Computing Global Strategies for Multi-Market Commodity Trading

Milos Hauskrecht, Luis Ortiz, Ioannis Tsochantaridis and Eli Upfal
Computer Science Department, Box 1910
Brown University
Providence, RI 02912
{*milos,leo,it,eli*}*@cs.brown.edu*

Abstract

The focus of this work is the computation of efficient strategies for commodity trading in a multi-market environment. In today's "global economy" commodities are often bought in one location and then sold (right away, or after some storage period) in different markets. Thus, a trading decision in one location must be based on expectations about future price curves in all other relevant markets, and on current and future storage and transportation costs. Investors try to compute a strategy that maximizes expected return, usually with some limitations on assumed risk.

With standard stochastic assumptions on commodity price fluctuations, computing an optimal strategy can be modeled as a Markov decision process (MDP). However, in general such a formulation does not lead to efficient algorithms. In this work we propose a model for representing the multi-market trading problem and show how to obtain efficient structured algorithms for computing optimal strategies for a number of commonly used trading objective functions (Expected NPV, Mean-Variance, and Value at Risk).

Introduction

Investment is the act of incurring immediate cost in the expectation of future reward. Investment options represent various tradeoffs between risk and expected profit. Investors try to maximize their expected return subject to the risk level that they are willing to assume. Modern economics theory models the uncertainty of future rewards as a stochastic process defining future price curves. The process is typically Markovian, thus investment decision can be modeled as a Markov decision process (MDP) (Bellman 1957; Howard 1960; Puterman 1994) where a state of the underlying process needs only to include the current investment portfolio and current prices. While the MDP gives a succinct formalization of the investment decision processes it does not necessarily imply efficient algorithms for computing optimal strategies. A challenging goal in this research area is to characterize special cases of the general investment paradigm that are interesting enough from the application point of view while simple enough to allow efficiently computable analytic solutions.

We focus in this paper on commodity trading. Past work has mainly dealt with single market trading problems (see (Dixit & Pindyck 1994; Hauskrecht, Pandurangan, & Upfal 1999) and the references there), where commodity is bought, stored and eventually sold at the same location. Here we address a more realistic scenario in today's "global economy", that of a multi-site trading problem where a commodity can be bought in one location, stored at a second location and eventually sold at a third market. Prices at different locations may be different, and they may have different future price curves. Transportation costs also vary in time. While there can be large gaps in spot prices in different locations, future prices are more correlated - the future price of the commodity at site X cannot be larger than the price at site Y plus the cost of transportation between Y and X. Trading in a "global economy" is significantly more complex, since a local trading decision must be based on expectations about future price curves in all other relevant markets, as well as transportation and storage costs.

Modeling the multi-site commodity trading as a Markov decision process leads to a large state space, and a large action space. Nevertheless, we show in this work that under several commonly used trading utility functions an optimal strategy can still be computed efficiently.

A standard assumption in mathematical economics is that commodity prices (e.g., oil and copper) are best modeled as a *mean reverting* stochastic process (Dixit & Pindyck 1994). In our case, prices in all locations follow the mean reverting process but with different set of parameters for different sites. To solve the trading problem we first consider the *expected net present value (ENPV)* objective function, where the goal is to maximize expected gain with no consideration to risk. Under this objective function the optimization problem becomes myopic and can be computed by considering only current and next step prices. This allows us to design global optimal portfolio allocation algorithms that are polynomial in the number of sites in each trading step.

Building on the myopic property of the ENPV objective function we extend the result to two commonly used objective functions that combine ENPV maximization with limits on assumed risk at any one step. In the *Mean-Variance* function the goal is to maximize a weighted difference of the expected gain and the variance. The *Value at Risk* function maximizes expected gain subject to a (probabilistic) limit on the possibility of a large loss at any one step. Since both functions include a term that is linear in the variance of the process, the optimization problems in both cases lead to a constrained quadratic optimization problem. However, the computational complexities of the two problems are different. The mean-variance function has a particular structure that allows for polynomial time solution. The complexity of the optimization problem for the value at risk function varies, some special cases have polynomial time solutions. To improve the computational efficiency of both methods even further we present structure-based algorithms exploiting the special structure and regularities of the problem.

The Model

We consider investment problems with one type of commodity that is traded at n different sites. Once the commodity is bought it can be either stored in each of the locations or transported between any two locations.

Price model

We assume that trading occurs at discrete time steps. To model commodity price fluctuations we adopt a discrete time version of the mean-reverting model (Dixit & Pindyck 1994):

$$p^{(t+1)} = \mu - e^{-\eta}(\mu - p^{(t)}) + \epsilon^{(t)}, \qquad (1)$$

where μ is the long term average price of the commodity i.e., a value to which the process reverts, η is the speed of reversion and $\epsilon^{(t)}$ is a sequence of independent random variables following normal distribution $N(0, \sigma_\epsilon)$.[1]

Commodity prices at all locations follow mean reverting processes, each with different parameters and with possible correlations between their random components ϵ's. Their combined fluctuations are fully described by a multivariate normal distribution $N(\mathbf{0}, \Sigma)$, with a zero mean vector and a covariance matrix Σ. We assume that price movements are independent of our trading activities. Also, there is no fee for trading and buy and sell prices are the same.[2]

There are natural capacity constraints on the number of commodity units we can transport (store) between the two locations at any time step. However, there are no constraints on buy and sell activities.

