Modeling, Simulation, and Optimization of a Production Line

Ludwig J. Cromme, Cottbus, and **Stephan Würll,** Berlin

Summary. Modern production facilities are complex systems involving thousands of components and system parameters. In this real world application we model a line for production and packaging of brown coal briquette bundles. We concentrate here on the problem of how to produce bundles with guaranteed (minimum) weight with minimal input of raw material. Variations of the incoming raw material parameters, the mechanics of the involved machinery, and complex interdependencies are mathematically modeled and analyzed. Based on this a computer model allows for simulation, parameter optimization, and minimization of input material. This results in a better understanding of the overall performance of the facility and of the sensitivity and interdependence of the parameters. Based on these investigations a considerable increase in production output could be realized by the industrial partner.

AMS subject classification: 00A72, 62P30, 90B30, 90C31, 90C90, 26B05, 65K05, 68U20

Keywords: Modeling, simulation, optimization, production line, packaging industry, brown coal, briquette bundles, guaranteed weight, computer integrated manufacturing CIM

1 Introduction

Computer integrated manufacturing (CIM) has become a powerful tool to control production processes and cut production costs. This is also true for the highly competitive energy market. Brown coal industry therefore developed new high quality products like brown coal dust for use in power plants and made great efforts to redesign classical products according to market needs and reduce production costs.

In an extensive project the production of brown coal briquette bundles was thoroughly investigated including the mechanical and thermal treatment of raw brown coal, pressing of briquettes, assembly of layers and bundles, wrapping, palletizing and stacking of the produced material. Problems to be tackled include control of briquette thickness and bundle weight, optimal conveyor belt system design and control of transport velocities and delay phases, optimal palletizing schemes, heat reduction for packed material (fire protection!), quality control for long term stock etc.

The project included experiments to determine system parameters, mathematical and computer based modeling, simulation (see fig. 1), and design of procedures for systems control and optimization.

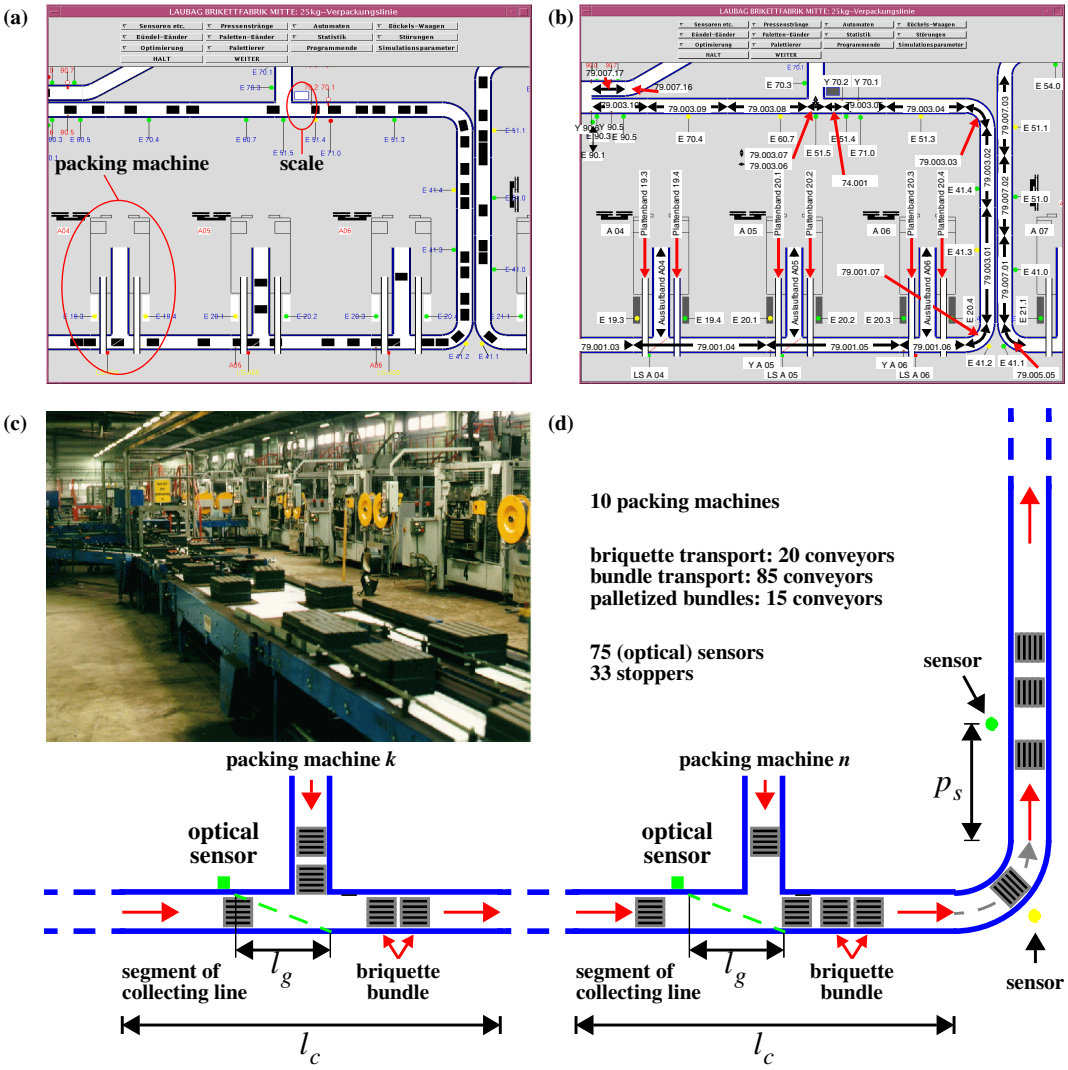

Fig. 1. (a) User interface (GUI) of simulation environment, (b) GUI with notation of the segments – illustrating the complexity of the plant, (c) automated line for 25 kg-bundles, (d) schematic representation of a part of the conveyor system

