Goodwin, Graham C., Seron, María M., Mayne, David Q., "Optimization opportunities in mining, metal and mineral processing ". Originally published in Annual Reviews in Control Vol. 32, Issue 1, p. 17-32 (2008).

Available from: http://dx.doi.org/10.1016/j.arcontrol.2008.02.002

OPTIMIZATION OPPORTUNITIES IN MINING, METAL AND MINERAL PROCESSING [1]

Graham C. Goodwin * **María M. Seron** *
David Q. Mayne **

* *Centre for Complex Dynamic Systems and Control*
School of Electrical Engineering & Computer Science
The *University* of *Newcastle, Australia.*
** *Department of Electrical and Electronic Engineering*
Imperial College London, UK

Abstract: This paper examines the opportunities arising from the use of optimization in the fields of mining, metal and mineral (MMM) processing. A brief overview of optimization is given. Our main goal in this paper is to raise awareness to the use of optimization as a key enabling technology in the MMM field. *Copyright 2007 IFAC*

Keywords: optimization, scheduling, estimation, control

1. INTRODUCTION

Arguably every design problem involves some form of optimization. However, often the optimization is implicit rather than explicit. We argue here that making optimization explicit has many advantages including the provision of a clear articulation of the design objectives and trade-offs.

Optimization can play a role at many levels in an enterprise including design of elementary feedback loops, coordination of feedback loops, interconnection of unit operations, supply chain management and long term corporate investment strategies.

Examples of the use of optimization to solve real world design problems can be found in many areas. Our goal in this paper is to give a brief overview of optimization. We will also reflect on several case studies arising from our own experience in the fields of mining, metal and mineral processing.

2. BACKGROUND TO MODERN OPTIMIZATION

2.1 General Issues

A general formulation of a typical optimization problem is as follows (Bazaraa et al., 1993; Boyd and Vandenberghe, 2003; Nocedal and Wright, 1999; Nash and Sofer, 1996; Floudas, 1995; Fiacco and McCormick, 1990; Fletcher, 2000; Luenberger, 1989; Gill et al., 1981; Abadie, 1967; Borwein and Lewis, 2000):

$$\begin{aligned}&\text{Minimize } f(x)\\ &\text{subject to:}\\ &g_i(x) \leq 0 \text{ for } i = 1, \ldots, m\\ &h_i(x) = 0 \text{ for } i = 1, \ldots, \ell\\ &x \in X\end{aligned} \tag{1}$$

The set X can contain continuous variables or integer variables. Note that there are many ways of

[1] Originally presented as a plenary paper with the title "Optimization: A key tool for advanced design in scheduling, estimation and control" at IFAC MMM '07.

setting up an optimization problem including the choice of the optimization variable, x, the inequality constraints $\{g_i(x)\}$, the equality constraints $\{h_i(x)\}$ and the allowable set X. For example, a given set of constraints could be described using $\{g_i(x)\}$ and/or $\{h_i(x)\}$ or using the set X.

Different names are associated with different forms of the optimization problem, e.g.,

LP: when f, g_i, and h_i are linear functions of x.
QP: when f, is quadratic in x and g_i and h_i are linear.
MILP: when f, g_i, and h_i are linear and x can take integer and/or real values.
MINLP: when f, g_i, and h_i are (possibly) nonlinear and x can take integer and/or real values.

An important class of problems arises when some of the variables defining the problem are uncertain. In this case, probabilistic descriptions can be used for the uncertain variables. This leads to the class of problems known an "stochastic optimization" for which special techniques are available.

Engineering design problems typically fall into the MINLP class since one is often required to make architectural choices (i.e., select from a discrete set of alternatives) and then to specify a number of continuous variables (i.e., select variable component sizes).

In all optimization problems, convexity is an important issue. Indeed, modern theory shows that suitably structured convex optimization problems (including most linear and quadratic problems) are tractable. Indeed, Interior Point Algorithms utilizing Newton type iterations often find an acceptable solution with relatively few iterations. Clearly this is a major "selling point" for using optimization strategies.

2.2 Modelling

One of the key challenges in applying optimization to engineering problems is that of modelling. Indeed, the judicious choice of variables, constraints and cost function is a crucial, and often formidable, step. In recent years, the key step of modelling has been facilitated by the advent of interface software including GAMS (Brooke et al., 1998), CUTE (Bongartz et al., 1995), and AMPL (Fourer et al., 1993).

Naturally, there exists a strong interplay between the problem formulation (i.e. modelling) and the solution method. We advocate that it is often a good idea to start simple and add extra features as one gains confidence in both the problem and its solution. We also advocate the judicious use of "toy problems" to gain a feel for the question and its solution. (A "toy problem" is one with just a few variables but which captures important features of the problem under study.) Toy problems can give valuable insight and avoid wasted effort dealing with the thousands (or even millions) of variables that typically occur in real world design problems.

2.3 Solution Methods

Once one has modelled the problem, then the next step is to seek a suitable solution strategy. The solution strategy is a function of the nature of the variables, cost function and constraints. A brief overview is given below. Further details may be found in, for example, Biegler and Grossmann (2004).

2.3.1. Continuous Variable Optimization When only continuous variables appear in an optimization problem then this is a substantial advantage. The simplest class of problems of this type are LP problems for which the standard algorithm is the simplex algorithm. Another simple class is when the cost is quadratic and the constraints linear. These problems are termed quadratic programs (QP) and can typically be solved in a finite number of steps.

For more general nonlinear programs (NLP) one must rely upon iterative solvers such as those that utilize Newton type steps to satisfy necessary conditions of optimality. There is a wide variety of algorithms available (active set, interior point, etc.), see for example the package TOMLAB® (Holmströn et al., 2006). Having a large number of constraints can be problematic. However, problems with a large number of nonlinear inequality constraints may often have few active constraints. This issue can be addressed in certain cases. For example, Polak et al. (2007) has described specially designed algorithms for on-line MPC (see Section 3.1) which selects a small number of active constraints using outer approximations.

Another important class of algorithm (termed Simulated Annealing (Kirkpatrick et al., 1983)) is motivated by a thermodynamic cooling analogy. In these algorithms a "temperature" parameter is used to adjust the probability of accepting new points even though they do not improve the cost function. Thus, one can "jump out of" local minima.

Another class of algorithms (termed Genetic Algorithms (GA) (Goldberg, 1989)) is motivated by genetics. In this type of algorithm, new trial solutions are generated by crossover (i.e. randomly swapping elements in given vector trial solutions) or by mutation (i.e. randomly adding components to elements of trial solutions).

2.3.2. Discrete Variable Optimization When discrete choices appear in optimization problems then this leads to inherent difficulties. For example, these problems are certainly non-convex. Branch and Bound methods (Biegler and Grossmann, 2004) are commonly deployed to solve these problems. A tree search is typically used such that the integer space is successively partitioned into relaxed problems (i.e. where the integer constraint is replaced by an interval constraint) at each node of the tree. Simplifications are introduced by preprocessing, e.g. by eliminating variables or by removing certain constraints.

2.3.3. Stochastic Optimization Uncertainty is often a key issue in optimization. In this case, there are different modelling approaches that one can use. For example, one can deploy deterministic uncertainly bounds or probabilistic descriptions. Either approach will tend to add significant complexity to the underlying optimization problem. A simple strategy is to base the solution only on the nominal (or expected) value for the uncertain variable. However, this can be misleading since the resultant solution will, almost certainly, be non-optimal, or, indeed, even infeasible (i.e., violate key constraints), under certain reasonable realizations for the uncertain variables. If there are hard constraints then one must examine each possible realization to ensure constraints are not violated for this realization.

2.3.4. Receding Horizon Optimization Many optimization problems have a temporal character (i.e. one wants to optimize over some future time horizon). Also, it is often true that actions planned for the far future have a diminishing effect on what is the best action now. In this case, it makes sense to reduce the time window to capture the horizon over which the current decision has its greatest impact. Having carried out the optimization over this restricted horizon, one can lock in the current action and then move the window forward starting at the next time step. The set of decisions designed with one horizon becomes a good initial guess for the decisions over the next displaced horizon. Surprisingly, it turns out that it is sometimes possible to use very short planning horizons. For example, it is known that horizon one optimization in control correspond to widely used anti-windup strategies to deal with actuator saturation. The latter are very simple but are known to give excellent performance in many cases (Goodwin et al., 2005b).

The idea of receding horizon optimization opens up many other possibilities for simplifying computations. For example, within a given optimization window, one need not use uniform time discretization. Based on the idea that future actions have only a second order effect on the best action to take now, then one can use time steps that grow larger towards the end of the interval. If one locks in the current action and moves the horizon forward a small time step followed by another non-uniform quantization of time, then one can design a finely quantized policy by solving a (rolling horizon) sequence of simpler problems.

2.3.5. Optimization Software There exists a substantial body of software for solving optimization problems; some commercial, some freeware. A good first step for practitioners looking for optimization software is the NEOS server (www-neos.mcs.anl.gov). Control engineers often first try the optimization routines available in Matlab®. These are generic routines and typically work well at least on simple problems. However, they are not tailored to capitalize on the structure inherent in specific problem classes. In some cases it makes sense for users to code their own software. However, it is usually preferable to use well-tested and well-designed software rather than trying to code ad-hoc algorithms oneself.

2.3.6. Caveats A user of optimization needs to be cautious of the blind application of nonlinear optimization methods. For example, typical algorithms are iterative in nature and thus there always exists the possibility of non-optimality of solutions due to bad initialization or premature termination. Also, one can easily get trapped in a local (as opposed to global) optimum. Special precautions are necessary to safeguard against these possibilities.

3. SOME OPTIMIZATION PROBLEMS ARISING IN MINING, METAL AND MINERAL PROCESSING

In this section we briefly review several classes of optimization problems that frequently arise in mining, metal and mineral processing.

3.1 Constrained Control

One of the best known, and most successful applications of optimization in mining, metal and mineral processing has been constrained Model Predictive Control (MPC) (Camacho and Bordons, 1999; Maciejowski, 2002; Borrelli, 2003; Rossiter, 2003; Goodwin et al., 2005b). Many thousands of successful applications have been reported and hundreds of vendors sell general tools for carrying out the required computations (Qin and Badgwell, 1997). For example, the PACTmpc® software sold by Matrikon provides a complete solution ranging from estimating models from closed loop data to implementing controllers that account for both input and state constraints. Similar packages are sold by other companies.

