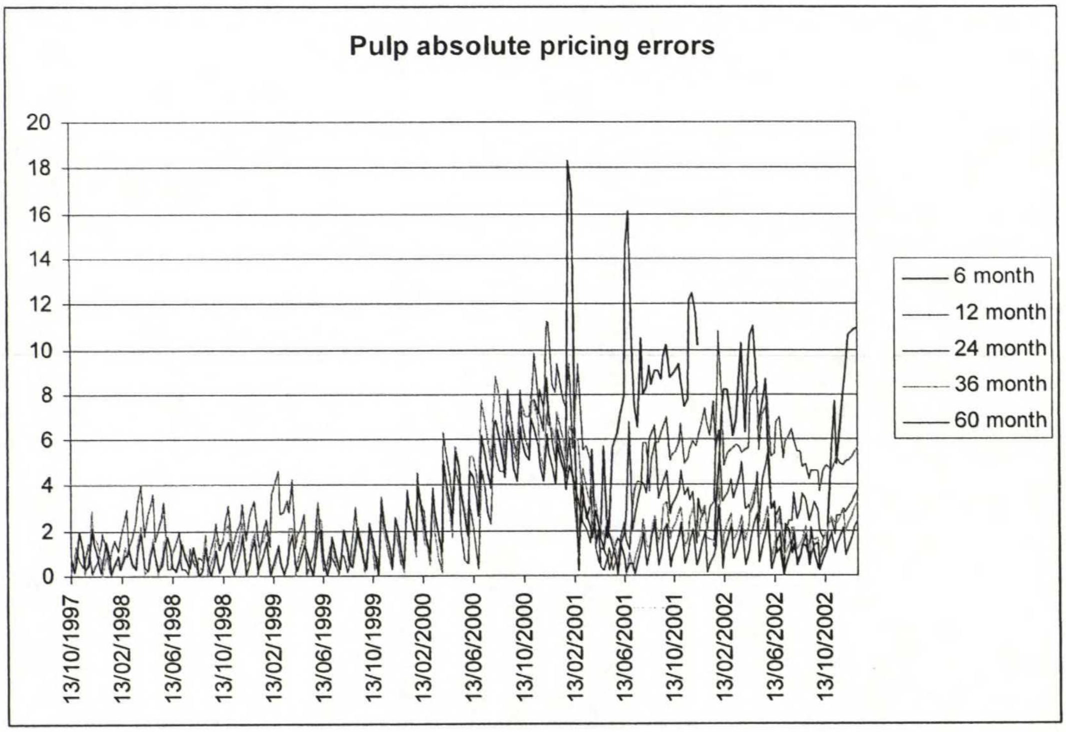


Figure 5-4 Pulp absolute pricing errors

The plot of pricing errors over time (figure 5-2) gives quite mixed results. First of all, long maturity contracts have fairly low pricing errors in the beginning of the time series (1995-2000) how ever towards the end of the period it looks like pricing errors increase at the end of 2000.

Mean absolute pricing errors for pulp are displayed in the next table.

	Mean absolute error				
Maturity	**6**	**12**	**24**	**36**	**60**
MAE	6.378	3.590	3.478	2.262	1.813
% of average market price	1.209%	0.615%	0.619%	0.381%	0.308%

Table 5-6 Mean absolute pricing errors for pulp

Even though pulp has much larger mean absolute errors than oil, percentage amounts are to a great extent smaller than those for oil. Pulp and oil has similar results, when concerning mean absolute errors, in the sense that mean errors decrease when maturity increases.