Valuation

Profit is measured by the standard *expected net present value (ENPV)* (see e.g. (Brealey & Myers 1991; Trigeorgis 1996)):

$$V^\pi(s) = E(\sum_{t=0}^{T} \gamma^t m^{(t)} | \pi, s) \qquad (2)$$

where s denotes an initial state, π is the trading strategy, $\gamma = \frac{1}{1+r}$ is a discount factor, with r denoting the interest rate (present cost of money), T is the decision horizon, and $m^{(t)}$ is the cash flow at time t. We focus primarily on problems with infinite horizon ($T \to \infty$).

Markov decision process formulation of the problem

A *Markov decision process (MDP)* (Bellman 1957; Howard 1960; Puterman 1994) describes a stochastic controlled process represented by a 4-tuple (S, A, T, R), where S is a set of process states; A is a set of actions; $T : S \times A \times S \to [0, 1]$ is a probabilistic transition model describing the dynamics of the modeled system; and $R : S \times A \times S \to \mathcal{R}$ models rewards assigned to transitions.

In the multi-site commodity trading problem the state of a process is determined by a price vector

$$\mathbf{p} = \{p_1, p_2, \cdots, p_i, \cdots, p_n, p_{11}, p_{12}, \cdots, p_{nn}\},$$

where the p_i's give the commodity price at location i, the $p_{i,j}$'s give the transportation price from i to j, and the $p_{i,i}$'s give the storage price at site i. Actions represent trading activities at a specific time step, and are defined as

$$\mathbf{a} = \{a_{11}, a_{12}, \cdots, a_{ij}, \cdots a_{nn}\},$$

where a_{ij} is the amount of commodity to be transported between i and j, or stored at location i if $j = i$. Thus, actions define allocations of commodity to different transportation (storage) edges. [3]

The transition model is defined by a set of mean-reverting price functions (Equation 1), one for each location. For example, the price movements for location i is

$$p'_i = \mu_i - e^{-\eta_i}(\mu_i - p_i) + \epsilon_i,$$

where p_i and p'_i is the current and next step price, η_i and μ_i are the parameters of the mean-reverting process and ϵ_i is the random component.

[1] We note that normally distributed random components of the price process may lead to negative prices. One way to deal with this issue is to use a geometric version of the mean reverting process, where the logarithm of the price follows the mean reverting model. However, the behavior of such a model is quite different, and price curves of the standard model are more realistic.

[2] In the more general setting (not considered here) prices can also fluctuate based on our demand and supply for the commodity or transportation service.

[3] It is easy to see that the number of units to be transported between different locations is sufficient to define all trading activities. Simply, the number of units to buy and sell at different locations can be obtained by comparing the number of units currently held and the number of units to be transported from that location in the next step.

Rewards represent partial profits from applying the strategy and are modeled in terms of *step-wise gains.* The gain for transporting one unit of commodity between location i and j is defined by

$$g_{ij}(\mathbf{p}) = -p_i - p_{ij} + \gamma p'_j,$$

where p_i is the current price of the commodity in location i, p_{ij} is the cost of transportation and p'_j is the price of the commodity in location j in the next step. The gain for an action $\mathbf{a}$ that allocates commodity to different transportation edges is the combination of partial gains

$$g_{\mathbf{a}}(\mathbf{p}) = g(\mathbf{p}).\mathbf{a} = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}(\mathbf{p})a_{ij}.$$

Using our model, a sequence of cash flows for any strategy can be expressed in terms of step-wise gains (rewards) rather than actual money inflow and outflow. Intuitively, we can replicate payoffs from any strategy by buying the commodity at the beginning of a decision step and selling it at the end of that step. Therefore, the expected NPV model from Equation 2 for a strategy π can be expressed in terms of gains as

$$V^{\pi}(\mathbf{p}) \quad = \quad \lim_{T\to\infty} E(\sum_{t=0}^{T} \gamma^t g^{(t)} | \pi, \mathbf{p}), \tag{3}$$

where $g^{(t)}$ is the gain at time t. This is exactly the discounted, infinite-horizon criterion used commonly in MDPs (Puterman 1994). Thus, our multi-site investment problem for expected NPV model can be expressed and solved as a Markov decision problem.

The optimal trading strategy for the discounted, infinite horizon Markov decision problem is stationary (see (Bellman 1957; Puterman 1994)) and maps states of the process to actions. Therefore, the optimal strategy for our problem is $\pi^* : R^n \times R^{n^2} \rightarrow R^{n^2}$, mapping the current commodity and transportation prices to amounts of units to be allocated to different transportation/storage edges.

Solving the expected NPV problem

Using the MDP formulation, Equation 3 for the expected NPV model and a fixed policy π can be rewritten in Bellman's form (Bellman 1957) as

$$V^{\pi}(\mathbf{p}) = E(g_{\pi(\mathbf{p})}(\mathbf{p})) + \gamma \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} V^{\pi}(\mathbf{p}')f(\mathbf{p}'|\mathbf{p})d\mathbf{p}', \tag{4}$$

where $E(g_{\pi(\mathbf{p})}(\mathbf{p}))$ is the expected one-step gain for $\pi(\mathbf{p})$ and $f(\mathbf{p}'|\mathbf{p})$ is the conditional probability density function of the next step prices.