A core idea used in MPC is that of receding horizon optimization (see Section 2.3.4). This idea can be summarized as follows for control problems:

(1) At time i and for the current state x_i, solve an optimal control problem over a fixed future interval, say $[i, i+N-1]$, taking into account the *current* and *future* constraints.
(2) Apply only the first step in the resulting optimal control sequence.
(3) Measure the state reached at time $i+1$.
(4) Repeat the fixed horizon optimisation at time $i+1$ over the future interval $[i+1, i+N]$, starting from the (now) current state x_{i+1}.

In particular, for time-invariant systems and functions in the optimization problem, we can set $i = 0$ in the formulation of the open loop control problem without loss of generality. Then at the current time, and for the current state x, we solve:

$$\mathcal{P}_N(x): \qquad V_N^{\text{OPT}}(x) \triangleq \min V_N(\{x_k\}, \{u_k\}), \quad (2)$$

subject to the equations describing the system, e.g.,

$$x_{k+1} = f(x_k, u_k) \quad \text{for } k = 0, \ldots, N-1, \quad (3)$$
$$x_0 = x, \quad (4)$$

together with

$$\text{input constraints: } u_k \in \mathbb{U} \quad \text{for } k = 0, \ldots, N-1, \quad (5)$$
$$\text{state constraints: } x_k \in \mathbb{X} \quad \text{for } k = 0, \ldots, N, \quad (6)$$
$$\text{terminal constraints: } x_N \in \mathbb{X}_f \subset \mathbb{X}. \quad (7)$$

The cost function typically takes the following form:

$$V_N(\{x_k\}, \{u_k\}) \triangleq F(x_N) + \sum_{k=0}^{N-1} L(x_k, u_k), \quad (8)$$

where $\{x_k\}$, $x_k \in \mathbb{R}^n$, $\{u_k\}$, $u_k \in \mathbb{R}^m$, denote the state and control sequences $\{x_0, \ldots, x_N\}$ and $\{u_0, \ldots, u_{N-1}\}$, respectively, and $\mathbb{U} \subset \mathbb{R}^m$, $\mathbb{X} \subset \mathbb{R}^n$, and $\mathbb{X}_f \subset \mathbb{R}^n$ are constraint sets. All sequences $\{u_0, \ldots, u_{N-1}\}$ and $\{x_0, \ldots, x_N\}$ satisfying the constraints (3)–(7) are called *feasible* sequences. A pair of feasible sequences $\{u_0, \ldots, u_{N-1}\}$ and $\{x_0, \ldots, x_N\}$ constitute a *feasible solution* of (2)–(8). The functions F and L in the objective function (8) are the *terminal state weighting* and the *per-stage weighting*, respectively. Typical choices for the weighting functions F and L are quadratic functions of the form $F(x) = x^{\mathrm{T}}Px$ and $L(x, u) = x^{\mathrm{T}}Qx + u^{\mathrm{T}}Ru$, where $P = P^{\mathrm{T}} \geq 0$, $Q = Q^{\mathrm{T}} \geq 0$ and $R = R^{\mathrm{T}} > 0$. More generally, one could use functions of the form $F(x) = \|Px\|_p$ and $L(x, u) = \|Qx\|_p + \|Ru\|_p$, where $\|y\|_p$ with $p = 1, 2, \ldots, \infty$, is the p-norm of the vector y.

If we denote the minimising control sequence, which is a function of the current state x_i, by

$$\mathscr{U}_{x_i}^{\text{OPT}} \triangleq \{u_0^{\text{OPT}}, u_1^{\text{OPT}}, \ldots, u_{N-1}^{\text{OPT}}\}; \quad (9)$$

then the control applied to the plant at time i is the first element of this sequence, that is,

$$u_i = u_0^{\text{OPT}}. \quad (10)$$

Time is then stepped forward one instant, and the above procedure is repeated for another N-step-ahead optimisation horizon. The first element of the new N-step input sequence is then applied. The above procedure is repeated endlessly. We can see that one is continually looking ahead to judge the impact of current and future decisions on the future response before one "locks in" the current input by applying it to the plant.

The above receding horizon procedure *implicitly* defines a time-invariant control policy $\mathcal{K}_N : \mathbb{X} \to \mathbb{U}$ of the form

$$\mathcal{K}_N(x) = u_0^{\text{OPT}}. \quad (11)$$

If the model and objective function are time invariant, then it is clear that the same input u_0^{OPT} will result whenever the state takes the same value. That is, the receding horizon optimisation strategy is really an "alibi" for generating a particular time-invariant feedback control law. Indeed, in some cases, it makes sense to evaluate $\mathcal{K}_N(x)$ off-line and then to use a table look up for on-line use (Goodwin et al., 2005b).

An important aspect of receding horizon constrained control is that closed loop stability can often be guaranteed provided one chooses the final state weighting and final constraint set appropriately (Sznaier and Damborg, 1987, 1990; Keerthi and Gilbert, 1988; Mayne and Michalska, 1990; Rawlings and Muske, 1993; Bemporad et al., 1995; Scokaert and Rawlings, 1998; Mayne et al., 2000). The key idea is to utilize the value function of the optimisation problem as a Lyapunov function. In early work, stability analyses were restricted to the nominal case (known model and zero disturbances). More recent work has focused on guaranteeing robust stability in the presence of disturbances and model uncertainty. For example, recent work reported in Løvaas et al. (2007), shows how one can guarantee robust stability in the presence of model error and disturbances. These kinds of results give theoretical support for (and hence comfort to users of) MPC.

3.2 Constrained State Estimation

Constraints can also be important in state estimation problems. For example, it may be *a-priori* known that certain variables lie in given ranges. Clearly, it can be helpful to build this kind of *a-priori* knowledge into the associated algorithm (Michalska and Mayne, 1995; Rao et al., 2001, 2003; Goodwin et al., 2005a). This represents a natural extension of unconstrained (Jazwinski, 1970; Kalman, 1960; Bryson and Frazier, 1963; Cox, 1964) state estimation.

To illustrate the key ideas, we consider the case of a linear system satisfying the following linear Markov Model:

$$\begin{aligned} x_{k+1} &= Ax_k + Bw_k, \\ y_k &= Cx_k + v_k, \end{aligned} \tag{12}$$

where $x_k \in \mathbb{R}^n$, $w_k \in \mathbb{R}^m$, $y_k \in \mathbb{R}^r$ and $v_k \in \mathbb{R}^r$. Suppose that $\{w_k\}$, $\{v_k\}$, x_0 are i.i.d. sequences having *truncated* Gaussian distributions, where $w_k \in \Omega_1, v_k \in \Omega_2, x_0 \in \Omega_3$.

We consider estimation using data from time 1 to N and define

$$\mathbf{y}_N = \begin{bmatrix} y_1^{\mathrm{T}} & \dots & y_N^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}, \tag{13}$$

$$\mathbf{y}_N^d = \begin{bmatrix} y_1^{d^{\mathrm{T}}} & \dots & y_N^{d^{\mathrm{T}}} \end{bmatrix}^{\mathrm{T}}, \tag{14}$$

$$\mathbf{x}_N = \begin{bmatrix} x_0^{\mathrm{T}} & \dots & x_N^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}, \tag{15}$$

$$\hat{\mathbf{x}}_N = \begin{bmatrix} \hat{x}_0^{\mathrm{T}} & \dots & \hat{x}_N^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}. \tag{16}$$

where $\{y_k^d\}$ denotes the observed output sequence and $\hat{x}_k$ denotes the state estimates. When the matrix B in (12) is nonsingular, the joint probability density function for $\mathbf{y}_N$ and $\mathbf{x}_N$ satisfies

$$\begin{aligned} &p_{\mathbf{y}_N,\mathbf{x}_N}(\mathbf{y}_N = \mathbf{y}_N^d, \mathbf{x}_N = \hat{\mathbf{x}}_N) \\ &= \text{constant} \times \exp\left\{-\frac{1}{2}\sum_{k=0}^{N-1} \hat{w}_k^{\mathrm{T}} Q^{-1} \hat{w}_k\right\} \\ &\quad \times \exp\left\{-\frac{1}{2}\sum_{k=1}^{N} \hat{v}_k^{\mathrm{T}} R^{-1} \hat{v}_k\right\} \\ &\quad \times \exp\left\{-\frac{1}{2}(\hat{x}_0-\mu_0)^{\mathrm{T}} P_0^{-1}(\hat{x}_0-\mu_0)\right\}, \end{aligned} \tag{17}$$

whenever

$$\begin{aligned} \hat{w}_k &\in \Omega_1 \quad \text{for } k = 0,\dots,N-1, \\ \hat{v}_k &\in \Omega_2 \quad \text{for } k = 1,\dots,N, \\ \hat{x}_0 &\in \Omega_3, \end{aligned}$$

where

$$\begin{aligned} \hat{x}_{k+1} &= A\hat{x}_k + B\hat{w}_k \quad \text{for } k = 0,\dots,N-1, \\ \hat{v}_k &= y_k^d - C\hat{x}_k \quad \text{for } k = 1,\dots,N. \end{aligned}$$

The estimation problem is: Given the observations $\mathbf{y}_N^d = [y_1^{d^{\mathrm{T}}} \dots y_N^{d^{\mathrm{T}}}]^{\mathrm{T}}$, make some statement about the states $\mathbf{x}_N = [x_0^{\mathrm{T}} \dots x_N^{\mathrm{T}}]^{\mathrm{T}}$. From the joint probability density function (17), we can express the *a posteriori distribution* of $\mathbf{x}_N$ given $\mathbf{y}_N$ as follows:

$$p_{\mathbf{x}_N|\mathbf{y}_N}(\hat{\mathbf{x}}_N|\mathbf{y}_N^d) = \frac{p_{\mathbf{y}_N,\mathbf{x}_N}(\mathbf{y}_N^d, \hat{\mathbf{x}}_N)}{p_{\mathbf{y}_N}(\mathbf{y}_N^d)}, \tag{18}$$

where $p_{\mathbf{y}_N}(\mathbf{y}_N^d)$ is a data dependent term which does not depend on $\mathbf{x}_N$. The a posteriori distribution $p_{\mathbf{x}_N|\mathbf{y}_N}(\hat{\mathbf{x}}_N|\mathbf{y}_N^d)$ summarises "what we know about $\mathbf{x}_N$ given the observations $\mathbf{y}_N^d$." Our aim is to find the *joint a posteriori most probable* [JAPMP] state estimates $\hat{\mathbf{x}}_N = [\hat{x}_0^{\mathrm{T}} \dots \hat{x}_N^{\mathrm{T}}]^{\mathrm{T}}$ given the observations $\hat{\mathbf{y}}_N^d$; that is,

$$\hat{\mathbf{x}}_N^* \triangleq \arg\max_{\hat{\mathbf{x}}_N} \; p_{\mathbf{x}_N|\mathbf{y}_N}(\hat{\mathbf{x}}_N|\mathbf{y}_N^d). \tag{19}$$

Note that (19) is equivalent to maximising the joint probability density function, since, as noticed in (18), both functions are related by a term that does not depend on $\mathbf{x}_N$. Thus, the joint maximum a posteriori estimate is given by

$$\begin{aligned} \hat{\mathbf{x}}_N^* &\triangleq \arg\max_{\hat{\mathbf{x}}_N} \; p_{\mathbf{x}_N|\mathbf{y}_N}(\hat{\mathbf{x}}_N|\mathbf{y}_N^d) \\ &= \arg\max_{\hat{\mathbf{x}}_N} \; p_{\mathbf{y}_N,\mathbf{x}_N}(\mathbf{y}_N^d, \hat{\mathbf{x}}_N) \\ &= \arg\min_{\hat{\mathbf{x}}_N} \; -\ln p_{\mathbf{y}_N,\mathbf{x}_N}(\mathbf{y}_N^d, \hat{\mathbf{x}}_N). \end{aligned} \tag{20}$$

The preceding discussion leads to the following constrained optimisation problem.