In this paper we investigate how with a minimum of raw material input (minimum of rejects) bundles of a given minimum weight can be produced. For some of the other aspects of the investigation see [2,3]. Briquettes are pressed in electric brown coal presses and have fixed width and height. Their thickness was verified to be approximately normally distributed with an electronically controllable mean value (adjustable target thickness). Varying weight and thickness of briquettes causes great problems in pro-

duction. 9 briquettes are assembled to a layer; the length of a layer (equal to the summed thickness of the 9 briquettes) must be at least $\delta_{\min}$ and at most $\delta_{\max}$ to be handable by the machines and to guarantee stable bundles. 5 such layers form a bundle which must weigh at least 25 000 g – the guaranteed minimum selling weight. A layer should therefore weigh 5 000 g on average. The weight of a layer varies with the length of the layer. Therefore, instead of discarding a possible first layer with weight below 5 000 g (and realizing the loss of 9 briquettes) one might accept it hoping for one of the following 4 layers to make up for the missing weight. But if the first layer is far below 5 000 g many layers might have to be tested (and rejected) before the first layer's missing weight is compensated. Obviously, the realization of an immediate loss of a layer needs to be carefully weighed up against the possible loss of many layers in the future; the machinery neither allows to reintegrate rejected layers into the production process nor is it possible to reject layers – once they are accepted – at a later state.

This means an (optimal) criterion γ_1 needs to be fixed: If a test layer weighs more than γ_1 it can be accepted as layer 1 of a new bundle, if it weighs less than γ_1 it is rejected.

More formally we have the following quality-control-algorithm (see also fig. 2,3). For given $0 < \delta_{\min} < \delta_{\max}$ and $\gamma_0 := 0 < \gamma_1 < \ldots < \gamma_4 < \gamma_5 := 25\,000$ g:

(A1) $i := 1$, $w^1_{\text{targ}} := \gamma_1$

(A2) A layer l is grasped by the claw arm.

(A3) If the length of layer l is not within $[\delta_{\min}, \delta_{\max}]$ reject it and go to (A2).

(A4) If the weight of layer l is below w^i_{targ} reject it and go to (A2).

Otherwise, the layer with weight w_i is accepted as layer i of the bundle.

(A5) If $i = 5$ STOP, a new bundle has been packed.

(A6) Set $i := i + 1$, $w^i_{\text{targ}} := \gamma_i - \sum_{k=1}^{i-1} w_k$, go to (A2).

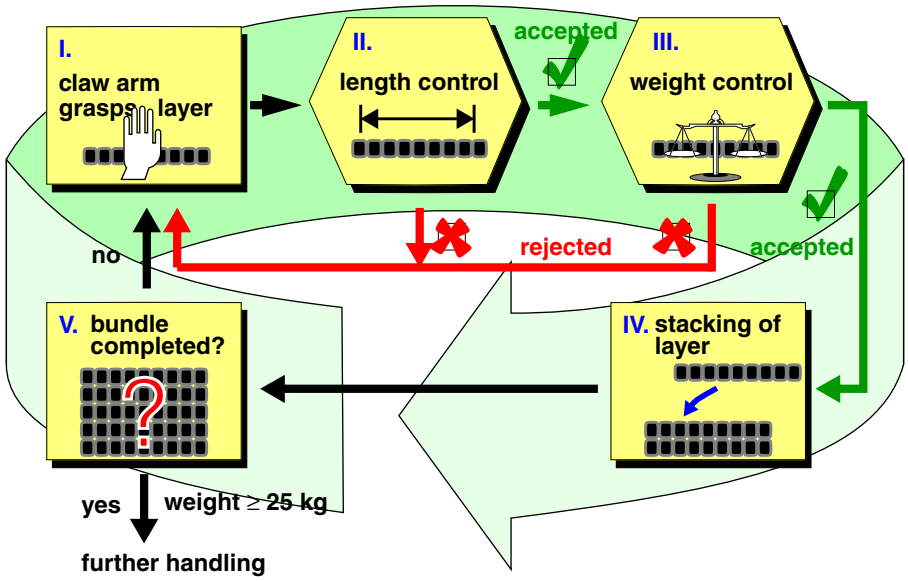

Fig. 2. Schematic representation of the quality control.

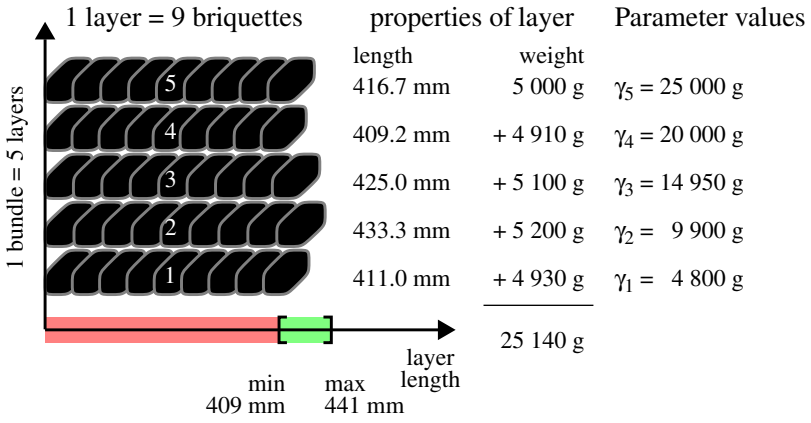

Fig. 3. Example of a packed briquette bundle; only layers that could be accepted for the given values of γ_i are shown.

Results of measurements confirmed that the briquette thickness (length of briquette layers) may be considered normally distributed and that the weight of a briquette (briquette layer) is approximately proportional to the briquette thickness (length of a layer) with the proportionality factor ρ ("specific briquette mass") in grams per mm constant on a short term scale.