Myopic property

We see that $V^{\pi}(\mathbf{p})$ is hard to compute exactly. However, despite this difficulty the optimal strategy that maximizes ENPV can be computed efficiently. A key feature of our model is that prices change independently of our trading decisions (see Equation 4). Thus, the optimal policy is *myopic* (a greedy one-step policy is globally optimal) and can be easily computed (see (Hauskrecht, Pandurangan, & Upfal 1999)).

Theorem 1 *The optimal trading strategy for the expected NPV model is myopic.*

Proof The value of the optimal trading strategy is obtained from Equation 4 by maximizing over all possible actions

$$V^*(\mathbf{p}) = \max_{\mathbf{a}} \left[E(g_{\mathbf{a}}(\mathbf{p})) + \gamma \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} V^*(\mathbf{p}')f(\mathbf{p}'|\mathbf{p})d\mathbf{p}' \right].$$

As the next step prices are independent of the action choice, the value can be rewritten as

$$V^*(\mathbf{p}) = \max_{\mathbf{a}}[E(g_{\mathbf{a}}(\mathbf{p}))] + \gamma \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} V^*(\mathbf{p}')f(\mathbf{p}'|\mathbf{p})d\mathbf{p}'.$$

We see that in order to get the optimal solution for $\mathbf{a}$ it is sufficient to optimize $\mathbf{a}$ only with regard to $E(g_{\mathbf{a}}(\mathbf{p}))$. Thus the optimal strategy is myopic. □

The myopic property of the optimal investment strategy is critical for computing the solution for the commodity problem. The complete optimal investment strategy $\pi : R^n \times R^{n^2} \rightarrow R^{n^2}$ allocates the commodity units to different transportation edges for every price vector $\mathbf{p}$. As the number of possible prices and corresponding allocations is very large, it is not feasible to represent and store the optimal policy.

One way to avoid the computation of the complete policy is to compute individual price-specific allocations on-line. The on-line algorithm is invoked repeatedly in every step. In the general case, the on-line phase may be very time consuming as it may require to examine multiple price trajectories spanning multiple time steps. The myopic property of the decision process (Theorem 1) assures that we can obtain the optimal solution just by looking on what can happen in the next step. Simply, in order to decide the best allocation of investment for some price vector $\mathbf{p}$ it is sufficient to choose the allocation with the best one-step expected gain, and it is not necessary to consider more distant future and possible later price movements.

Optimal allocation

To find the optimal trading strategy for the expected NPV model it is sufficient to optimize expected one-step gains. Let $\mathbf{a}$ be some allocation of units to different transportation edges. The expected gain for $\mathbf{a}$ is

$$E(g_{\mathbf{a}}(\mathbf{p})) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}E(g_{ij}(\mathbf{p})).$$

To maximize the expectation we need to maximize the components of the sum. Assuming that C_{ij} is the constraint on the number of units we can transport between location i and j, the optimal allocation of a_{ij} is easy:

$$a^*_{ij} = \begin{cases} C_{ij} & \text{if } E(g_{ij}(\mathbf{p}) > 0; \\ 0 & \text{otherwise.} \end{cases}$$

Simply, we invest the limit to every edge with a positive expected gain.

Objective functions with one-step risk models

Once risk is taken into account, the above strategy of investing the limit on all edges with positive expected gains may not be optimal anymore.

Investment risk can be incorporated into the model in various ways. We focus here on objective functions that penalize or bound risk in any single step. In particular, we investigate:

- Mean-Variance model (Markowitz 1991; Alexander & Francis 1986; Bodie, Kane, & Marcus 1992) that explicitly relates expected one-step gain and the gain variance;
- Value at Risk (VaR) model (Jorion 1996) which maximizes the expected present value of the investment, but at the same time limits possible step losses.

The important property of both models is that their value function is time-decomposable and can be expressed in the form similar to the expected NPV model

$$\begin{aligned} V^*(\mathbf{p}) &= \qquad (5) \\ &= \max_{\mathbf{a}} \left[h(g_{\mathbf{a}}(\mathbf{p})) + \gamma \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} V^*(\mathbf{p}') f(\mathbf{p}'|\mathbf{p}) d\mathbf{p}' \right] \\ &= \max_{\mathbf{a}} \left[h(g_{\mathbf{a}}(\mathbf{p})) \right] + \gamma \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} V^*(\mathbf{p}') f(\mathbf{p}'|\mathbf{p}) d\mathbf{p}'. \end{aligned}$$

Here, $h(g_{\mathbf{a}}(\mathbf{p}))$ is a function of a one-step gain (a random variable), not just its expectation. Different risk models use different forms of h. Note that the optimal policies must be myopic for this formalization.