Percentage distribution of absolute pricing errors

Scale / Maturity	6	12	24	36	60
0	0.00%	0.00%	0.00%	0.00%	0.00%
1	10.89%	18.32%	5.31%	28.21%	37.73%
2	9.90%	19.05%	11.50%	29.30%	32.23%
3	7.92%	10.26%	26.55%	19.78%	11.36%
4	3.96%	8.79%	29.20%	6.96%	5.49%
5	2.97%	10.26%	9.73%	3.30%	5.49%
6	4.95%	16.12%	6.19%	3.66%	4.40%
7	10.89%	7.69%	6.19%	4.40%	3.30%
8	8.91%	4.03%	3.54%	4.03%	0.00%
9	13.86%	2.93%	1.77%	0.37%	0.00%
10	8.91%	1.47%	0.00%	0.00%	0.00%
11	9.90%	0.37%	0.00%	0.00%	0.00%
12	0.99%	0.73%	0.00%	0.00%	0.00%
13	1.98%	0.00%	0.00%	0.00%	0.00%
14	0.00%	0.00%	0.00%	0.00%	0.00%
15	0.99%	0.00%	0.00%	0.00%	0.00%
16	0.00%	0.00%	0.00%	0.00%	0.00%
17	1.98%	0.00%	0.00%	0.00%	0.00%
18	0.00%	0.00%	0.00%	0.00%	0.00%
19	0.99%	0.00%	0.00%	0.00%	0.00%
20	0.00%	0.00%	0.00%	0.00%	0.00%

Table 5-7 Distribution of absolute pricing errors (pulp)

The effect that absolute pricing errors decrease when maturity of par swaps increases is even more easily observed in the case of pulp than it was in the case of oil. As it can be seen from the table, all the pricing errors of 5 year swaps (60 month) are smaller or equal to 7 where as for 6 month swaps over 48% of the errors are greater than 7.

5.4 Pulp curve dynamics

In this section we will discuss shortly the explanatory power of principal components to pulp curve variations. Complete PCA details can be found in the appendices, but here we discuss the results of the analysis. The following table summarizes the most important findings.

Principal component	Trend	Tilt	Convexity
Value	0.000339631	0.000223124	6.93904E-05
Explanatory power	50.408%	33.116%	10.299%

Table 5-8 Summary of PCA results for pulp

Parallel shifts (i.e. trend) explain slightly over 50% of curve variations which is very similar result to oil. However, tilt nearly 7% units more of the variations in the case of pulp than what were found for oil. On the other hand, the difference in explanatory power of convexity is opposite when compared to oil, i.e. for pulp convexity explains only about 10% whereas for oil it explains over 17% of the forward curve variations.

5.5 Reasons for pricing errors in the model

As was mentioned above, optimization produces statistically significant pricing errors for both of the commodities analyzed here, oil and pulp. In section we discuss potential reasons for pricing errors and also remedies to the model.

First of all it is good to notice that the model won't produce zero pricing errors in any circumstances because of the smoothing factor discussed in section 2.5.4. If smoothing factor was omitted from the model, the resulting forward curves would be saw toothed and would have large short time variations which are not a good thing when concerning the pricing ability of the curve. Despite the fact that smoothing induces pricing errors, we can decrease the importance of smoothing by reducing the weight (λ) of smoothing factor. This way we can improve the importance of pricing accuracy and the sum of pricing errors will have larger penalty factor in optimization.

Another source of pricing errors is the calculation of monthly averages with the help of alpha factor discussed in chapter 2. As Järvinen (2004) explained, we need to use alpha factor because we have only month end prices as variables in optimization but we need to calculate monthly averages for pricing purposes. If we could increase the number of price points per month, we could actually increase the pricing accuracy. However, the computational capacity, as discussed earlier, introduces some limitations to the increase of the number of optimization variables.

6. Summary and conclusion

This paper has studied optimization as forward curve construction method with datasets of oil and pulp. The main objective was to test whether optimization generates forward curves which can be used to value commodity swaps correctly in the view point of statistical significance. In addition we studied the dynamics of the resulting forward curves by principal component analysis. Oil and pulp data was used in the analysis and the first data points date to mid 1990s and the last prices date to the end of 2004.

It was found that forward curves generated with optimization model produce statistically significant pricing errors in all but one case. Only 1 year swaps in the case of oil produce pricing errors which statistically significantly zero with 95% confidence level. If we would have chosen the confidence level to be 99% also 5 year swaps in the case of pulp would have produced similar results but the pricing errors of 5 year pulp swaps are considered to be statistically significantly different from zero which was the null hypothesis.