2 Problem Formulation

The basic problem is to find optimal weight limits $\gamma_1, \ldots, \gamma_4$; optimal in the sense that a production process according to (A1) to (A6) will result in a minimum of rejects or – equivalently – that a maxi-

mum number of bundles is produced from the input for given process parameters: Normal distribution $L \sim N(\mu_L, \sigma_L^2)$ of the layer length, mean layer length μ_L, specific briquette mass ρ, minimal acceptable layer length $\delta_{\min}$ and maximal layer length $\delta_{\max} > \delta_{\min}$. We introduce $\gamma_0 := 0$ which helps to simplify some formulae and $\gamma_5 = 25\,000$ g is the (fixed) required minimal bundle weight.

For $\rho \geq \gamma_5/(5\delta_{\min})$ the optimal parameters are easily seen to be $\gamma_i := i\delta_{\min}\rho$; that is, any layer within the layer length tolerance can be accepted.

Additional requirements are needed to ensure that the problem is well-defined. The assumption

$$\rho \geq \frac{\gamma_5}{5\delta_{\max}}$$

guarantees, for example, the existence of producible bundles. The following assumptions guarantee finiteness of algorithm (A1) to (A6) and are not redundant:

$$\begin{aligned} \gamma_0 := 0, \qquad 0 < \delta_{\min} < \delta_{\max}, \qquad 0 < \gamma_5 < 5\delta_{\max}\rho, \\ \delta_{\min}\rho \leq \gamma_{i+1} - \gamma_i < \delta_{\max}\rho, \qquad i = 0, 1, \ldots, 4 \end{aligned} \tag{1}$$

The restrictions in (1) describe a linearly bounded partly open, partly closed set of feasible points $(\gamma_1, \ldots, \gamma_4)$. Existence of optimal points is therefore nontrivial; see section 4.

3 Mathematical Model

As motivated above we assume now that the layer lengths L and corresponding weights W are independently and identically normally distributed, i.e.

$$L \sim N(\mu_L, \sigma_L^2), \qquad W = \rho \cdot L, \qquad W \sim N(\mu_W, \sigma_W^2),$$

with mean value $\mu_W = \mu_L\rho$ and standard deviation $\sigma_W = \sigma_L\rho$. The density of the random variable W then is

$$\varphi(x) := \frac{1}{\sigma_W\sqrt{2\pi}} \cdot \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu_W}{\sigma_W}\right)^2},$$

and the accompanying distribution function is defined by the following expression:

$$\Phi(x) := \int_{-\infty}^{x} \varphi(u)du = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{x-\mu_W}{\sigma_W\sqrt{2}}\right)\right).$$

$\Phi(x)$ is the probability that a layer picked at random has weight x at most. Here $\mathrm{erf}(\cdot)$ denotes the error function closely related to the Gauss error integral:

$$\mathrm{erf}(x) \;:=\; \frac{2}{\sqrt{\pi}} \cdot \int_0^x \mathrm{e}^{-t^2} dt$$

The probability that a gripped layer is accepted in step (A4), i.e. it has at least the weight w^i_{targ} and its length lies within the tolerance interval $[\delta_{\min}, \delta_{\max}]$, is:

$$\begin{aligned} p(w^i_{\mathrm{targ}}) &:= P(L \in [\delta_{\min}, \delta_{\max}] \wedge W \geq w^i_{\mathrm{targ}}) \\ &= P(W \in [\max\{\delta_{\min}\rho, w^i_{\mathrm{targ}}\}, \delta_{\max}\rho]) \\ &= \begin{cases} 0 & \text{if } w^i_{\mathrm{targ}} \geq \delta_{\max}\rho, \\ \Phi(\delta_{\max}\rho) \;-\; \Phi(\max\{\delta_{\min}\rho, w^i_{\mathrm{targ}}\}) & \text{otherwise} \end{cases} \end{aligned} \tag{2}$$

The acceptance probability $p(\cdot)$ is continuous on $\mathbb{R}$, but the derivative of $p(\cdot)$ has jump discontinuities in $\delta_{\min}\rho$ and $\delta_{\max}\rho$. The minimum weight a layer must have to be acceptable as layer i in step (A4) of the algorithm is w^i_{targ} and it depends on the weight of the already accepted layers

$$w^1_{\mathrm{targ}} \;:=\; \gamma_1, \qquad\qquad w^i_{\mathrm{targ}} \;:=\; \gamma_i - \sum\nolimits_{k=1}^{i-1} w_k\,. \tag{3}$$

Let $A^i := A^i(\gamma_1, \ldots, \gamma_i) \subseteq \mathbb{R}^i$ denote the set of acceptable tuples of layer weights $(w_1, \ldots, w_i)^T \in \mathbb{R}^i$, i.e.

$$A^1 \;:=\; [\max\{\delta_{\min}\rho, \gamma_1\}, \delta_{\max}\rho] \;=\; [\gamma_1, \delta_{\max}\rho]$$

$$A^i \;:=\; \left\{ (w_1, \ldots, w_i)^T \;\middle|\; (w_1, \ldots, w_{i-1})^T \in A^{i-1},\;\; w_i \in \left[\max\left\{ \delta_{\min}\rho, \gamma_i - \sum_{k=1}^{i-1} w_k \right\}, \delta_{\max}\rho \right] \right\}$$

$$i \;=\; 2, 3, 4, 5$$

From the above definition of A^i (or the definition of the algorithm in section 1) we conclude

$$\gamma_{i-1} \;\leq\; \sum_{k=1}^{i-1} w_k$$

which together with (1) implies

$$\gamma_i - \sum_{k=1}^{i-1} w_k \;\leq\; \gamma_i - \gamma_{i-1} \;<\; \delta_{\max}\rho,$$

such that the A^i have nonempty interior, are closed and linearly constrained, in particular convex ($1 \leq i \leq 5$).