Mean-Variance (MV) model

The mean-variance model (Markowitz 1991; Alexander & Francis 1986; Bodie, Kane, & Marcus 1992) quantifies the risk in terms of the gain volatility. The model is additive and combines the expected one-step gain and the gain volatility into a single objective function $h_{\mathbf{a}}(\mathbf{p})$:

$$h_{\mathbf{a}}(\mathbf{p}) = \alpha E(g_{\mathbf{a}}(\mathbf{p})) - \beta Var(g_{\mathbf{a}}(\mathbf{p})), \qquad (6)$$

where $\alpha, \beta \geq 0$. Intuitively the function reflects the fact that investors like the mean to be large but dislike the variance. Parameters α, β quantify this relation. We note that this valuation corresponds to the quadratic utility function (Markowitz 1991).

Using the valuation function from equation 6, our goal is to find the allocation of commodity maximizing it. That is:

$$\pi^*(\mathbf{p}) = \arg\max_{\mathbf{a}} \left[\alpha E(g_{\mathbf{a}}(\mathbf{p})) - \beta Var(g_{\mathbf{a}}(\mathbf{p})) \right], \qquad (7)$$

subject to constraints $C_{ij} \geq a_{ij} \geq 0$ for all a_{ij}. The variance of the gain for $\mathbf{a}$ is:

$$Var(g_{\mathbf{a}}(\mathbf{p})) = \mathbf{a}^T \boldsymbol{\Sigma}' \mathbf{a},$$

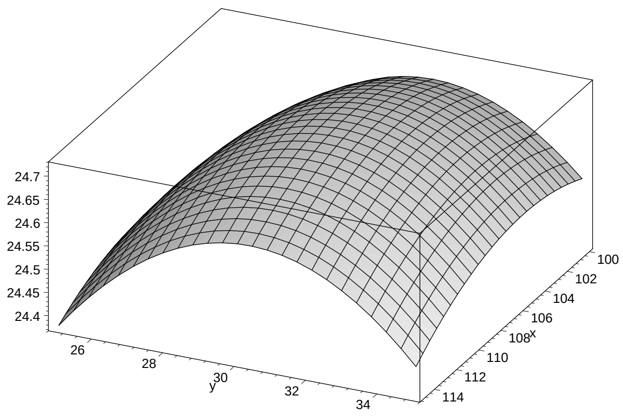

Figure 1: An example of a concave quadratic function for two dimensions.

where $\boldsymbol{\Sigma}'$ is the gain covariance matrix obtained from the price covariance matrix $\boldsymbol{\Sigma}$ as:

$$\boldsymbol{\Sigma}'_{(ij)(kl)} = Cov(g_{ij}(\mathbf{p}), g_{kl}(\mathbf{p})) = \gamma^2 Cov(\epsilon_j, \epsilon_l) = \gamma^2 \boldsymbol{\Sigma}_{jl}.$$

The allocation weights in $\mathbf{a}$ must be non-negative since there is no meaning in our model to negative investment.[4] Also, weights a_{ij} should have only integer values. However, to simplify the problem and its solution we approximate the integer problem by allowing continuous allocation weights.

Solution for the model Equation 7 defines a quadratic optimization problem with linear constraints. The important property of this problem is that the h function has a unique global optimum solution. We can observe this from the fact that the Hessian of our function is a constant negative definite matrix (equal to $-2\beta\boldsymbol{\Sigma}'$).[5] Therefore, the function is concave. Figure 1 illustrates the shape of the function for the 2-dimensional case. This special case of Quadratic Programming is known to have a polynomial time solution (Vavasis 1991).

Exploiting the structure Solving the optimization problem requires to optimize all n^2 possible allocation weights. We show that this optimization can be carried out more efficiently by taking advantage of the problem structure and by solving a sequence of optimizations of smaller complexity.

The idea of our solution is to exploit the regularities of the covariance matrix $\boldsymbol{\Sigma}'$ of one-step gains for all transportation edges, in particular the fact that random

[4] We note that in some of the problems in finance, similar to our problem (e.g. portfolio optimization), constraints on weights can be lifted. This is the case when short-selling of an asset or security is possible. In that case, negative weights in the portfolio will reflect a short position.

[5] Recall that the covariance matrix Σ' is symmetric, positive definite.

components of transportation links leading to the same location are fully correlated. Combining this property with the MV criterion makes it possible to find the optimal allocation incrementally. The idea of the approach is based on the following theorem.

Theorem 2 *Let $\mathbf{a}^*$ be the optimal allocation of commodity maximizing expected gains (returns) and penalizing risk (volatility). Let, (i,j) and (k,j) be two different transportation links ending in the same target location j such that $-p_k - p_{kj} < -p_i - p_{ij}$ holds. Then $a^*_{kj} > 0$ only if $a^*_{ij} = C_{ij}$, otherwise $a^*_{kj} = 0$.*

Proof Gains from transporting one unit of commodity from i to j and k to j are

$$g_{ij}(\mathbf{p}) = -p_i - p_{ij} + \gamma\left[\mu_j - e^{\eta_j}(p_j - \mu_j) + \epsilon_j\right]$$
$$g_{kj}(\mathbf{p}) = -p_k - p_{kj} + \gamma\left[\mu_j - e^{\eta_j}(p_j - \mu_j) + \epsilon_j\right]$$