Estimation Problem: Given the observations $\{y_1^d, \dots, y_N^d\}$ solve:

$$\mathcal{P}_e : V_N^{\mathrm{OPT}}(\mu_0, \{y_k^d\}) \triangleq \min V_N(\{\hat{x}_k\}, \{\hat{v}_k\}, \{\hat{w}_k\}), \tag{21}$$

subject to:

$$\hat{x}_{k+1} = A\hat{x}_k + B\hat{w}_k \quad \text{for } k = 0,\dots,N-1, \tag{22}$$

$$\hat{v}_k = y_k^d - C\hat{x}_k \quad \text{for } k = 1,\dots,N, \tag{23}$$

$$\hat{w}_k \in \Omega_1 \quad \text{for } k = 0,\dots,N-1, \tag{24}$$

$$\hat{v}_k \in \Omega_2 \quad \text{for } k = 1,\dots,N, \tag{25}$$

$$\hat{x}_0 \in \Omega_3, \tag{26}$$

where

$$\begin{aligned} &V_N(\{\hat{x}_k\}, \{\hat{v}_k\}, \{\hat{w}_k\}) \\ &\triangleq \frac{1}{2}(\hat{x}_0-\mu_0)^{\mathrm{T}} P_0^{-1}(\hat{x}_0-\mu_0) \\ &\quad + \frac{1}{2}\sum_{k=0}^{N-1} \hat{w}_k^{\mathrm{T}} Q^{-1} \hat{w}_k + \frac{1}{2}\sum_{k=1}^{N} \hat{v}_k^{\mathrm{T}} R^{-1} \hat{v}_k. \end{aligned} \tag{27}$$

The above problem can be seen to be a standard QP problem and is closely related to the constrained control problem described in section 3.1.

Further to the above developments one can also utilize rolling horizon ideas for constrained state estimation. For example one can use the observation interval $(1, N)$ to estimate the state $x(N)$. Then one can move the estimation interval forward one step and consider data on the interval $(2, N+1)$ to estimate $x(N+1)$ and so on. When doing this, a key factor becomes the "entry cost" i.e. the last term in equation (17). For example, one could remember the state estimate for $x(0)$ obtained using data from $(N-1)$ to 0 and use this for μ_0. In principle one also needs P_0. One option is to use an unconstrained Kalman filter to supply P_0.

3.3 Scheduling Problems

Another class of problems of interest in mining, mineral and metal processing is that of scheduling. These problems typically involve discrete variables and thus fall into the MILP or MINLP categories.

To illustrate, let $x_i^t (i = 1, \ldots, N)$ denote a variable which takes the value '1' if a resource is used in periods 1 to t and is zero otherwise. Let c_i^t denote the value returned by utilizing resource x_i at time t. Then, over a horizon of T, the net value returned is

$$f(x) = \sum_{i=1}^{N} c_i^T x_i^T + \sum_{t=2}^{T} \sum_{i=1}^{N} \left(c_i^{t-1} - c_i^t\right) x_i^{t-1} \quad (28)$$

Also, in typical problems, there will be constraints on the x_i's of the form

$$g_i(x) = \sum_{t=2}^{T} \sum_{i=1}^{N} d_i^{t-1} x_i^{t-1} \leq 0; j = 1, \ldots, M_1 \quad (29)$$

$$h_\ell(x) = \sum_{t=2}^{T} \sum_{i=1}^{N} e_i^{t-1} x_i^{t-1} \leq 0; \ell = 1, \ldots, M_2 \quad (30)$$

The fact that $\{x_i^t\}$ can take only integer values renders this an MILP.

4. DUALITY

An important concept in optimization is that of duality. Duality often gives insight into optimization problems and is exploited in some optimization algorithms. In this section, we will briefly review the idea of duality and illustrate it by describing the symmetric dual relationship that exists between constrained control and estimation.

4.1 Brief Introduction to Lagrangian Duality

Consider the primal problem as in (1).

The *Lagrangian dual problem* is then defined as the following nonlinear programming problem:

Lagrangian Dual Problem D

$$\begin{aligned} &\text{maximise } \theta(\alpha, \beta) \\ &\text{subject to :} \\ &\alpha \geq 0 \end{aligned} \quad (31)$$

where $\theta(\alpha, \beta)$ is the solution of the *Lagrangian dual subproblem* defined by

$$\begin{aligned} \theta(\alpha, \beta) = &\inf\{f(x) + \sum_{i=1}^{m} \alpha_i g_i(x) \\ &+ \sum_{i=1}^{\ell} \beta_i h_i(x) : x \in X\} \end{aligned} \quad (32)$$

In the dual problem (31)-(32), the vectors α_i and β_i have, as their components, the Lagrange multipliers α_i for $i = 1, \ldots, m$, and β_i for $i = 1, \ldots, \ell$. Note that the Lagrange multipliers α_i, corresponding to the inequality constraints $g_i(x) \leq 0$, are restricted to be nonnegative, whereas the Lagrange multipliers β_i, corresponding to the equality constraints $h_i(x) = 0$, are unrestricted in sign.

Given the primal P (1), several Lagrangian dual problems D of the form of (31)-(32) can be devised, depending on which constraints are handled as $g_i(x) \leq 0$ and $h_i(x) = 0$, and which constraints are handled by the set X. Hence, an appropriate selection of the set X must be made, depending on the nature of the problem and the goal of formulating or solving the dual problem D.

An interesting geometric interpretation of the dual problem can be made by considering a simpler problem with only one inequality constraint and no equality constraint. Consider the following primal problem P:

Primal Problem P

$$\begin{aligned} &\text{Minimize } f(x) \\ &\text{subject to:} \\ &g(x) \leq 0 \\ &x \in X \end{aligned} \quad (33)$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$, and define the following set in $\mathbb{R}^2$:

$$G = \{(y, z) : y = g(x), z = f(x) \text{ for some } x \in X\} \quad (34)$$

that is, G is the image of X under the (g, f) map. Figure 1 shows an example of the set G. Then, the primal problem consists of finding a point in G with $y \leq 0$ that has minimum ordinate z. Obviously this point in Figure 1 is $(\bar{y}, \bar{z})$.

Now, consider the Lagrangian dual problem D:

$$\begin{aligned} &\text{maximise } \theta(\alpha) \\ &\text{subject to :} \\ &\alpha \geq 0 \end{aligned} \quad (35)$$

where

$$\theta(\alpha) = \inf\{f(x) + \alpha g(x) : x \in X\} \quad (36)$$

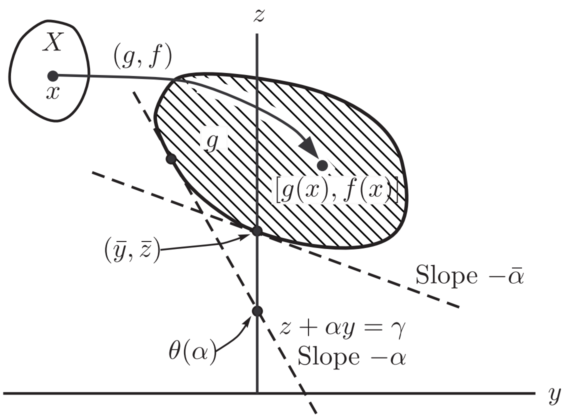

Fig. 1. Geometric interpretation of Lagrangian duality: case with no duality gap.

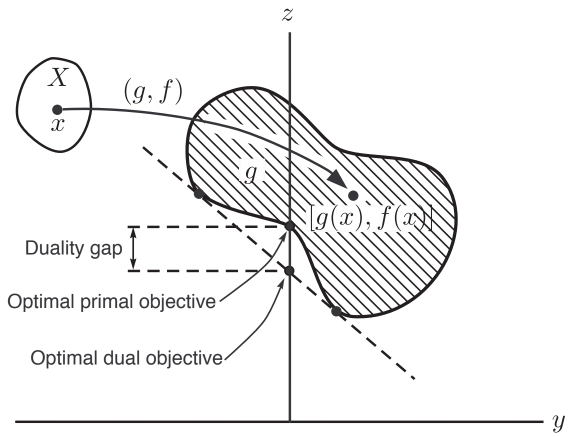

Fig. 2. Geometric interpretation of Lagrangian duality: case with duality gap.