X_i denotes the random variable which describes how many randomly picked layers have to be tested until one is accepted at the i^{th} location of the bundle. W^i_{targ} (probability density function $\varphi^i_{\text{targ}} : \mathbb{R}^i \to [0, 1]$) is the random variable whose realization is w^i_{targ}. The probability distribution of the discrete random variable X_i, for given value w^i_{targ}, is then given by the following expression.

$$P(X_i = j \,|\, W^i_{\text{targ}} = w^i_{\text{targ}}) \;=\; [1 - p(w^i_{\text{targ}})]^{j-1} p(w^i_{\text{targ}}), \qquad j \in \mathbb{N}$$

Summing over j gives

$$E(X_i \,|\, W^i_{\text{targ}} = w^i_{\text{targ}}) \;=\; \sum_{j=1}^{\infty} j \cdot P(X_i = j \,|\, W^i_{\text{targ}} = w^i_{\text{targ}}) \;=\; \frac{1}{p(w^i_{\text{targ}})} \tag{4}$$

and the expectation value is given by the integrals

$$E(X_1) \;=\; \frac{1}{\Phi(\delta_{\max}\rho) - \Phi(\gamma_1)}$$

$$E(X_{i+1}) \;=\; \frac{\int_{A^i} \varphi(w_1) \cdot \ldots \cdot \varphi(w_i) \cdot \dfrac{1}{p(w^i_{\text{targ}})} dA^i}{\int_{A^i} \varphi(w_1) \cdot \ldots \cdot \varphi(w_i) dA^i} \qquad i \;=\; 1, 2, 3, 4 \tag{5}$$

The formula for $E(X_1)$ can be concluded directly from (4) with (2) and (3), since the only possible value of W^i_{targ} is γ_1 in this case. The formula for $E(X_{i+1})$ results from the following consideration. The observed weights $(w_1, \ldots, w_i)$ of layers already accepted by the algorithm are realizations of an i-dimensional random variable $(W_1, \ldots, W_i)$. The corresponding density is given by

$$\Psi(w_1, \ldots, w_i) := \begin{cases} \dfrac{\varphi(w_1) \cdot \ldots \cdot \varphi(w_i)}{\int_{A^i} \varphi(\tilde{w}_1) \cdot \ldots \cdot \varphi(\tilde{w}_i) dA^i} & \text{for } w_1, \ldots, w_i \in A^i \\ 0 & \text{otherwise.} \end{cases}$$

Then by definition the expectation value $E(X_{i+1})$ is given by

$$E(X_{i+1}) \;=\; \int_{\mathbb{R}^i} [E(X_i \,|\, W^i_{\text{targ}} = w^i_{\text{targ}}) \cdot \Psi(w_1, \ldots, w_i)] \;=\; \frac{\int_{A^i} \varphi(w_1) \cdot \ldots \cdot \varphi(w_i) \cdot \dfrac{1}{p(w^i_{\text{targ}})} dA^i}{\int_{A^i} \varphi(w_1) \cdot \ldots \cdot \varphi(w_i) dA^i}$$

which proves (5). For an explicit description of the $E(X_i)$ see section 6.1.

It is far from obvious that the above expressions are well-defined. The constraints $\gamma_{i+1} - \gamma_i < \delta_{\max}\rho$ (see (1)) come into play here. The estimate $w^i_{\text{targ}} = \gamma_i - \sum_{k=1}^{i-1} w_k \leq \gamma_i - \gamma_{i-1} < \delta_{\max}\rho$ shows that the lower integration limits $\max\{\delta_{\min}\rho, w^i_{\text{targ}}\}$ never coincide with the corresponding upper limits. For the same reason probability $p(w^i_{\text{targ}})$ is always positive. Thus, all denominators are positive and the expectation values $E(X_i)$ are well-defined.

In case $\gamma_i = \gamma_{i-1} + \delta_{\min}\rho$ for some i $(1 \leq i \leq 5)$ the weight limit γ_i is not significant since any layer within length tolerance $[\delta_{\min}, \delta_{\max}]$ is acceptable. In analogy to (4) the expectation reduces in this case to

$$E(X_i \,|\, W^i_{\text{targ}} = \delta_{\min}\rho) = \sum_{j=1}^{\infty} j \cdot P(X_i = j \,|\, W^i_{\text{targ}} = \delta_{\min}\rho) = \frac{1}{p(\delta_{\min}\rho)}.$$

Our aim is to fix the real parameters $\gamma_1, \ldots, \gamma_4$ such that the expected number of layers needed for the composition of a bundle becomes minimal. This can now be formalized as the following nonlinear optimization problem.

Optimization Problem

For fixed real numbers $\delta_{\min}$, $\delta_{\max}$, ρ, γ_5 and $\gamma_0 := 0$ with $0 < \delta_{\min} < \delta_{\max}$ and $0 < \gamma_5 < 5\delta_{\max}\rho$ minimize the objective function

$$\Omega(\gamma_1, \gamma_2, \gamma_3, \gamma_4) := \sum_{i=1}^{5} E(X_i)$$

$$\text{on} \qquad C := \left\{ \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{pmatrix} \middle| \begin{array}{c} \delta_{\min}\rho \leq \gamma_{i+1} - \gamma_i < \delta_{\max}\rho, \\ i = 0, 1, \ldots, 4 \end{array} \right\} \tag{6}$$

where $E(X_i)$ is defined as in (5) or (9) to (12).

4 The Objective Function

Numerical optimization methods are, as a rule, more efficient if far-reaching requirements on the objective function (continuous, differentiable) and the feasible point set (bounded, closed) are fulfilled. We therefore check the objective function for some of these properties in the sequel.