As the two gains share the same stochastic component and their difference is always deterministic

$$g_{ij}(\mathbf{p}) - g_{kj}(\mathbf{p}) = -p_i - p_{ij} - [-p_k - p_{kj}].$$

Moreover their covariance terms in Σ are the same. Thus, if $p_k - p_{kj} > -p_i - p_{ij}$, there is no value in allocating the commodity to the transport link choice from k before we allocate the maximum, C_{ij}, to a_{ij}. Therefore if $a^*_{kj} > 0$, a^*_{ij} must be saturated ($a^*_{ij} = C_{ij}$). By similar argument, $a^*_{ij} < C_{ij}$ implies $a^*_{kj} = 0$. $\square$

By using this result we can perform the allocation of commodity to different transportation edges incrementally by allocating commodity to edges according to their expected gains, i.e. edges with higher expected gains for the same target location are allocated first. This approach translates to a sequence of quadratic optimization problems with at most n variables.

The algorithm works as follows: the optimization starts by considering only transportation choices with the highest expected gains, one for each target location. We refer to these edges as *active edges*. The optimization procedure for the MV model is then applied to active edges. The solution gives an allocation of units to all active edges. During the optimization a transportation edge can reach its maximum capacity; we say that the edge becomes *saturated*. Once an edge is saturated it is removed and no longer considered as a choice. After the removal, the transportation edge with the next highest expected gain (and the same target location) becomes active and the optimization process continues with the next step. This is repeated until all edges have been exhausted or when none of the edges were saturated in the last step.

The optimization steps are not independent. In particular, every optimization step must take into consideration results of all previous (partial) allocations. The dependencies between the current and previous steps are summarized by:

- a vector of target allocations $\mathbf{s} = \{s_1, s_2, \cdots, s_n\}$, reflecting, for each target location, the number of units of commodity already allocated to edges incident to that location;
- adjusted capacity constraints $\{D_{11}, \cdots, D_{nn}\}$ representing the remaining capacity of all edges, i.e., the original capacity less the capacity already allocated in all previous solutions.

To find the optimal allocation of commodity to active set of edges we solve a quadratic program (with n variables). Let $\tilde{\mathbf{a}} = \{\tilde{a}_1, \tilde{a}_2, \cdots, \tilde{a}_n\}$ denote a vector of allocations for the current set of active edges and $E(g_j(\mathbf{p}))$ be the expected gain for the active edge for target j. Then the optimization task corresponds to:

$$\max_{\tilde{\mathbf{a}}} \left\{ \alpha E(g_{\tilde{\mathbf{a}}}(\mathbf{p})) - \beta\left[(\mathbf{s}+\tilde{\mathbf{a}})^T \tilde{\boldsymbol{\Sigma}} (\mathbf{s}+\tilde{\mathbf{a}})\right] \right\} \qquad (8)$$

subject to constraints:

$$\tilde{D}_j \geq \tilde{a}_j \geq 0 \text{ for all } \tilde{a}_j,$$

where $E(g_{\tilde{\mathbf{a}}}(\mathbf{p})) = \sum_{j=1}^{n} E(g_j(\mathbf{p}))\tilde{a}_j$ is the expected gain for the portfolio of active edges and $\tilde{\boldsymbol{\Sigma}}$ is the *reduced gain covariance matrix*, an $n \times n$ matrix of the gain fluctuations for target locations ($\tilde{\boldsymbol{\Sigma}}_{kl} = \gamma^2 \boldsymbol{\Sigma}_{kl}$). $\tilde{D}_j$ denotes an adjusted capacity constraint corresponding to the transportation link for a target location j which is subject to optimization (is active).

During the computation process we keep track of the number of units allocated to each transportation link (staring from zero allocations at the beginning). That is, after every optimization step we apply the following update:

$$a^*_{ij} \leftarrow \begin{cases} a^*_{ij} + \tilde{a}^*_j & \text{if link } (i,j) \text{ is active;} \\ a^*_{ij} & \text{if not active.} \end{cases}$$

This allows us to recover the optimal allocation $\mathbf{a}^*$ at the end. In addition, we update $\mathbf{s}$ quantities and adjust dynamic capacity constraints:

$$D_{ij} \leftarrow \begin{cases} D_{ij} - \tilde{a}^*_j & \text{if link } (i,j) \text{ is active;} \\ D_{ij} & \text{if not active.} \end{cases}$$

Example Figures 2, 3 and Table 1 illustrate and compare the performance of strategies for different optimality criteria (the Value at Risk criterion is discussed in the next section) on a problem with 5 trading sites. Figure 2 shows the actual step-wise gains obtained for these criteria using a fixed 50-step trajectory of prices' fluctuations; each price following a mean-reverting process. Table 1 summarizes the results in Figure 2 by showing real gain averages and their standard deviations. Finally, Figure 3 compares expectations of gains under different strategies. We see that ENPV always leads to the maximum expected gain and it also achieves higher real gains on average. However, step-wise gains for ENPV are also subject to higher fluctuations. On the other hand, Mean-Variance (MV) criterion yields gains that fluctuate less, but at the same time lead to considerable lower expected gains and also real gains on average.