Even though pricing errors were found to be significantly different from zero, we find that the distribution of pricing errors is very favorable as most of the pricing errors are quite small especially in the case of longer swap contracts. This can be easily seen especially in the case pulp, whereas oil has somewhat mixed results in this sense. The density of small pricing errors do increase when the maturity of oil swaps increase but 5 year oil swaps have also much more large pricing errors than 2,3 or 4 year swaps. This is somewhat surprising and it would be good to study the reason for this further in later studies.

In curve dynamics analysis, the main finding was that trend (i.e. parallel shifts) explains the most of the variations in both cases; oil and pulp forward curves. However, in the case of oil tilt explains less than in the case of pulp and the situation is opposite when convexity is analyzed, i.e. convexity explains more in the case of oil than it does in the case of pulp.

In this study we study only the pricing accuracy of forward curved produced by optimization, i.e. we don't examine whether they systemically over or under value swaps. In addition, we don't analyze how pricing errors relate to bid ask spreads because we have only mid prices as data. If mid price curves are situated inside actual market bid-ask spread, pricing errors aren't too important and the curves can be used effectively in valuation of commodity derivatives. Other interesting topics which could be studied further are also the optimal weight of smoothing factor and the estimation of monthly average prices.

7. References

Alexander, C., 2001. Market models: A guide to financial data analysis. Chichester, UK, John Wiley & Sons.

Allen, D., E., Cruickshank, S., 2002. Empirical testing of the Samuelson hypothesis: An application to futures markets in Australia, Singapore and the UK. Working paper.

Allen, D., E., Cruickshank, S., Morkel-Kingsbury, N., Souness,.N., 1999. Backward to the futures: a test of three futures markets. Working paper.

Bessembinder, H., Coughenour, J., F., Seguin, P., J., Smoller, M., M., 1995. Mean reversion in equilibrium asset prices: evidence from the futures term structure. Journal of Finance, vol 50:1, 316-375.

Bhar, R., Chiarella, C., 1995. Estimating the term structure of volatility in futures yield – a maximum likelihood approach. Working paper. University of Technology, Sydney, Australia.

Blanco, C., Soronow, D., Strfiszyn, P., 2002. Multi-factor model for forward curve analysis: An introduction to principle component analysis. Commodities Now, June, 76-78.

Bluman, A., G., 1997. Elementary statistics: A step by step approach, 3rd ed. Boston, McGraw-Hill.

Bollen, N., P., B., 1997. Derivatives and the price of risk. The Journal of Futures Markets, Vol. 17, No. 7, 839-854.

Brennan, M., J., 1991. The price of convenience and the valuation of commodity contingent claims. In: Lund, D., Øksendal, B., (Eds.). Stochastic models and option values. Elsevier Science Publishers, North Holland, 33-71.

Casassus, J., Collin-Dufresne, P., 2002. Maximal affine model of convenience yields implied from interest rates and commodity futures, Carnegie Mellon University. Working paper.

Castelino, M., Francis, J., C., 1982. Basis speculation in commodity futures: The maturity Effect. Journal of Futures Markets, 2, 195-206.

Clark, E., Lesourd, J-B., Thiéblemont, R., 2001. International commodity trading. Wiley, London.

Cohen, B., H., Lea, R., B., 2004. Essentials of statistics for the social and behavioural sciences. NJ, John Wiley & Sons.

Cortazar, G., Naranjo, L., 2003. A multi-factor stochastic model and estimation procedure for the valuation and hedging of commodity contingent claims. Working paper.

Cortazar, G., Schwartz, E., S., 1994. The valuation of commodity contingent claims. Journal of derivatives, summer 1994, 27-39.

Cox, J., C., Ingersoll, J., E., Ross, S., A., 1981. The relation between forward prices and futures price. Journal of Financial Economics, 9, 321-346.

Deaves, R., Krinsky, I., 1995. Do futures prices for commodities embody risk premiums. Journal of Futures Markets 15, 637-648.

De Jong, A., De Roon, F., Veld, C., 1997. Out-of-sample hedging effectiveness of currency futures for alternative models and hedging strategies. The Journal of Futures Markets, Vol. 17, No. 7, 817-837.

De Rossi, G., 2004. Kalman filtering of consistent forward rate curves: a tool to estimate and model dynamically the term structure. Journal of Empirical Finance, Volume 11, Issue 2, 277-308.