Differentiability

The objective function Ω is continuous on C since $p(\cdot)$ is positive on C. Furthermore, Ω is positive on C and it can even be shown that $E(X_i) > 1$, since the integration limits do not coincide, the functions $\varphi(\cdot)$ and $p(\cdot)$ are positive on the domain of integration and $p(\cdot) < 1$ on C ($i = 2, 3, 4, 5$). To investigate the differentiability of the function Ω with respect to the variables $\gamma_1, \ldots, \gamma_4$ turns out to be much more difficult.

Let us take a closer look at $E(X_3)$. For the partial derivative of the inner integral of the denominator of $E(X_3)$ with respect to γ_2 we get

$$\frac{\partial}{\partial \gamma_2} \int_{\max\{\delta_{\min}\rho,\, \gamma_2 - w_1\}}^{\delta_{\max}\rho} \varphi(w_1)\varphi(w_2)\, dw_2 \;=\; \begin{cases} -\varphi(w_1)\varphi(\gamma_2 - w_1) & \text{if } \gamma_2 > w_1 + \delta_{\min}\rho \\ 0 & \text{if } \gamma_2 < w_1 + \delta_{\min}\rho \end{cases}.$$

This follows from Leibniz' rule for the derivative of an integral with variable integration limits (see section 6.2). Since we get different limits for the two cases the integral *does not have* a continuous derivative with respect to γ_2 and it seems reasonable to suspect that Ω is not differentiable.

However, a detailed examination shows that Ω is (totally) differentiable in contrast to this expectation. This is a result of the fact that the jump discontinuity in the derivative of the integral with respect to γ_2 is smoothened by the following integration with respect to w_1 (!!). We get:

Lemma 1 *The objective function of problem (6) is (totally) differentiable with respect to the parameters* γ_i, $i = 1, \ldots, 4$. *The set of all feasible points is bounded and the intersection of a finite number of partly open, partly closed half-spaces.*

The proof of Lemma 2 is presented in section 6.3.

Existence of Minima

Since the constraint set C of our optimization problem (6) is not compact existence of optimal points cannot be guaranteed at this point. The problem is not just theoretical: If an optimization algorithm generates a sequence $\vec{\gamma}^n = (\gamma_1^n, \gamma_2^n, \gamma_3^n, \gamma_4^n)_{n \in \mathbb{N}}$ of parameter values approaching a point $\lim_{n \to \infty} \vec{\gamma}^n = \vec{\gamma}^* \in \partial C$ on the boundary of C, the expectation value $E(X_i)$ is not defined for the limit $\vec{\gamma}^*$, if $\lim_{n \to \infty} (\gamma_i^n - \gamma_{i-1}^n) = \delta_{\max}\rho$, $\gamma_0^n := 0$ holds. In this case $p(\gamma_i{}^* - w_{\text{targ}}^i) = 0$ would appear in the denominator of equation (5). In this section we prove that the level sets of (6) are closed which guarantees that the problems sketched above cannot occur and that optimal points exist.

Let us try to understand the situation intuitively: With weight limits γ_i^n, γ_{i-1}^n, $\gamma_i^n \approx \delta_{\max}\rho + \gamma_{i-1}^n < \delta_{\max}\rho + \gamma_i^n$ and accepted layers with weight $w_1^n, \ldots, w_{i-1}^n$ and $\sum_{j=1}^{i-1} w_j^n - \gamma_{i-1}^n \approx 0$ a tested layer must weigh almost exactly $\delta_{\max}\rho$ to be acceptable as layer i in this situation: A weight above $\delta_{\max}\rho$ is not acceptable because the length tolerance criterion is violated and only a weight close to $\delta_{\max}\rho$ guarantees $\sum_{j=1}^{i} w_j^n \geq \gamma_i^n$. This means that on average many layers have to be tested until one is acceptable as layer i. On the other hand the situation $\sum_{j=1}^{i-1} w_j^n \approx \gamma_{i-1}^n$ itself is unlikely and intuition is not reliable to decide whether $E(X_i)$ remains finite or not. The precise answer is given in the following

Lemma 2 *Let* $\vec{\gamma}^n = (\gamma_1^n, \gamma_2^n, \gamma_3^n, \gamma_4^n)_{n \in \mathbb{N}}$ be a sequence of feasible parameter values. If there exists at least one index i with $\lim_{n \to \infty} (\gamma_i^n - \gamma_{i-1}^n) = \delta_{\max}\rho$ where $\gamma_0^n := 0$ the sequence of the objective function values is unbounded: $\lim_{n \to \infty} \Omega(\gamma_1^n, \gamma_2^n, \gamma_3^n, \gamma_4^n) = \infty$.

A detailed proof of Lemma 2 is given in section 6.4.

5 Numerical Solution of the Optimization Problem

The two main problems in connection with the numerical treatment of the optimization problem (6) are

(i) the numerical evaluation of the objective function;

(ii) the choice of a suitable optimization algorithm.

5.1 Numerical evaluation of the objective function

To evaluate the objective function (6) a multidimensional numerical integration method is required. This could be based on recursive application of the following one-dimensional quadrature formula.

$$\int_{a_1}^{b_1}\int_{a_2(x_1)}^{b_2(x_1)}\cdots\int_{a_n(x_1,\ldots,x_{n-1})}^{a_n(x_1,\ldots,x_{n-1})} f(x_1,\ldots,x_n)dx_n\ldots dx_1 = \int_{a_1}^{b_1} g(x_1)dx_1$$
$$g(x_1) := \int_{a_2(x_1)}^{b_2(x_1)}\cdots\int_{a_n(x_1,\ldots,x_{n-1})}^{b_n(x_1,\ldots,x_{n-1})} f(x)dx_n\ldots dx_2 \qquad (7)$$

However, since the numerical effort grows fast for increasing n, even for smooth functions and adaptive knot choice, this method suffers from inefficiency. Furthermore, error bounds are not transferable to the recursive case from one-dimensional quadrature methods directly, since the function $g(x_1)$ in formula (7) itself can only be approximated and cannot be determined with machine precision. In practice this leads to slow convergence. Therefore one might prefer the exact quadrature of a linear or second order approximation of $f(x_1, \ldots, x_n)$, if no high precision is demanded.