Besides the experiments shown here, we have tested the performance of the MV model for different combinations of parameters α and β. As expected, higher

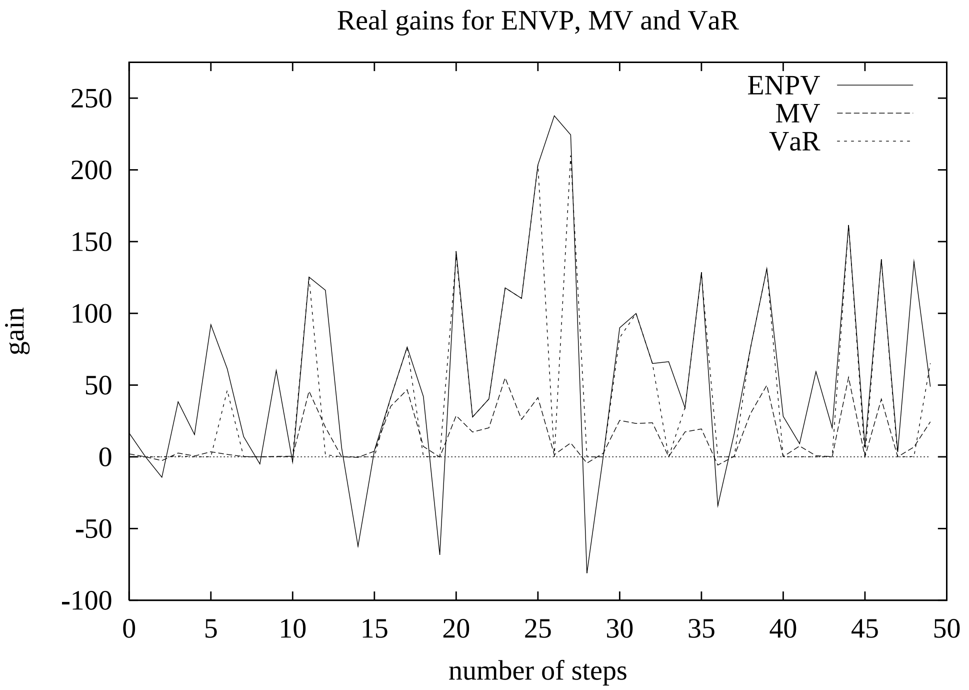

Figure 2: Comparison of three different optimization criteria: Expected NPV (ENPV), Mean-Variance (MV) and Value at Risk (VaR) on a problem with 5 trading sites and 50-step trajectory of prices' fluctuations. each following a mean-reverting process. For each step we plot the real gains for that step. The parameters of the MV model we use are $\alpha = 1$ and $\beta = 0.01$. We use $K = 0$ and $\delta = 0.0005$ for the VaR model.

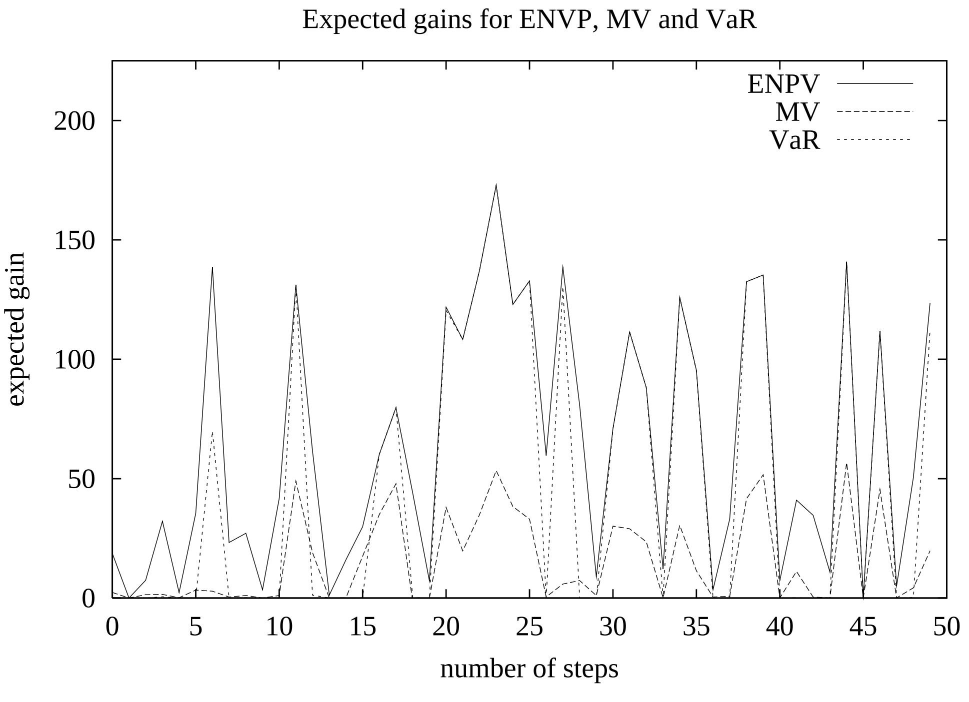

Figure 3: Comparison of expected gains for three different optimization criteria: Expected NPV (ENPV), Mean-Variance (MV) and Value at Risk (VaR) on a problem with 5 trading sites and 50 step long prices' trajectories.