Diebold, F., X., Li, C., 2003. Forecasting the term structure of government bond yields. Working paper. University of Pennsylvania, USA.

Dong, L., Liu, H., 2003. Equilibrium forward contracts on nonstorable commodities in the presence of market power. Working paper.

Dusak-Miller, K., 1979. The relation between volatility and maturity in futures contracts. Leuthold, R., M., (ed.) Commodity Markets and Futures Prices, Chicago Mercantile Exchange, 25-36.

Easton, V., J., McColl, J., H. Statistics glossary v1.1. http://www.stats.gla.ac.uk/steps/glossary/nonparametric.html, November 23, 2005.

Fleten, S-E., Lemming, J., 2001. Constructing forward price curves in electricity markets. Working paper. Riso National Laboratory, Denmark.

Geman, H., Roncoroni, A., 2004. Understanding the fine structure of electricity prices. Working paper.

Girma, P., B., Paulson, A., S., 1998. Seasonality in petroleum futures spreads. The Journal of Futures Markets, Vol. 18, No. 5, 581-598.

Gorton, G., Rouwenhorst, K., G., 2005. Facts and fantasies about commodity futures. Working paper. Yale ICF, USA.

Gourieroux, C., Jasiak, J., 2001. Financial econometrics. Princeton University Press.

Grauer, F., L., A., 1977. Equilibrium in commodity futures markets: Theory and tests. Dissertation, Stanford University.

Hagan, P., West, G., 2004. Interpolation methods for yield curve construction. Working paper. School of Computational and Applied Mathematics, Johannesburg, South Africa.

Heath, D., Jarrow, R., Morton, A., 1992. Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. Econometrica, vol 60:1, 77-105.

Herbert, J., H., 1995. Trading volume, maturity and natural gas futures price volatility. Energy Economics 17, 293-299.

Hull, J., 2002. Options, futures and other derivatives, fifth edition. Prentice Hall.

Humphreys, B., 2005. The ABC of PCA. Energy Risk, June 2005, 60-61.

Imarex-website, www.imarex.no, June 29, 2005.

International Monetary Fund, 2000. The impact of higher oil prices on the Global Economy. www.imf.org/external/pubs/ft/oil/2000/oilrep.pdf. November 15, 2005.

Jarrow, R., Turnbull, S., 1996. Derivative securities, 1st ed. South-Western College publishing, Ohio.

Jazwinski, A., H., 1970. Stochastic processes and filtering theory. Academic Press, New York.

Järvinen, S., 2004. Essays on pricing commodity derivatives. Dissertation. Helsinki School of Economics.

Kaldor, N., 1939. Speculation and economic stability. Review of Economic Studies, 7, 1-27.

Kellerhals, B,. P., 2001. Pricing electricity forwards under stochastic volatility. Working paper.

Keynes, J., 1930. A treatise on money: Volume II, The applied theory of money. Macmillan, London.

Keynes, J., 1937. The general theory of employment. Macmillan, London.

Koekebakker, S., Ollmar, F., 2001. Forward curve dynamics in the Nordic electricity markets. Working paper. Norwegian School of Economics and Business Administration, Norway.

Kolb, R.W., 1992. Is normal backwardation normal? Journal of Futures Markets 12, 75-91.

Lehtonen, T., 1998. Tilastotieteen jatkokurssi. Hakapaino Oy, Helsinki.

Leistikow, D., 1989. Announcements and futures price variability. Journal of Futures Markets 9, 477-486.

Linton, O., Mammen, E., Nielsen, J., P., Tanggaard, C., 2001. Yield curve estimation by kernel smoothing methods. Journal of Econometrics 105, 185-223.

Litzenberger, R., H., Rabinowitz, N., 1995. Backwardation in oil futures markets: theory and evidence. Journal Of Finance 5, 1517-1645.

Maybeck, P., S., 1982. Stochastic models, estimation and control. Academic Press, London.

Milonas, N., T., 1986. Price variability and the maturity effect in futures markets. Journal of Futures Markets 6, 443-460.