Given $\alpha \geq 0$, problem (36) is equivalent to minimising $z + \alpha y$ over points (y, z) in G. Note that $z + \alpha y = \gamma$ is the equation of a straight line with slope $-\alpha$ that intercepts the z-axis at γ. Thus, in order to minimise $z+\alpha y$ over G we need to move the line $z+\alpha y = \gamma$ parallel to itself as far down as possible, whilst it remains in contact with G. The last intercept on the z-axis thus obtained is the value of $\theta(\alpha)$ corresponding to the given $\alpha \geq 0$, as shown in Figure 1. Finally, to solve the dual problem (35), we have to find the line with slope $-\alpha(\alpha \geq 0)$ such that the last intercept on the z-axis, $\theta(\alpha)$, is maximal. Such a line is shown in Figure 1. It has slope $-\alpha$ and supports the set G (see equation (34)) at the point $(\bar{y}, \bar{z})$. Thus, the solution to the dual problem (35) is $\bar{\alpha}$. It can be seen that, in the example illustrated in Figure 1, the optimal primal and dual objective values are equal. In such cases, it is said that there is no *duality gap*. More generally, it is readily seen that any feasible solution to the dual problem always provides a lower bound for the objective functions for any feasible solution to the primal problem. This is established below:

Theorem 1. (Weak Duality Theorem). Consider the primal problem P given by (1) and its Lagrangian dual problem D given by (31), (32). Let x be a feasible solution to P; that is, $x \in X, g(x) \leq 0$, and $h(x) = 0$. Also, let (α, β) be a feasible solution to D; that is, $\alpha \geq 0$. Then:

$$f(x) \geq \theta(\alpha, \beta) \tag{37}$$

PROOF. We use the definition of θ given in (32), and the facts that $x \in X, \alpha \geq 0, g(x) \leq 0$ and $h(x) = 0$. We then have

$$\begin{aligned} \theta(\alpha, \beta) &= \inf\{f(\bar{x}) + \alpha^T g(\bar{x}) + \beta^T h(\bar{x}) : \bar{x} \in X\} \\ &\leq f(x) + \alpha^T g(x) + \beta^T h(x) \leq f(x) \end{aligned} \tag{38}$$

and the result follows.

Corollary 2.

$$\begin{aligned} &\inf\{f(x) : x \in X, g(x) \leq 0, h(x) = 0\} \\ &\geq \sup\{\theta(\alpha, \beta) : \alpha \geq 0\} \end{aligned} \tag{39}$$

Note from (39) that the optimal objective value of the primal problem is greater than or equal to the optimal objective value of the dual problem. If (39) holds as a *strict* inequality, then it is said that there exists a *duality gap*. Figure 2 shows an example for the primal and dual problems defined in (33) and (35)–(36), respectively. Notice that, in the case shown in the figure, there exists a duality gap. We see, by comparing Figure 2 with Figure 1, that the presence of a duality gap is due to the nonconvexity of the set G. Indeed if some suitable convexity conditions are satisfied, then there is no duality gap between the primal and dual optimization problems.

4.2 Duality between Constrained Estimation and Control

We have argued in Section 3, that two major application areas of optimization in mining, metal and mineral processing are constrained control (see Section 3.1) and constrained state estimation (see Section 3.2). It is thus of considerable interest to further study these two classes of problems. It is obvious from the brief description of these problems given in Section 3 that they are closely related. What is perhaps less obvious is that the two problems actually bear a beautifully symmetric dual relationship. This result has only recently been established (see Goodwin et al. (2005a)). To outline the key ideas, we refer to the finite horizon constrained state estimation problem described by equations (21) to (27).

For simplicity, we consider only the case where ω_k is constrained to a convex set Ω. However, similar results also hold when v_k and $\hat{x}_k$ are also constrained (Müller et al., 2006). The following result establishes (Lagrangian) duality between the constrained estimation problem $\mathcal{P}_e$ and a particular optimal control problem with projected variables.

Theorem 3. (Dual Problem). Assume Ω is a non-empty closed convex set. Given the primal constrained fixed horizon estimation problem $\mathcal{P}_e$ defined by equations (21)–(27), the Lagrangian dual problem is

$$\mathcal{D}_e: \quad \phi^{\text{OPT}}(\mu_0, \{y_k^d\}) \triangleq \min_{\lambda_k, u_k} \phi(\{\lambda_k\}, \{u_k\}), \tag{40}$$

subject to:

$$\lambda_{k-1} = A^{\text{T}}\lambda_k + C^{\text{T}}u_k \quad \text{for } k = 1, \cdots, N, \tag{41}$$

$$\lambda_N = 0, \tag{42}$$

$$\zeta_k = B^{\text{T}}\lambda_k \quad \text{for } k = 0, \cdots, N-1, \tag{43}$$

$$\bar{\zeta}_k = Q^{-1/2}\Pi_{\tilde{\Omega}}Q^{1/2}\zeta_k \quad \text{for } k = 0, \cdots, N-1. \tag{44}$$

In (40) the objective function is:

$$\begin{aligned}
&\phi(\{\lambda_k\}, \{u_k\}) \\
&\triangleq \frac{1}{2}(A^{\text{T}}\lambda_0 + P_0^{-1}\mu_0)^{\text{T}}P_0(A^{\text{T}}\lambda_0 + P_0^{-1}\mu_0) \\
&\quad + \frac{1}{2}\sum_{k=1}^{N}(u_k - R^{-1}y_k^d)^{\text{T}}R(u_k - R^{-1}y_k^d) \\
&\quad + \sum_{k=0}^{N-1}\left[\frac{1}{2}\bar{\zeta}_k^{\text{T}}Q\bar{\zeta}_k + (\zeta_k - \bar{\zeta}_k)^{\text{T}}Q\bar{\zeta}_k\right] + \gamma
\end{aligned} \tag{45}$$

where γ is the constant term given by

$$\gamma \triangleq -\frac{1}{2}\mu_0^{\text{T}}P_0^{-1}\mu_0 - \frac{1}{2}\sum_{k=1}^{N}(y_k^d)^{\text{T}}R^{-1}y_k^d. \tag{46}$$

In (44), $\Pi_{\tilde{\Omega}}$ denotes the minimum Euclidean distance projection onto $\tilde{\Omega} \triangleq \{z : Q^{1/2}z \in \Omega\}$, that is,

$$\begin{aligned}
\Pi_{\tilde{\Omega}} : \mathbb{R}^m &\longrightarrow \tilde{\Omega} \\
s &\longmapsto \bar{s} = \Pi_{\tilde{\Omega}}s \triangleq \arg\min_{z \in \tilde{\Omega}} \|z - s\|.
\end{aligned} \tag{47}$$

Moreover, there is no duality gap, that is, the minimum achieved in (21) is equal to minus the minimum achieved in (40).

Proof: See Goodwin et al. (2005a).

We can think of (41)–(44) as the state equations of a system (running in reverse time) with input u_k and output ξ_k. The above theorem then shows that the dual of the estimation problem with constraints on the system inputs (the process noise ω_k) is a control problem having *projected outputs* in the objective function. A striking symmetry between the two problems is revealed (see Goodwin et al. (2005a)) when one realizes that the estimation problem with constraint $\omega_k \in \Omega$ can also be described by an equivalent optimization problem (as below) based on the use of *projected inputs*:

Equivalent Estimation Problem
(with Projected Variables)

$$\begin{aligned}
\mathcal{P}_e': \quad \min_{\hat{x}_k, \hat{v}_k, \hat{w}_k} &\left\{\frac{1}{2}(\hat{x}_0 - \mu_0)^{\text{T}}P_0^{-1}(\hat{x}_0 - \mu_0)\right. \\
&+ \frac{1}{2}\sum_{k=1}^{N}\hat{v}_k^{\text{T}}R^{-1}\hat{v}_k + \sum_{k=0}^{N-1}\Big[\frac{1}{2}\bar{w}_k^{\text{T}}Q^{-1}\bar{w}_k \\
&\left. + (\hat{w}_k - \bar{w}_k)^{\text{T}}Q^{-1}\bar{w}_k\Big]\right\},
\end{aligned}$$

subject to:

$$\hat{x}_{k+1} = A\hat{x}_k + B\bar{w}_k \quad \text{for } k = 0, \cdots, N-1,$$

$$\hat{v}_k = y_k^d - C\hat{x}_k \quad \text{for } k = 1, \cdots, N,$$

$$\bar{w}_k = Q^{1/2}\Pi_{\tilde{\Omega}}Q^{-1/2}\hat{w}_k \quad \text{for } k = 0, \ldots, N-1.$$

Comparison of Problem $\mathcal{D}_e$ and $\mathcal{P}_e'$ reveals the remarkable symmetry that exists between the two problems.

5. ILLUSTRATIVE CASE STUDIES

In this section we will briefly overview three case studies drawn from the mining, mineral and metal processing area. We will utilize these case studies to highlight some of the general points made in previous sections. In particular, we will discuss the interplay between modelling and solution methodology. Various other observations arising from the case studies will be highlighted.

5.1 Air knife Control in Continuous Galvanizing Lines

This case study is an example of continuous variable optimization. Earlier attempts to solve this problem were based on conventional control methods. However, the fact that the problem is inherently nonlinear with hard constraints makes optimization a natural choice. Further details may be found in Rojas et al. (2006).

Continuous strip hot-dip galvanizing lines represent a significant and complex industrial control application (Edwards et al., 1976; Jacobs, 1995). Some of the main difficulties associated with these types of processes is that they are multivariable with a large number of inputs and outputs, (30×30 transfer functions are typical) they are nonlinear and are required to operate subject to tight input constraints and at fast sampling rates. This

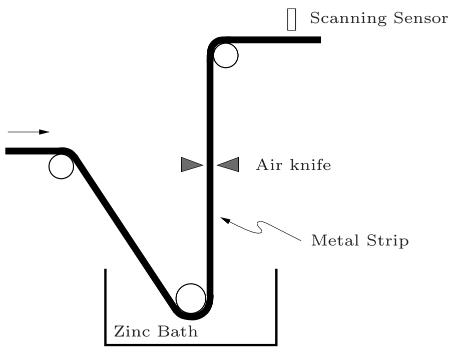

Fig. 3. General galvanizing line process scheme.

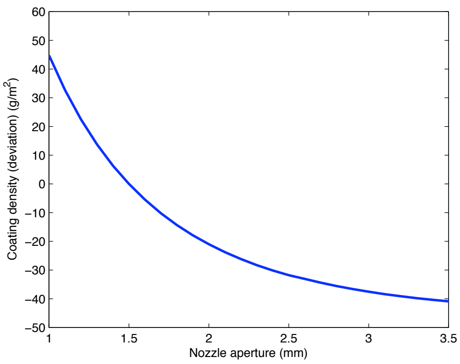

Fig. 4. Gain variation from nozzle aperture to coating density deviation on the strip.

suggests that nonlinear MPC could be a suitable strategy.

A simplified diagram of a continuous strip hot-dip galvanizing line is depicted in Figure 3. To be specific, we consider a metal strip with a width of 0.75 m. The air knife comprises a set of 10 air nozzles spaced at equal intervals of 10 cm. The scanning sensor that measures the coating density on the metal strip is positioned downstream from the air knife and the spatial separation between the measurements is 10 mm.