	ENPV	MV	VaR
average real gains	57.31	13.69	42.47
standard deviation	70.64	17.49	60.54

Table 1: Average of the real gains and their standard deviation for ENPV, MV and VaR criteria and data from Figure 2.

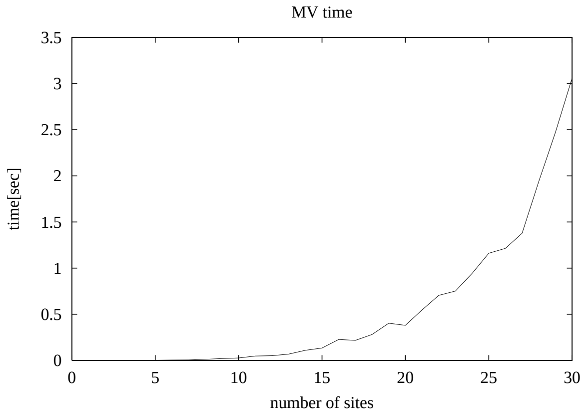

Figure 4: Average running times for markets with varying number of trading sites.

values of β lead to smaller average gains and smaller gain fluctuations. Simply, for higher values of β we penalize the variance more and thus we are likely to sacrifice the opportunity to capture higher gains.

One concern in applying our approach is that the optimization is carried on-line in every step, and thus it may lead to large reaction delays for larger problems (with many trading sites). To see the effect of the size of the multi-site market on the actual running time of the optimization problem we ran a set of experiments, varying the number of trading sites. For each market size we ran 1000 different parameter settings and averaged them. To solve the quadratic optimization problem we use ISML C/Math/Library implementation based on (Goldfarb & Idnani 1983). Figure 4 shows average running times, obtained for different market sizes. The running time (in seconds) increases moderately with the number of sites. In particular, the solution for 30 different trading sites, which is about the practical limit, can be obtained very quickly (in about 3 seconds on a SUN Ultra-10).

Value at Risk (VaR) model

Let K be a loss threshold and δ the maximum probability of losing K or more units. The value of K is called the *value at risk* for δ (see (Jorion 1996)).

This optimization problem has the form of Equation 5, where we maximize

$$h(g_{\mathbf{a}}(\mathbf{p})) = E(g_{\mathbf{a}}(\mathbf{p}))$$

subject to

$$\begin{aligned} C_{ij} \geq a_{ij} \geq 0 &\text{ for all } a_{ij} \\ P(g_{\mathbf{a}}(\mathbf{p}) \leq -K) &\leq \delta. \end{aligned} \tag{9}$$

This is a linear optimization problem with linear and quadratic constraints. Inequality 9 reduces to a quadratic constraint by the properties of the normal distribution. Let x be a normally distributed random variable with mean μ and variance σ^2. Let k be a value such that $P(x \leq \mu - k\sigma) \leq \delta$ holds. The value of k measures the distance from the mean in terms of a standard deviation σ, such that values smaller than $\mu - k\sigma$ occur with probability less than δ. In the case of a normal distribution, k is only a function of δ, and it is independent of μ and σ. Therefore, in order to limit the losses of more than K units with probability $1-\delta$, we set the value of k_δ such that it satisfies $\mu - k_\delta \sigma \geq -K$.

Therefore, the constraint 9 can be rewritten as

$$[E(g_{\mathbf{a}}(\mathbf{p})) + K]^2 - k_\delta^2 Var(g_{\mathbf{a}}(\mathbf{p})) \geq 0 \tag{10}$$

which is quadratic in allocation weights $\mathbf{a}$. We can rewrite the constraint in terms of mean one-unit gains (vector $\boldsymbol{\mu}$) and covariances ($\boldsymbol{\Sigma}'$) as:

$$\mathbf{a}^T \left[\boldsymbol{\mu}\boldsymbol{\mu}^T - k_\delta^2 \boldsymbol{\Sigma}'\right] \mathbf{a} + 2K\boldsymbol{\mu}^T \mathbf{a} + K^2 \geq 0. \tag{11}$$

Let $W = \left[\boldsymbol{\mu}\boldsymbol{\mu}^T - k_\delta^2 \boldsymbol{\Sigma}'\right]$ be the $n^2 \times n^2$ matrix defining the quadratic term. We note that if the matrix W is negative definite, the problem corresponds to the linear optimization over the convex space. Thus, it can be solved efficiently in polynomial time (Papadimitriou & Stieglitz 1998). However, when the matrix W is not negative definite we have a non-convex space over which we optimize. To solve this problem we can apply standard augmented Lagrangian techniques (see e.g. (Bertsekas 1995)).

Using structure to solve the VaR model The optimization of VaR criterion can be performed more efficiently by solving a sequence of optimization problems of smaller complexity. This is the same idea as used for the structured solution of the Mean-Variance model and Theorem 2 also applies to this case. Simply, the only sources of stochasticity are price fluctuations at different target locations. Thus, if two different transportation edges share the same target location, their stochastic component is the same and for the rational and risk averse investor the transportation choice with better expected gain should be chosen first. Therefore, under transportation capacity constraints, the global optimization can be carried incrementally by solving a sequence of optimization problems with n variables, instead of the optimization with n^2 variables. The globally optimal solution is then constructed from results of partial solutions.