There also exist algorithms for multiple integrals analogously to the development of adaptive quadrature methods for one-dimensional integration. But here the following difficulties arise:

- As a rule, multidimensional numerical integration methods are formulated only for special integration areas e.g. for rectangular areas. IMSL® [11] offers only the possibility to calculate integrals of the form $\int_{a_1}^{b_1}\int_{a_2(x)}^{b_2(x)} f(x, y)dydx$ or $\int_{a_1}^{b_1}\ldots\int_{a_n}^{b_n} f(x_1, \ldots, x_n)dx_n\ldots dx_1$.
- Adaptive multidimensional integration methods which allow a reliable error estimation are offered only in few, mostly commercial, numerical libraries (e.g. IMSL®).
- The value of the objective function on the given integration area cannot be determined directly with any of the currently available numerical libraries.

For these reasons two different approaches were pursued and the results were compared to each other: Romberg's one-dimensional quadrature formula was implemented recursively as one approach and alternatively algorithms from the class library Cubpack++ (Cools et al. 1994) were used because of their easy applicability to areas with almost arbitrary boundaries.

Romberg's quadrature formula

When evaluating the objective function Ω (see equation (6)) by use of Romberg's one-dimensional quadrature method for $f \in C^{2k+2}[a, b],\ k \geq 0$, the following estimate can be used to calculate an upper bound of the error of the numerical integration, by means of Euler-Maclaurin's summation formula (Werner 1992, [10]):

$$\left| \int_a^b f(x)\ dx - T^{(k)}(f;a, b) \right| \leq 3(b-a)^{2k+3} \frac{|B_{2k+2}|}{(2k+2)!} \max_{x \in [a, b]} \left| f^{(2k+2)}(x) \right|$$

in which the term B_{2k+2} is a Bernoulli number, defined as coefficient of the power series expansion $\frac{x}{e^x - 1} = \sum_{k=0}^{\infty} B_k \frac{x^k}{k!},\ |x| < 2\pi$.

Unfortunately, this result is not valid for the case of the multidimensional (recursive) integration.

Adaptive quadrature / cubature by means of DQNC79 and Cubpack++

As an alternative to Romberg's quadrature method, the integral (9) was evaluated by applying routine DQNC79 of the SLATEC numerical library. (More precisely: The FORTRAN code of DQNC79 was translated by f2c into C.) This is an adaptive quadrature method based on a 7 point Newton-Côtes formula. This way the quadrature error and the number of the function evaluations could be checked precisely.

The C++ class library Cubpack++ [4] provides the user with routines for automatic cubature. By the object-oriented implementation of Cubpack++ this library is not static but can easily be enlarged. Those areas which are not included in Cubpack++ yet can simply be added by inheritance. For these areas the user must carry out a transformation of the integral on an area for which a cubature formula exists.

For these reasons the integrals (9)-(12) were evaluated by a combination of Cubpack++ and DQNC79.

5.2 Optimization

As seen the optimization problem (6) has a differentiable objective function, is linearly constrained with partly closed (≤), partly open (<) constraints. In view of Lemma 2 (closed level sets) descent algorithms are therefore appropriate for the problem. Since derivatives of the objective function are expensive to evaluate (see the preceding section) we prefer methods that make no explicit use of derivatives. This leaves methods with no assumptions about differentiability at all like simulated annealing, genetic

algorithms, methods of Hooke and Jeeves, Nelder and Mead etc. and methods that assume differentiability in their philosophy but do not demand evaluation of derivatives explicitly (conjugate directions algorithm by Powell).

As starting point we used the parameter set used in the plant before this investigation

$$(\tilde{\gamma}_1, \tilde{\gamma}_2, \tilde{\gamma}_3, \tilde{\gamma}_4) = (4800\mathrm{g}, 9900\mathrm{g}, 14950\mathrm{g}, 20000\mathrm{g}) \tag{8}$$

if $(\tilde{\gamma}_1, \tilde{\gamma}_2, \tilde{\gamma}_3, \tilde{\gamma}_4) \in C$. Alternatively starting parameters were chosen according to the following scheme

$$(\tilde{\gamma}_1, \tilde{\gamma}_2, \tilde{\gamma}_3, \tilde{\gamma}_4) = (5000\mathrm{g} - 4w, 10000\mathrm{g} - 2w, 15000\mathrm{g} - w, 20000\mathrm{g}),\ w \in 0 \ldots 50\mathrm{g}\,.$$

The experience from our numerical investigations can be summarized as follows: the conjugate direction method of Powell as described in [1, pp. 464-467] and Hooke and Jeeves [7] gave acceptable results with the least cost. The Downhill-Simplex-Method by Nelder and Mead [8] also gave acceptable results but needed around 50% to 150% more (costly) function evaluations; simulated annealing was even slower with a factor of 10 for the number of necessary function evaluations.

We also experimented with algorithms that use gradient information and approximated the gradients numerically. Altogether our experience with numerical approximation of gradients was not encouraging for this particular objective function.

Extremely important for practical applications is a good knowledge of the sensitivity of optimal parameters from possibly varying input parameters. While $\delta_{\min}$, $\delta_{\max}$, and the mean μ_W are constant over longer periods of time, σ_W and ρ sometimes vary within hours. The table in fig. 4 demonstrates that the values of the minimum of the objective function change rapidly with changing values of σ_W and ρ. We discuss the consequences of this observation in the following section.