Miltersen, K., R., Schwartz, E., S., 1998. Pricing of options on commodity futures with stochastic term structures of convenience yields and interest rates. Journal of Financial and Quantitative Analysis, 33:1, 33-59, March 1998.

Nielsen, L., T., 1999. Pricing and hedging of derivative securities. Oxford University Press.

Pitkäranta, J., 2005. TKK:n laaja matematiikka 1. Edita, Helsinki.

Ribeiro, D., Hodges, S., 2004. A two-factor model for commodity prices and futures valuation. Working paper.

Roach, S., 2004. When China sneezes. Morgan Stanley, Global Economic Forum. http://www.morganstanley.com/GEFdata/digests/20040503-mon.html, November 15, 2005.

Routledge, B., R., Seppi, D., J., Spatt, C., S., 2000. Equilibrium forward curves for commodities. The Journal of Finance, Vol 3, 1297-1338.

Samuelson, P., A., 1965. Proof that properly anticipated prices fluctuate randomly. Industrial Management review, 6, 41-49.

Sheskin, D., J., 1997. Handbook of parametric and nonparametric statistical procedures. New York, CRC Press.

Toivonen, H., 2005. Modeling, valuation and risk management of commodity derivatives. Dissertation. Helsinki School of Economics.

Willmott, P., 1998. Derivatives. Wiley, England.

Working, H., 1949. The theory of the price of storage. American Economic Review, 39, 1254-1262.

8. Appendices

Appendix 1: PCA calculation for oil

PCA Calculation

Covariance Matrix

	Spot	1Yr	2Yr	3Yr	4Yr	5Yr
Spot	0.002717022	0.00079328	0.000432716	0.00032346	0.0002737	0.0003061
1Yr	0.00079328	0.001620659	0.001024751	0.00076735	0.00062668	0.00053683
2Yr	0.000432716	0.001024751	0.000902615	0.00077263	0.00066961	0.00059814
3Yr	0.000323462	0.000767345	0.000772628	0.00076413	0.00070975	0.00066036
4Yr	0.000273696	0.000626676	0.000669611	0.00070975	0.00072581	0.00070119
5Yr	0.000306105	0.000536831	0.000598144	0.00066036	0.00070119	0.00267771

Eigen Values	0.0024866	0.0047458	0.0000648	0.0000117	0.0004668	0.0016323
Sorted Eigen Values	0.004745844	0.002486592	0.001632251	0.00046683	6.4754E-05	1.1672E-05
% Variance Explained	**50.45%**	**26.43%**	**17.35%**	4.96%	0.69%	0.12%