An interesting aspect of the problem is that the gain relationship between the aperture of one nozzle on the air knife and the coating mass density on the strip is nonlinear. This nonlinear gain is determined by several variables including the strip speed, the separation between the air knife and the strip, the air jet pressure, etc. The nonlinear gain is usually approximately known (say within an error of less than 5%). A typical curve is shown in Figure 4.

In addition, the pressurised air blown by each nozzle not only affects the coating mass directly underneath the nozzle but it also affects the coating mass density at nearby positions on the strip. This phenomenon is typical of systems with a spatial distribution similar to that exhibited by the galvanizing line process. Figure 5 shows a typical spatial profile of the gain associated with a

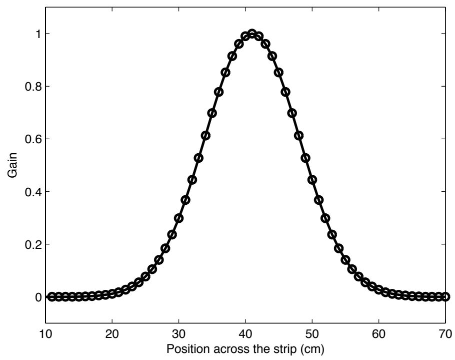

Fig. 5. Spatial gain profile generated by one nozzle.

single nozzle. Physical limitations impose a series of constraints on the nozzle apertures. Typical constraints include:

(1) input amplitude constraint:

$$0.5 \leq u_i \leq 4 \quad (mm) \quad i = 1, \ldots, 10 \tag{48}$$

(2) first difference constraint:

$$|u_{i+1} - u_i| \leq 1 \quad (mm) \quad i = 1, \ldots, 9 \tag{49}$$

(3) second difference constraint:

$$|u_{i+2} - 2u_{i+1} + u_i| \leq 1.5 \quad (mm) \quad i = 1, \ldots, 8 \tag{50}$$

The air pressure in the system is regulated by a separate control loop. To minimise the interaction between the coating mass control loop and the air pressure control loop, the following equality constraint is imposed on the air nozzle aperture:

(4) constant average nozzle opening:

$$\frac{1}{10} \sum_{i=1}^{10} u_i = r_n \tag{51}$$

where r_n is the desired value of the average nozzle opening.

Similarly, the average coating density across the strip is mainly controlled by the separation between the air knife and the strip. Thus, in order to prevent the coating mass control loop from modifying the average coating density across the strip, the following equality constraint is imposed:

(5) constant average coating density:

$$\frac{1}{N} \sum_{i=1}^{N} y_i = r_c \tag{52}$$

where N is the number of components in the measurement vector $\mathbf{y}$ and r_c is the desired average coating density across the strip.

A mathematical model relating the nozzle apertures with the coating density on the strip can be derived directly from the available information on the spatial interaction shown in Figure 5 and

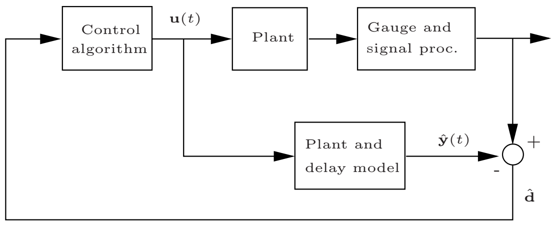

Fig. 6. Control architecture for the air knife control problem.

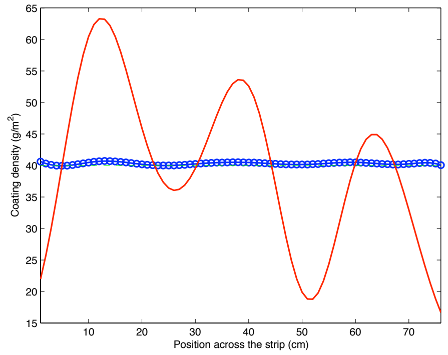

Fig. 7. Coating density profile across the strip (circles and continuous line) and applied output disturbance profile (continuous line).

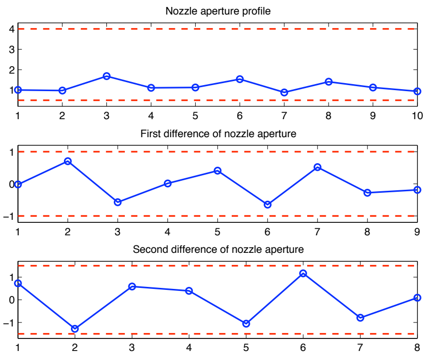

Fig. 8. Nozzle opening profiles. The horizontal dashed lines are the constraint limits.

the nonlinear gain relation shown in Figure 4. In particular, let $\mathbf{u}$ be a column vector that contains the aperture for each nozzle on the air knife and let $\mathbf{y}$ be a column vector containing the coating density measurements across the strip. Then, we can write:

$$\mathbf{y} = G\, g\{\mathbf{u}\} + \mathbf{d} \tag{53}$$

where $g\{\cdot\}$ is the nonlinear gain and $\mathbf{d}$ is an unknown output disturbance. In addition, G is a constant matrix that models the combined effect of all nozzles on the coating density profile. Each column of the matrix G is a shifted version of the profile shown in Figure 5. Notice that the model in (53) does not include dynamics, since any longitudinal or actuator dynamics are assumed to be faster than the controller sampling period of 0.1 s. Potential picketing difficulties (i.e. adjacent actuators "fighting each other") exist in these types of problems. There are, in general, several ways of dealing with picketing. We rely on the first and second difference constraints in (49) and (50) to avoid large changes in the profile of the nozzle apertures to mitigate these effects.

A difficulty is that the system has a complex delay structure due to the nature of the scanning gauge that provides the coating density measurements. At a given sampling time, the coating density measurements all have different time delays. It is thus a challenging modelling problem to deal with these delays. We adopted an internal model approach to this problem as shown in Figure 6. We observe that since the process model contains also a model of the system's delay structure we have that the signal fed-back to the controller is essentially an estimation of the output disturbance vector $\mathbf{d}$. Some on-line filtering of $\mathbf{d}$ may be desirable to isolate different components. Constrained state estimation (see Section 3.2 is a potentially useful tool in this regard. However, at the current time, we have only used simple methods in this particular application. For example, pre-filtering is used to ensure that the estimated output disturbance profile $\hat{\mathbf{d}}$ contains no constant nor tilt components. These components are removed and used as disturbance data for separate control loops.

Having modelled the problem, the next step is to specify a suitable algorithm. Our first choice was to approximate the nonlinearity by a set of piecewise linear models. Piecewise affine ideas were then used to convert the optimization problem into a set of interconnected QP problems (Borrelli, 2003). This gave a satisfactory solution save for one major issue, namely, the required computation time for each step in the rolling horizon optimization far exceeded the available sampling time of 0.1 seconds. Hence, the real time aspects of the problem were not satisfactorily addressed.

A very simple suboptimal strategy was then tried where a gradient of the cost function was evaluated and a linear search was employed in the gradient direction to minimize the cost whilst respecting the constraints. Only one gradient evaluation was conducted per time step. Of course, this is a highly suboptimal strategy. However, for this particular problem, it was found to give results that were very close to those obtained by the much more complex piecewise affine strategy.

Typical output and input profiles are shown in Figures 7 and 8, respectively. We observe that the controller has successfully compensated the output disturbance profile. Owing to the higher spatial modes included in the disturbance, we see in Figure 8 how the nozzle first and second difference profiles have larger variations. However, these variations are still inside the constraint limits. In conclusion, this provided a satisfactory solution to the practical problem.

5.2 Deterministic Optimal Mine Planning

These problems contain integer variables and thus fall into the MILP framework. We will provide a brief introduction to the problem below [see also Appiah et al. (1990); Appiah and Sturgul (1990); Ataei and Osanloo (2003); Brazil et al. (2003); Caccetta and Hill (2003); Clement and Vagenas (1994); Darwen (2001a); Denby and Schofield (1994, 1995a,b); Denby et al. (1998); Stone et al. (2004); Thomas (1996)]. The essence of the problem is as follows: One has preliminary available data on the location of an ore-body in a particular geological volume. Given the data, one would like to know 'where' and 'when' to dig so as to optimize the 'net present value'. A typical mining operation can span 15 to 20 years (or more) and hence there is a temporal aspect to these questions. Also, the optimization needs to respect a host of constraints, e.g., mining capacity in each year and slope constraints on the mine walls (to avoid collapses), mining constraints (e.g., on the order that material can be mined), mining capacity constraints, etc.

The basic idea of open cut mining can be visualised in Figure 9, which shows the ultimate pit of a typical mine, that is, the opening left in the ground after mining operations have been completed. For simplicity of exposition, we represent the potential mine by the 'box' shown in Figure 10, where the 'surface' is divided into $(J+1) \times (K+1)$ rectangles.

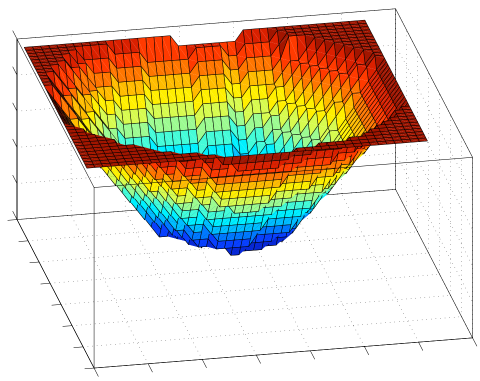

Fig. 9. Illustration of a typical mine ultimate pit.

The above mine planning problems are typically formulated as MILP problems Darwen (2001b), Denby and Schofield (1994), Denby and Schofield (1995a). One can also model the problem as a (large scale) control problem. An advantage of using the "control" model formulation is that the problem then appears more familiar to control engineers (including the authors of the paper). In the sequel, we use the term 'control' to describe those variables that can be selected (at each time step) by the optimizer.

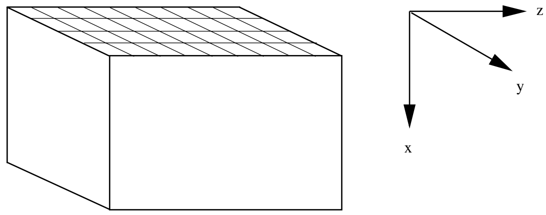

Fig. 10. Simplified representation of the mine.