Fig. 4. Objective function minima for varying specific briquette mass and standard deviation of the layer length and default parameter settings $\mu_L = 47\,\mathrm{mm}$, $\delta_{\min} = 409\,\mathrm{mm}$, $\delta_{\max} = 441\,\mathrm{mm}$

5.3 Interpretation

As seen above an optimal parameter set $(\gamma_1^*, \gamma_2^*, \gamma_3^*, \gamma_4^*)$ allows for a significant increase in briquette production output. At the same time the optimal parameter set is highly dependent on parameter values like σ_W and ρ. This means that even minor changes in these parameters may result in significant production losses unless the parameters $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ are also readjusted.

This leads to the following important recommendation: If parameters ρ and σ_W change the weight limits should be readjusted with the help of a computer-aided system or look-up table.

In the following we check alternative approaches.

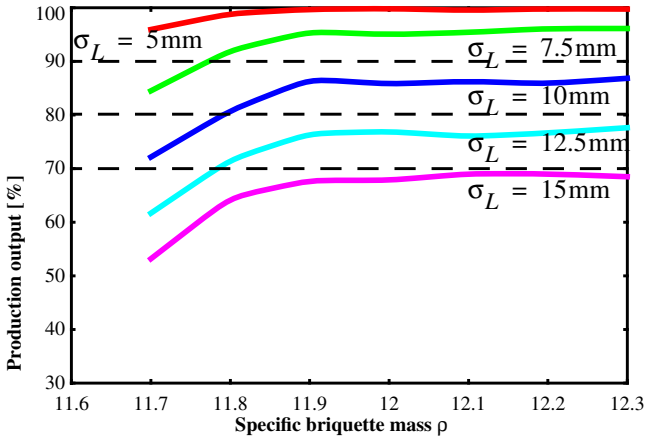

Fig. 5. Maximum production output for different values of σ_L and for optimal μ_L and weight limits $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$, depending on specific briquette mass ρ. Lower and upper length limit were chosen as $\delta_{\min} = 409\text{mm}$, $\delta_{\max} = 441\text{mm}$.

μ_L can be controlled by a well performing 3-point briquette thickness control at the briquette presses. We therefore examined how the values of the objective function are influenced by the mean layer length μ_L. The maximum attainable packing quota can be seen in fig. 5. It is obvious that a good choice of the mean layer length is very important for a good production rate. However, this again is no longer true for a non-optimal choice of the weight limits.

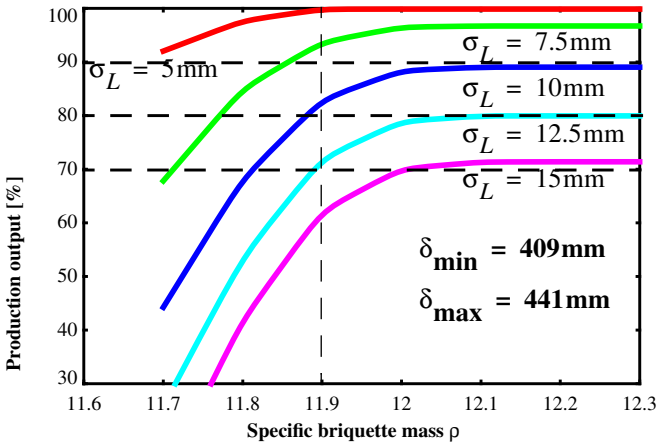

Fig. 6. Maximum production for optimal μ_L without weight control, depending on specific briquette mass ρ. Lower and upper length limit were chosen as $\delta_{\min} = 409\text{mm}$, $\delta_{\max} = 441\text{mm}$.

Therefore, we examined the results that can be achieved if there is no weight control of single layers and the mean briquette thickness is controlled instead. The attainable packing quotas can be seen in fig. 6. In specific situations the strategy “Bundling without weight control” proves to be competitive to the quality surveillance procedure based on a check of single layer weights. Furthermore the strategy is less dependent on changing production parameters.

These calculated optimal values of the reference value setting, presented in fig. 7, have the feature of being almost independent of the scatter of the value σ_L of a certain production unit. In addition, to guarantee a nearly optimal production output, it is sufficient to adjust the reference value at the center of the interval, limited by $\delta_{\min}$ and $\delta_{\max}$, as long as the specific briquette mass doesn't fall below 12 g/mm.

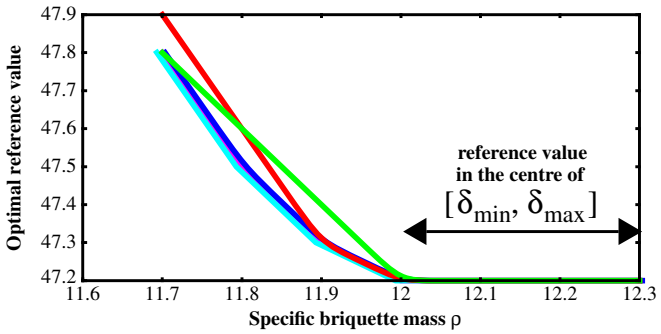

Fig. 7. Optimal setting of the reference value in dependence of the specific briquette mass ρ. Lower and upper length limit were chosen as

$$\delta_{\min} = 409\,\text{mm}, \delta_{\max} = 441\,\text{mm}$$

This may even lead to the recommendation to run the production without weighing each layer if production parameters like ρ and σ_L vary strongly and if the parameters $\gamma_1, \ldots, \gamma_4$ cannot be held optimal or nearly optimal at all times.