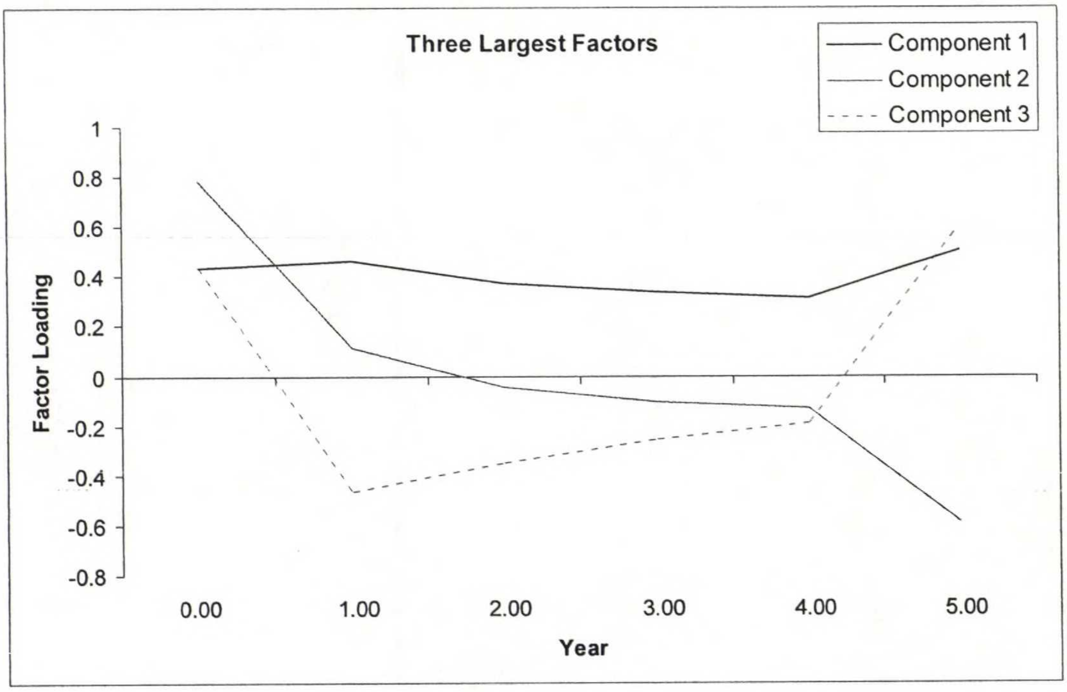
Eigen Vectors

0.780864	0.431334	0.019747	-0.002164	0.111950	0.437350
0.111907	0.461976	-0.312237	0.081141	-0.670656	-0.469255
-0.043107	0.371626	0.723004	-0.451238	0.103055	-0.350806
-0.103297	0.335332	0.120952	0.779773	0.432777	-0.258671
-0.131752	0.310678	-0.603316	-0.426301	0.552499	-0.187463
-0.589781	0.503442	0.027647	0.004171	-0.185786	0.602832

3 Largest Eigen Vectors

	Component 1	Component 2	Component 3
0.00	0.431334088	0.780863971	0.437349807
1.00	0.461975885	0.11190723	-0.469255399
2.00	0.371626466	-0.043107333	-0.350806091
3.00	0.335332073	-0.103297102	-0.258671427
4.00	0.310678422	-0.131751879	-0.187463
5.00	0.503442423	-0.589780586	0.602831917

Appendix 2: Three largest eigenvectors plotted for oil



Appendix 3: PCA calculation for pulp

PCA Calculation

Covariance Matrix

	Spot	0.5Yr	1Yr	2Yr	3Yr	5Yr
Spot	0.000230212	5.17455E-05	1.10236E-05	5.21429E-06	1.35913E-05	-1.39886E-07
0.5Yr	5.17455E-05	9.64669E-05	9.86662E-05	6.66957E-05	3.91794E-05	3.00111E-05
1Yr	1.10236E-05	9.86662E-05	0.000116948	8.46463E-05	5.09857E-05	3.89798E-05
2Yr	5.21429E-06	6.66957E-05	8.46463E-05	8.28477E-05	6.62245E-05	3.89605E-05
3Yr	1.35913E-05	3.91794E-05	5.09857E-05	6.62245E-05	8.2592E-05	4.26227E-05
5Yr	-1.39886E-07	3.00111E-05	3.89798E-05	3.89605E-05	4.26227E-05	6.47035E-05

Eigen Values	0.0002231	0.0000020	0.0003396	0.0000060	0.0000694	0.0000336
Sorted Eigen Values	0.000339631	0.000223124	6.93904E-05	3.36479E-05	6.02347E-06	1.95358E-06
% Variance Explained	**50.41%**	**33.12%**	**10.30%**	4.99%	0.89%	0.29%

Eigen Vectors

0.919665	-0.135207	0.345451	0.069145	0.108697	0.001367
-0.011823	0.730509	0.489896	-0.201101	-0.421455	0.090291
-0.228265	-0.666144	0.521921	-0.235386	-0.407557	0.101171
-0.231314	0.024557	0.433326	0.806954	0.101073	-0.311012
-0.160436	0.047563	0.348566	-0.491846	0.627504	-0.463493
-0.150748	0.038242	0.242885	0.079263	0.490389	0.818569

3 Largest Eigen Vectors

	Component 1	Component 2	Component 3
0.00	0.345451442	0.919665386	0.108697097
0.50	0.489896197	-0.011822628	-0.421454695
1.00	0.521921444	-0.228265388	-0.407557075
2.00	0.433326219	-0.231313843	0.10107303
3.00	0.348565714	-0.160435648	0.627503785
5.00	0.242885476	-0.150748219	0.490388984

Appendix 4: Three largest eigenvectors plotted for pulp