Accordingly, we define the mine state as the set of pit depths at the locations of the surface. We represent the evolution of this state via a linear, discrete-time dynamic model where mining action is the control input. Specifically, we denote by $x_{jk}(t)$ the mine depth at location jk at time t. Similarly, we denote by $u_{jk}(t)$ the action to mine (or not) at time t in the location jk, $j \in \{0, \dots, J\}$, $k \in \{0, \dots, K\}$. We thus think of $u_{jk}(t)$ as an $(J+1) \times (K+1)$ input vector. A state model for the system can then be written as

$$\begin{aligned} x_{jk}(t+1) &= x_{jk}(t) + b_1 u_{jk}(t); \quad t \in \mathbb{N}_0 \\ x_{jk}(0) &= 0; \quad j \in \{0, \dots, J\}, \quad k \in \{0, \dots, K\} \end{aligned} \tag{54}$$

where b_1 is a constant that reflects the effect of one unit of mining action. The model (54) will appear very familiar to control engineers.

Constraints can be incorporated in a natural way in the state-space formulation presented. For example, note that $u_{jk}(t)$ can take either the value 1 or 0 indicating the action of mining or not at location jk at time t. Thus, $u_{jk}(t)$ is nonnegative and the model (54) readily ensures that the mine depth cannot decrease at any location. Also, slope constraints on the mining depth can be directly incorporated by means of state constraints of the form

$$|x_{\ell n}(t) - x_{jk}(t)| \leq b_2; \; t \in \mathbb{N}_0, \; |\ell - j| = 1, \; |n-k| = 1 \tag{55}$$

The mining capacity constraint can be easily handled by imposing an input constraint such that only a certain number of $u_{jk}(t)$ can be nonzero at any t. Other constraints, such as processing plant constraints, can also be modelled by introducing functions to model ore content. Finally, the state-space formulation presented here can be extended to more complex situations, such as multiple processing plants with variable capacities, multiple material stockpiles, variable material price, etc.

The value of the body of ore at different locations is typically obtained by preliminary drilling work. Using this information one can construct a value function $V_{jk}(x_{jk})$ which represents the value assigned to the material in location jk at depth x_{jk}. We also introduce a time discounting function d_t

to yield net present value and assume that the price of ore at time t is c_t. The cost function to use for mine planning, representing the net present value achieved by a given mining strategy over a planning horizon T, then takes the form

$$J := \sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{k=1}^{K} d_t c_t \, V_{jk}(x_{jk}(t)) \, u_{jk}(t-1) \quad (56)$$

Note that we multiply by $u_{jk}(t-1)$ in (56) since the value is only liberated after mining.

A key point about this problem is that it is too complex (in its raw form) to be solved in any reasonable sense. (The raw "block model" has 10^4 variables.) Hence, one must simplify the problem substantially before setting out to solve it. These simplifications amount to quantizing in both space and time. Some "art" as well as science is involved in this step. In our own work we have deployed many different approaches to solving the problem. Brief reflections are outlined below.

(1) We have made extensive use of GA's in the context of mine optimization. We have found GA's give rapid initial reduction in the cost function followed by slow final convergence. Thus, in this context, GA's seem to provide a useful way of getting good initial estimates for other solvers.
(2) We have tried several MILP solvers. In particular, we have made extensive use of CPLEX®. We have found these solvers to give good performance.
(3) We have also developed specific software ourselves based on the control orientated model. In developing our own software, we made several important observations. For example, we found that non-uniform time quantization (see Section 2.3.4) was extremely beneficial. This sped up the computations by an order of magnitude. This idea was subsequently adopted in the MILP based modelling approach since it allowed such larger problems to be solved in realistic time frames. (Typically real mine optimization problems can take several hours to solve on a fast computer.)
(4) As our confidence in the problem grew, so more challenging features were added to the model, e.g. ensuring proper access for trucks and adequate space at the bottom of each phase for mining equipment to operate. These required additional attention to modelling.

5.3 Optimal Mine Planning with Uncertainty

We recall that the model used in Section 5.2 assumed that $V_{jk}(x_{jk})$ and c_t were known functions. More realistically, however, one will not know the exact value of ore in the ground nor the future price that the ore will bring. Thus, one should, in principle at least, introduce this uncertainty into the problem. Since the deterministic problem is already very difficult to solve one needs to be rather careful how one models uncertainty so as not to render the problem intractable. Another important decision is to specify what information will be available and when this information will become available (Bertsekas, 2005). Three possible solution strategies are as follows:

Open Loop: Here one calculates the future "controls" based purely on the expected future value of uncertain variables. One then implements the control sequence blindly (irrespective of what actually happens).

Reactive: Here one calculates the future "controls" as for the open loop case. However, one only implements the first step. One then takes new measurements and updates the estimate of the current "state" and repredicts the value of the uncertain variables. One then redoes the optimization over another future horizon and implements the first step, as for MPC see section 3.1.

Closed Loop (or with recourse (Marti et al., 2004; Uryasev and Pardalos, 2001)). Here one calculates the control action based on the knowledge that *in the future* additional information will become available. This is usually a nontrivial exercise since one must effectively map all possible future "states" into control outcomes; i.e., one is designing a mapping from the information state to controls rather than a simple control sequence.

It is heuristically reasonable that the performance improves in the order open loop, reactive and closed loop. However, the difficulty of solving the problem increases in the same order. Thus, one needs to be sure that the more sophisticated strategies are truly worthwhile. This is problem dependent.

Three key issues arose in the context of applying stochastic optimization to the mining problem namely:

(i) how to formulate a closed loop solution,
(ii) how to model uncertainty and
(iii) whether reactive planning would suffice or whether one should consider a closed loop solution.

These issues are briefly addressed below.

5.3.1. Formulation of Closed Loop Strategies

To fix ideas, we will begin with a simplified problem. (Recall the comments made about "toy" problems in section 2.2.)

Consider a grossly simplified situation where the planning horizon has only 3 stages. Also, let us assume that there is only one uncertain variable and that there are 4 possible realizations ("scenarios") for this variable as shown in Figure 11.

In this figure, the price at stage 1 can only take the value $c_1 = v_1$; at stage 2, price can take the value $c_2 = v_2$, with probability α, or $c_2 = v_3$, with probability $1 - \alpha$; at stage 3, if price at stage 2 was v_2, then price can take either the value $c_3 = v_4$, with *conditional* probability β, or $c_3 = v_5$, with *conditional* probability $1 - \beta$; and similarly for the values v_6 and v_7. The price scenarios are then defined by each of the four branches of the scenario tree (for example, scenario 1 corresponds to $c_1 = v_1$, $c_2 = v_2$ and $c_3 = v_4$).

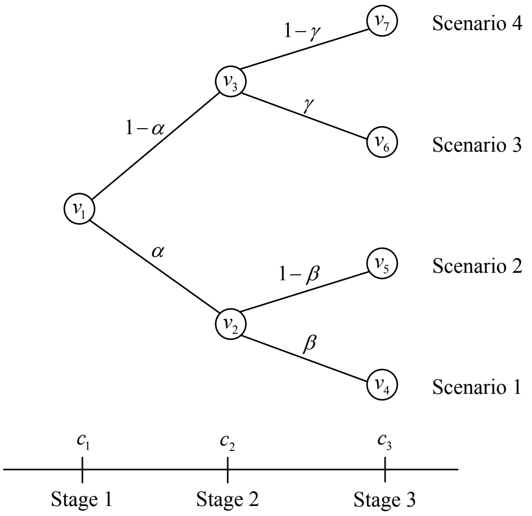

Fig. 11. An example of a scenario tree structure for closed-loop mine planning.

Next we consider the associated optimisation problems. To begin let us conceptually think of using 4 optimisation problems (each corresponding to a separate price scenario). Note that the scenarios have a tree like structure, i.e., $c_1(1) = c_1(2) = c_1(3) = c_1(4)$, $c_2(1) = c_2(2)$ and $c_2(3) = c_2(4)$.

Following the above reasoning, we introduce four corresponding optimisation problems with inputs $u(t, s)$, for $s \in \{1, 2, 3, 4\}$, where

- $u(t, s) = 1$ is the value of $u(t)$ under price scenario s.

It is important that we do not use information about price until it becomes available. This is captured by adding constraints that ensure that the optimisation variables are equal at each node of the price scenario tree. For the example above, these constraints have the form:

$$u(0,1) = u(0,2) = u(0,3) = u(0,4)$$
$$u(1,1) = u(1,2) \tag{57}$$
$$u(1,3) = u(1,4) \tag{58}$$

We see from the above "toy" example that one needs to, in effect, solve the optimization problem for all possible realizations of the uncertain variables. Obviously, in a complex problem, such as mine planning, one must restrict attention to a few "representative" realizations of the uncertain variables to make the overall problem tractable. This leads to the problem of scenario generation as discussed in the next section.

5.3.2. Scenario Generation Following the comments made in the previous section, we see that it is highly desirable that one approximate the uncertain variables by a small set of "representative" realizations. For example, if we consider the price of copper over a 10 year horizon, then this is potentially a real random variable in $\mathbb{R}^{10}$. Thus, one needs to think rather carefully how one will model the variable in the context of stochastic optimization. We have studied two alternatives.

(i) Monte Carlo's simulation: Here one simply draws realizations from the underlying probability distribution function for the uncertain variable. Unfortunately in the context of mine planning we have found that one needs to draw several thousand realizations to obtain a representative set in the sense that drawing another set of realizations of the same cardinality gives an answer to the overall problem within the required accuracy. Noting that a realistic deterministic mine planning problem can take several hours to solve, then dealing with 1000 realizations of the uncertain variable, potentially extends the solution time into many months which is unrealistic.

(ii) Deterministic scenario generation: In view of the time constraints outlined above, there is strong motivation to more carefully choose the "representative" realizations of the random variable than simply "tossing a coin". This problem is often called "scenario generation".

We refer the reader to the literature on the topic of scenario generation available in the mathematical finance and operations research fields (Beltratti et al., 1999; Dupačová, 1996; Dupačová et al., 2000; Høyland and Wallace, 2001; Keefer, 1994; Miller and Rice, 1983; Mulvey and Vladimirou, 1992; Pflug, 2001a, 1996, 2001b; Takriti et al., 1996; Yu et al., 2003).

The choice of best scenarios is a difficult problem. Indeed, it could be argued that the ultimate test of whether a given set of scenarios is "good" is to try them on the real problem. However, one is usually only motivated to consider scenario design when the real problem is very complex. Thus using the real problem as the "test bed" is usually not a feasible option. One is then forced to simplify the question of choosing scenarios to a related approximation problem. To illustrate, we refer to an algorithm for scenario generation which is related to "code book" design in signal processing.