To solve the problem, we optimize repeatedly the (reduced) problem with n variables:

$$\max_{\tilde{\mathbf{a}}} E(g_{\tilde{\mathbf{a}}}(\mathbf{p}))$$

subject to,

$$\tilde{D}_j \geq \tilde{a}_j \geq 0, \text{ for all } \tilde{a}_j;$$

$$[SM + E(g_{\tilde{\mathbf{a}}}(\mathbf{p})) + K]^2 - k_\delta^2 \left[(\mathbf{s} + \tilde{\mathbf{a}})^T \tilde{\mathbf{\Sigma}} (\mathbf{s} + \tilde{\mathbf{a}}) \right] \geq 0.$$

The notation used and the basic algorithm applied are the same as in the Mean-Variance case. The only difference is that for the VaR criterion we have to add constant SM which represents the sum of expected gains for all previous solution. This quantity is updated dynamically after every step and is needed to assure that the non-linear constraint is not violated during the optimization process.

Example Figures 2, 3 and Table 1 compare the VaR criterion to ENPV and MV criteria on a problem with 5 sites. We note that the VaR choices do not penalize a large variance when expectation is also high. Instead, it only tries to limit the probability of losses. Thus the real gains obtained for the VaR model vary more than those of the MV model and also tend to achieve higher gains (both under expectation and on average). From the graphs we observe that in many instances the allocations for the VaR criterion replicate exactly the ENPV choices. However, in some instances, when a chance of losses exceeds the confidence threshold, the approach is more conservative and the allocation it chooses is different. For example, in 50 simulation steps in Figure 2 the VaR approach (with threshold gain 0) never lead to the negative gain, while there are seven different cases of negative gains for ENPV and two for the MV criterion.

Conclusion

We addressed the complex problem of finding optimal strategies for trading commodity in a multi-market environment. We investigated various objective criteria based on expected net present value (ENPV) and risk preferences of the investor. Different criteria can lead to optimization problems of different complexity. We showed that under the assumption of equal buy and sell prices, a number of criteria lead to the myopic portfolio optimization problem. This is very important as the computation of the optimal strategy needs to take into account only the current and next step prices and not all possible future price trajectories.

We analyzed and solved the problem for the expected NPV criterion and two commonly used risk-based criteria: Mean-Variance and Value at Risk models. We showed that in both risk-based models the optimization problem reduces to some form of the quadratic optimization problem. To further improve the efficiency of the solution we exploited the structure of the covariance matrix, in particular the fact that gains for the same target locations are fully correlated. This allowed us to reduce a large optimization problem for both risk-based criteria into a sequence of problems of smaller complexity. The empirical results obtained for the mean-variance and value at risk models support the feasibility of the solution and its practical applicability.

We note that our results and algorithms can be applied directly to any multi-site model in which the next-step price fluctuations are normally distributed, and thus not necessarily mean-reverting. The current model can be extended in a number ways. For example, interesting issues will arise if we refine the market models and extend them to include price spreads, trading (buy, sell) constraints, prices sensitive to supply and demands, etc. Another interesting direction is the investigation and application of more complex risk models, reflecting different preferences of an investor.

Acknowledgement

We wish to thank Oliver Frankel of Goldman Sachs for introducing us to this problem and for valuable technical discussions.

References

Alexander, G. J., and Francis, J. C. 1986. *Portfolio Analysis.* Prentice Hall.

Bellman, R. E. 1957. *Dynamic Programming.* Princeton: Princeton University Press.

Bertsekas, D. P. 1995. *Nonlinear programming.* Belmont, MA: Athena Scientific.

Bodie, Z.; Kane, A.; and Marcus, A. J. 1992. *Investments.* Richard D. Irwin.

Brealey, R. A., and Myers, S. C. 1991. *Principles of Corporate Finance.* McGraw-Hill.

Dixit, A. K., and Pindyck, R. S. 1994. *Investment under Uncertainty.* Princeton: Princeton University Press.

Goldfarb, D., and Idnani, A. 1983. A numerical stable dual method for solving strictly convex quadratic programs. *Mathematical Programming* 27:1–33.

Hauskrecht, M.; Pandurangan, G.; and Upfal, E. 1999. Computing near optimal strategies for stochastic investment planning problems. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, 1310–1315.

Howard, R. A. 1960. *Dynamic Programming and Markov Processes.* Cambridge: MIT Press.

Jorion, P. 1996. *Value at Risk: The New Benchmark for Controlling Market Risk.* Irwin Professional Pub.

Markowitz, H. M. 1991. *Portfolio Selection.* Cambridge: Basil Blackwell.

Papadimitriou, C., and Stieglitz, K. 1998. *Combinatorial Optimization: Algorithms and Complexity.* Dover.

Puterman, M. L. 1994. *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* New York: Wiley.

Trigeorgis, L. 1996. *Real Options.* Cambridge: MIT Press.

Vavasis, S. A. 1991. *Nonlinear optimization: Complexity issues.* Oxford: Oxford University Press.