Let $X \in \Omega$ be a random variable and say that we fix the cardinality of the set of scenarios as K. Denote the scenarios as $y_1, \ldots, y_k$. Then a possible strategy for designing scenarios is to choose $y_1, \ldots, y_k$ such that each possible realization of X is close to at least one scenario. For example, one could choose and optimal set of scenarios, $\{y_i^o\} \triangleq \{y_1^o, \ldots, y_k^o\}$, by solving the following approximation problem:

$$\{y_i^o\} = \arg\min_{\{y_i\}} \sum_{\ell=1}^{k} \underset{X \in V_\ell}{E} \left[\|X - y_\ell\|^2 \right], \qquad (59)$$

where

$$V_\ell \triangleq \{x \in \Omega : \|x - y_\ell\|^2 \leq \|x - y_j\|^2 \text{ for } j \in \{1, \ldots, k\}, j \neq \ell\} \setminus \{x \in \Omega : \exists j < \ell \text{ such that } \|x - y_\ell\|^2 = \|x - y_j\|^2 \}.$$

By way of illustration, we refer to the problem of generating scenarios for the future price of copper. (Clearly this is motivated by mine planning problems.)

The USA copper price over the years 1965-2005 is shown in Figure 12.

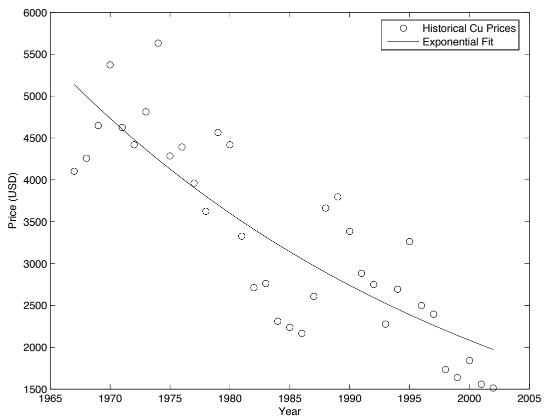

Fig. 12. USA Copper price (normalised by the CPI) and its exponential trend.

We will choose 125 scenarios and we will initially assume that the year is 1993. This will give us 9 further years (from 1994 to 2002) in which we can compare the generated scenarios with the actual price changes. (Of course, in practice, the scenarios are used to describe future uncertain variables.)

Figure 13 shows a set of 125 prices scenarios designed by using a criterion closely related to (59). Also shown is the true copper price over the years 1994 to 2002 and the best fitting (in a mean-square sense) scenario compared with the real copper price.

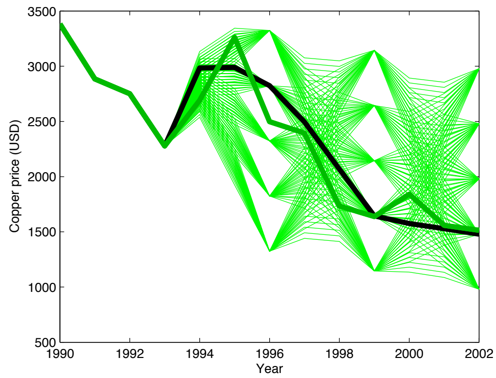

Fig. 13. Scenarios (thin lines) for the copper price, from 1994 to 2002, with the actual copper price (thick light line) and the closest fitted scenario in a mean-square sense (thick dark line).

The scenario tree shown in Figure 13 is in a form whereby one could use closed loop planning.

5.3.3. Reactive Versus Closed Loop Planning

From Sections 5.3.1 and 5.3.2, we see that closed loop planning represents a significant increase in modelling and solution complexity compared to open loop planning. Thus, one needs to have strong evidence that closed loop brings significant benefits to the problem of interest. We have carried out several trials in the context of mine planning to study this question. Our trials covered the following:

(i) A simple mine planning problem: Here the geometry of the mine was very simple and thus, one might anticipate that closed loop might not yield significant benefits compared with reactive planning. Indeed, this turned out to be the case (Rojas et al., 2007).

(ii) A very simple toy example: So as to gain insight into the condition under which closed loop planning would yield significant benefits over reactive, we took a "toy" mining problem consisting of a one-dimensional vertical homogenous ore-body with uncertain ore price. Details are given in Rojas et al. (2007). Surprisingly, it turned out that, for this "toy" problem, reactive solutions were often identical to closed loop solutions. Indeed, closed loop only showed a major difference for very special (and arguably unlikely) scenario sets. Thus, one might conclude that closed loop planning is not beneficial.

Nonetheless, this conclusion needs to be viewed with some caution. Indeed, we have studied other problems (e.g. Networked control with random packet loss) where closed loop planning gives major benefits over open loop and/or reactive planning (Quevedo et al., 2008). Thus, one needs to evaluate each problem carefully before adopting a simplified approach.

6. GOOD, BAD OR OPTIMAL?

Before concluding the paper, we wish to caution against blind adherence to "optimality" as a design criterion. The point here is that the word "optimal" is a loaded word. In reality it simply means that one has found the maximum (or minimum) of some mathematical criterion. An entirely different question is whether or not the criterion adequately captures important practical features of the problem. Our point is that it is often better to have a deep understanding of a vague question then an exact answer to the wrong question. This idea is related to old ideas presented in the control literature in Rosenbrock and McMorran (1971). Indeed, the authors of the current paper believe that it is always a good idea to have a firm comprehension of fundamental limitations in design so that the space of possible 'answers' can be understood before blindly accepting an, 'optimal' solution. For example, in control, it is very important to understand limitations imposed by non-minimum phase behaviour, time delays and model uncertainty - see, for example, Freudenberg and Looze (1985); Seron et al. (1997). As an illustration, we point to the recent interest by industry in Nonlinear Model Predictive Control. We applaud the use of this sophisticated tool that will undoubtedly lead to practical control solutions which would otherwise be totally unobtainable. However, we urge users to always perform checks on fundamental limits, bandwidths, robustness etc before blindly accepting the output of a computer programme.

Some questions (amongst other) that a user of optimization tools might ask are

(i) How should I model the problem? What variables are important? What are the key constraints? What is the cost function?
(ii) Is uncertainty a major factor, and if so, how should I model uncertainty?
(iii) Are we interested in average performance, worst case performance, average performance subject to constraints on worst case etc etc?
(iv) Has the software reached the global optimum or terminated prematurely?
(v) How sensitive is the solution to key variables (typical optimization software provides information on sensitivity as part of the solution)?
(vi) Are there benchmarks (e.g., *ad-hoc* solutions or bounds) which can be used to evaluate the answers provided by optimization routines?
(vii) Does my model capture key issues of importance?
(viii) Are the nonlinearities that I have included sensible?
(ix) Do we understand all of the performance limitations applicable to the problem (e.g. "optimization" will not prevent a non-minimum phase plant from exhibiting undershoot)?
(x) Would a radically different approach give a better answer? For example, we have several examples in rolling mill control, where one can only achieve the desired level of performance by using a radically different "architecture" for the controller. This takes one outside the realm of optimization.

7. CONCLUSIONS

This paper has examined the opportunities that arise from the application of optimization in the mining, metal and mineral processing field. Several examples, drawn from the authors' experience, have been used to illustrate the interplay between modelling and solution strategies. Also, we have used these examples to reflect on the pluses and minuses of using optimization. We believe that optimization offers huge potential in mining, metal and mineral processing, provided it is used thoughtfully and in the context of wider knowledge of performance limits.

8. ACKNOWLEDGEMENTS

In writing this paper we have drawn on the results of many of our friends, colleagues and coworkers. We wish to specifically mention Adrian Lewis, (a reviewer who allowed his name to be revealed), Rick Middleton, Claus Müller, Peter Stone, Merab Menabde, Mei-Mei Zhang, Bryan Hennessy, Osvaldo Rojas and Cristian Rojas.

REFERENCES

J. Abadie, editor. *Nonlinear Programming.* North Holland, Amsterdam, The Netherlands, 1967.

P. B. Appiah and J. R. Sturgul. Pareto optimal stochastic mine planning. *International Journal of Surface Mining, Reclamation and Environment*, 4(4):181–195, 1990.

P. B. Appiah, M. A. Rosenman, and J. R. Sturgul. Introduction to Pareto optimal mine planning.

Journal of Mining and Geological Engineering, 8(4):348–356, 1990.

M. Ataei and M. Osanloo. Using a combination of genetic algorithm and the grid search method to determine optimum cutoff grades of multiple metal deposits. *International Journal of Surface Mining, Reclamation and Environment*, 18(1):60–78, 2003.

M.S. Bazaraa, H.D. Sherali, and C.M. Shetty. *Nonlinear Programming: Theory and Algorithms*. Wiley, New York, 2nd edition, 1993.

A. Beltratti, A. Consiglio, and S.A. Zenios. Scenario modeling for the management of international bond portfolios. *Annals of Operations Research*, 85:227–247, 1999.

A. Bemporad, L. Chisci, and E. Mosca. On the stabilizing property of siorhc. *Automatica*, 30: 2013–2015, 1995.

D.P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 2005.

L.T. Biegler and I.E. Grossmann. Retrospective on optimization. *Computers and Chemical Engineering*, 28:1169–1192, 2004.

I. Bongartz, A. R. Conn, Nick Gould, and Ph.L. Toint. CUTE: Constrained and unconstrained testing environment. *ACM Transactions on Mathematical Software*, 21(1):123–160, 1995.

F. Borrelli. *Constrained Optimal Control of Linear and Hybrid Systems*. Springer, Heidelberg, 2003.

J. M. Borwein and A. S. Lewis. *Convex Analysis and Nonlinear Optimization: Theory and Examples*. Springer, 2000.

S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2003. See: `http://www.stanford.edu/`,.

M. Brazil, D. H. Lee, M. Van Leuven, J. H. Rubinstein, D. A. Thomas, and N. C. Wormald. Optimising declines in underground mines. *Mining Technology: Trans. of the Institution of Mining and Metallurgy, Section A*, 112:164–170, 2003.

A. Brooke, D. Kendrick, A. Meeraus, and R. Raman. Gams: A user's guide, 1998. URL `www.gams.com`.

A.E.J. Bryson and M. Frazier. Smoothing for linear and nonlinear dynamic systems. Technical Report ASD-TDR 63–119, Aero. Syst. Div., Wright-Patterson Air Force Base, Ohio, USA, 1963.

L. Caccetta and S. P. Hill. An application of branch and cut to open pit mine scheduling. *Journal of Global Optimization*, 27:349–365, 2003.

E. F. Camacho and C. Bordons. *Model Predictive Control*. Advanced Textbooks in Control and Signal Processing. Springer, London, 1999.

S. Clement and N. Vagenas. Use of genetic algorithms in a mining problem. *International Journal of Surface Mining, Reclamation and Environment*, 8:131–136, 1994.

H. Cox. Estimation of state variables via dynamic programming. In *Proceedings of the Joint Automatic Control Conference*, pages 376–381, Stanford, CA, USA, 1964.

P. J. Darwen. Genetic algorithms and risk assessment to maximize NPV with robust open-pit scheduling. In *Fourth Biennial Conference on Strategic Mine Planning*, pages 29–34, Perth, Western Australia, 2001a.

P.J. Darwen. Genetic algorithms and risk assessment to maximize NPV with robust open-pit scheduling. In *Fourth Biennial Conference on Strategic Mine Planning*, pages 29–34, Perth, Western Australia, 2001b.

B. Denby and D. Schofield. Open-pit design and scheduling by use of genetic algorithms. *Transactions of the Inst. Min. Metall., Section A: Min. industry*, 103:A21–A26, 1994.

B. Denby and D. Schofield. The use of genetic algorithms in underground mine scheduling. In *Proceedings of the 25th International Symposium Application of Computers Mineral Industries*, pages 389–394, Brisbane, Australia, 1995a.

B. Denby and D. Schofield. Inclusion of risk assessment in open pit design and scheduling. *Transactions of the Inst. Min. Metall., Section A: Min. industry*, 104:A67–A71, 1995b.

B. Denby, D. Schofield, and T. Surme. Genetic algorithms for flexible scheduling of open pit operations. In *Proceedings of APCOM 27th International Symposium Mathematics in the Mineral Industries*, London, UK, 1998.

J. Dupačová. Scenario-based stochastic programs: resistance with respect to sample. *Annals of Operations Research*, 64:21–38, 1996.

J. Dupačová, G. Consigli, and S. W. Wallace. Scenarios for multistage stochastic programs. *Annals of Operations Research*, 100:25–53, 2000.

W.J. Edwards, A.J. Carlton, G.F. Harvey, R.F.K. Evans, and P.J. McKerrow. Coating mass control system design for a continuous galvanizing line. *Automatica*, 12:225–235, 1976.

A.V. Fiacco and G.P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. Society for Industrial and Applied Mathematics, Philadelphia, 1990.

R. Fletcher. *Practical methods of optimization*. Wiley, New York, 2000.

C.A. Floudas. *Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications*. Oxford University Press, New York, 1995.

R. Fourer, D. Gay, and B. Kernighan. *AMPL: A Modeling Language for Mathematical Programming*. The Scientific Press, South San Francisco, 1993.

J. Freudenberg and D. Looze. Right half plane poles and zeros and design trade offs in feedback systems. *IEEE Transactions on Automatic Control*, 30(6):555–568, 1985.

P.E. Gill, W. Murray, and M.H. Wright. *Practical Optimization*. Academic Press, 1981.

D. Goldberg. *Genetic algorithms in search, optimization, and machine learning*. Addison-Wesley, Reading, MA, 1989.

G.C. Goodwin, J.A. De Doná, M.M. Seron, and X.W. Zhuo. Lagrangian duality between constrained estimation and control. *Automatica*, 41(6):935–944, 2005a.

G.C. Goodwin, M.M. Seron, and J.A. De Doná. *Constrained Control and Estimation: An Optimisation Approach*. Springer, New-York, 2005b.

K. Høyland and S. W. Wallace. Generating scenario trees for multistage decision problems. *Management Science*, 47(11):295–307, 2001.

K. Holmströn, A. O. Göran, and M. M. Edvall. *User's guide for TOMLAB*. Tomlab Optimization Inc., December 2006.

O.L.R. Jacobs. Designing feedback controllers to regulate deposited mass in hot-dip galvanizing. *Control Engineering Practice*, 3(11):1529–1542, 1995.

A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Mathematics in Science and Engineering. Academic Press, 1970.

R.E. Kalman. A new approach to linear filtering and prediction problems. *Transaction of the ASME, Journal of Basic Engineering*, 82:34–45, 1960.

D. L. Keefer. Certainty equivalents for three-point discrete-distribution approximations. *Management Science*, 40(6):760–773, 1994.

S.S. Keerthi and E.G. Gilbert. Optimal infinite horizon feedback laws for a general class of constrained discrete time systems: Stability and moving-horizon approximations. *Journal of Optimization Theory and Applications*, 57:265–293, 1988.

S. Kirkpatrick, C. D. Gelatt Jr., and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220, 4598:671–680, 1983.

C. Løvaas, M.M. Seron, and G.C. Goodwin. Robust model predictive control of input-constrained stable systems with unstructured uncertainty. In *European Control Conference*, Kos, Greece, 2-5 July 2007.

D.G. Luenberger. *Linear and Nonlinear Programming*. Addison-Wesley, 2nd edition, 1989.

J.M. Maciejowski. *Predictive Control with Constraints*. Prentice Hall, 2002.

K. Marti, G. Ermolieu, and G. Pflug. *Dynamic Stochastic Optimization*. Springer, 2004.

D.Q. Mayne and H. Michalska. Receding horizon control of nonlinear systems. *IEEE Transactions on Automatic Control*, 35(5):814–824, 1990.

D.Q. Mayne, J.B. Rawlings, C.V. Rao, and P.O. M. Scokaert. Constrained model predictive control: stability and optimality. *Automatica*, 36(6):789–814, 2000.

H. Michalska and D.Q. Mayne. Moving horizon observers and observer based control. *IEEE Trans. Automatic Control*, 40(6):995–1006, 1995.

A. C. Miller and T. R. Rice. Discrete approximations of probability distributions. *Management Science*, 29(3):352–362, 1983.

C. Müller, X. W. Zhuo, and J. A. De Doná. Duality and symmetry in constrained estimation and control problems. *Automatica*, 42(12):2183–2188, 2006.

J.M. Mulvey and H. Vladimirou. Stochastic network programming for financial planning problems. *Management Science*, 38(11):1642–1664, 1992.

S.G. Nash and A. Sofer. *Linear and Nonlinear Programming*. McGraw-Hill, 1996.

J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer, New York, 1999.

G. C. Pflug. *Optimization of Stochastic Models: The Interface Between Simulation and Optimization*. Kluwer Academic Publishers, Boston, 1996.

G. Ch. Pflug. Scenario tree generation for multiperiod financial optimization by optimal discretization. *Math. Program., Ser. B*, 89:251–271, 2001a.

G.Ch. Pflug. Scenario tree generation for multiperiod financial optimization by optimal discretization. *Math. Programming, Ser. B*, 89:251–257, 2001b.

E. Polak, H. Chung, and S. Sastry. An external active-set strategy for solving optimal control problems. Technical Report UCB/EECS-2007-90, EECS Department, University of California, Berkeley, Jul 2007. URL `http://www.eecs.berkeley.edu/Pubs/TechRpts/2007/EECS-2007-90.html`.

S.J. Qin and T.A. Badgwell. An overview of industrial model predictive control technology. In Chemical Process Control–V, CACHE, AIChE, pages 232–256, 1997.

D. E. Quevedo, E. I. Silva, and G. C. Goodwin. Control over unreliable networks affected by packet erasures and variable transmission delays. *IEEE Journal on Selected Areas in Communications*, 2008. To appear.

C. V. Rao, J. Rawlings, and J. Lee. Constrained linear state estimation - a moving horizon approach. *Automatica*, 37(10):1619–1628, 2001.

C.V. Rao, J.B. Rawlings, and D.Q. Mayne. Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations. *IEEE Transactions on Automatic Control*, 48(2):246–258, 2003.

J.B. Rawlings and K.R. Muske. Stability of constrained receding horizon control. *IEEE Transactions on Automatic Control*, 38(10):1512–1516, 1993.

C.R. Rojas, G.C. Goodwin, and M.M. Seron. Open-Cut Mine Planning via Closed Loop Receding Horizon Optimal Control. Chapter in *Identification and Control: The gap between theory and practice,* R. Sanchez-Pena and V. Puig and J. Quevedo, Eds. Springer-Verlag, 2007.

O.J. Rojas, G.C. Goodwin, and T. Domanti. Nonlinear constrained model predictive control (nmpc) approach to air knife control in galvanizing lines. In *Advanced Process Control Applications*, Vancouver, Canada, 8-10 May 2006.

H. H. Rosenbrock and P. D. McMorran. Good, bad, or optimal? *IEEE Transactions on Automatic Control*, 16(6):552–554, December 1971.

J. A. Rossiter. *Model-Based Predictive Control*. CRC Press, 2003.

P.O.M. Scokaert and J.B. Rawlings. Constrained linear quadratic regulation. *IEEE Transactions on Automatic Control*, 43(8):1163–1169, 1998.

M. M. Seron, J. H. Braslavsky, and G. C. Goodwin. *Fundamental Limitations in Filtering and Control*. Springer-Verlag, Berlin, 1997.

P. Stone, G. Froyland, M. Menabde, B. Law, R. Pasyar, and P. Monkhouse. Blasor–blended iron-ore mine planning optimisation at Yandi. In *Orebody Modelling and Strategic Mine Planning", Proceedings of the International Symposium, AIMM*, pages 285–288, Perth, Australia, 2004.

M. Sznaier and M.J. Damborg. Suboptimal control of linear systems with state and control inequality constraints. In *Proceedings of the American Control Conference*, pages 761–762, Los Angeles, CA, 1987.

M. Sznaier and M.J. Damborg. Heuristically enhanced feedback control of discrete-time linear systems. *Automatica*, 26:521–532, 1990.

S. Takriti, J.R. Birge, and E. Long. A stochastic model for the unit commitment problem. *IEEE Transactions on Power Systems*, 11(3):1497–1508, 1996.

G. S. Thomas. Optimization and scheduling of open pits via genetic algorithms and simulated annealing. In *Proceedings of the 1st International Symposium on Balkema Publisher*, pages 44–59, Rotterdam, The Netherlands, 1996.

S. Uryasev and P.M. Pardalos, editors. *Stochastic Optimization: Algorithms and Applications*. Kluwer Academic Publishers, 2001.

L. Y. Yu, X. D. Ji, and S. Y. Wang. Stochastic programming models in financial optimization: a survey. *Advanced Modeling and Optimization*, 5(1), 2003.