6 An Extended Leibniz Rule, Proof of Lemma 1 and Lemma 2

6.1 A detailed representation of the expectation values

The following equations are required in order to fully understand why we have to extend Leibniz' rule. Therefore we present the expectation values in (5) in detail:

$$E(X_2) = \frac{\int_{\gamma_1}^{\delta_{\max}\rho} \frac{\varphi(w_1)}{p(\gamma_2 - w_1)}\, dw_1}{\int_{\gamma_1}^{\delta_{\max}\rho} \varphi(w_1)\, dw_1} = E(X_1) \cdot \int_{\gamma_1}^{\delta_{\max}\rho} \frac{\varphi(w_1)}{p(\gamma_2 - w_1)}\, dw_1 \tag{9}$$

$$E(X_3) = \frac{\int_{\gamma_1}^{\delta_{\max}\rho} \int_{\max\{\delta_{\min}\rho,\, \gamma_2 - w_1\}}^{\delta_{\max}\rho} \frac{\varphi(w_1)\varphi(w_2)}{p(\gamma_3 - w_1 - w_2)}\, dw_2 dw_1}{\int_{\gamma_1}^{\delta_{\max}\rho} \int_{\max\{\delta_{\min}\rho,\, \gamma_2 - w_1\}}^{\delta_{\max}\rho} \varphi(w_1)\varphi(w_2)\, dw_2 dw_1} \tag{10}$$

$$E(X_4) = \frac{\int_{\gamma_1}^{\delta_{\max}\rho} \cdots \int_{\max\{\delta_{\min}\rho,\, \gamma_3 - w_1 - w_2\}}^{\delta_{\max}\rho} \frac{\varphi(w_1)\varphi(w_2)\varphi(w_3)}{p(\gamma_4 - w_1 - w_2 - w_3)}\, dw_3 .. dw_1}{\int_{\gamma_1}^{\delta_{\max}\rho} \cdots \int_{\max\{\delta_{\min}\rho,\, \gamma_3 - w_1 - w_2\}}^{\delta_{\max}\rho} \varphi(w_1)\varphi(w_2)\varphi(w_3)\, dw_3 .. dw_1} \tag{11}$$

$$E(X_5) = \frac{\int_{\gamma_1}^{\delta_{\max}\rho} \cdots \int_{\max\{\delta_{\min}\rho,\, \gamma_4 - w_1 - w_2 - w_3\}}^{\delta_{\max}\rho} \frac{\varphi(w_1)\varphi(w_2)\varphi(w_3)\varphi(w_4)}{p(\gamma_5 - w_1 - w_2 - w_3 - w_4)}\, dw_4 .. dw_1}{\int_{\gamma_1}^{\delta_{\max}\rho} \cdots \int_{\max\{\delta_{\min}\rho,\, \gamma_4 - w_1 - w_2 - w_3\}}^{\delta_{\max}\rho} \varphi(w_1)\varphi(w_2)\varphi(w_3)\varphi(w_4)\, dw_4 .. dw_1} \tag{12}$$

6.2 Leibniz' Rule

To prove the differentiability of the target function Ω of our optimization problem it would be helpful if differentiation and integration could be interchanged. But the classical Leibniz' rule is not applicable here, because the assumptions usually made do not hold. We therefore state an extended Leibniz' rule here with weaker assumptions which simplifies the proof a lot.

Lemma 3 (Leibniz' Rule [6]) *Let $f : D \to \mathbb{R}$ have continuous partial derivatives on $D = [a, b] \times [c, d]$. If the curves $\alpha : [c, d] \to \mathbb{R}$ and $\beta : [c, d] \to \mathbb{R}$ are continuously differentiable and the values of α and β are in $[a, b]$ then the following identity holds:*

$$\frac{\partial}{\partial\gamma} \int_{\alpha(\gamma)}^{\beta(\gamma)} f(g, \gamma) dg = \int_{\alpha(\gamma)}^{\beta(\gamma)} \frac{\partial}{\partial\gamma} f(g, \gamma) dg + \beta'(\gamma) f(\beta(\gamma), \gamma) - \alpha'(\gamma) f(\alpha(\gamma),$$

Lemma 4 (Extended Leibniz' Rule [5]) *The function* $f : \mathbb{R}^2 \to \mathbb{R}, f = f(g, \gamma)$, *be continuous with respect to the variables* g *and* γ *and have a continuous partial derivative* $\frac{\partial f}{\partial \gamma}$ *for* $\gamma - g \leq k$ *and for* $\gamma - g \geq k$, $k \in \mathbb{R}$. *Thus, there may be a jump discontinuity of the partial derivative for* $\gamma - g = k$. *Then* $\int_a^b f(g, \gamma) dg$ *has a partial derivative with respect to* γ. *The derivative* $\frac{\partial f}{\partial \gamma}$ *is a continuous function and the following holds:*

$$\frac{\partial}{\partial \gamma} \int_a^b f(g, \gamma) dg = \int_a^b \frac{\partial}{\partial \gamma} f(g, \gamma) dg.$$

6.3 Proof of Lemma 1

Lemma 1 *The objective function of problem (6) is (totally) differentiable with respect to the parameters* γ_i, $i = 1, \ldots, 4$. *The set of all feasible points is bounded and the intersection of a finite number of partly open, partly closed half-spaces.*

Proof The properties of C follow directly from (1) and (6). The continuity of Ω has already been shown. We prove now that all $E(X_i)$ have continuous partial derivatives for all γ_j.

For $E(X_1)$ this follows from

$$\frac{\partial}{\partial \gamma_1} E(X_1) = \frac{\varphi(\gamma_1)}{[\Phi(\delta_{\max}\rho) - \Phi(\gamma_1)]^2}.$$

Let us now look at the partial derivatives of $E(X_2)$. With Leibniz' rule (section 6.2) we get

$$\frac{\partial}{\partial \gamma_1} E(X_2) = \frac{\partial}{\partial \gamma_1} E(X_1) \cdot \int_{\gamma_1}^{\delta_{\max}\rho} \frac{\varphi(w_1)}{p(\gamma_2 - w_1)} dw_1 - E(X_1) \cdot \frac{\varphi(\gamma_1)}{p(\gamma_2 - \gamma_1)}.$$

Since $\gamma_2 - \gamma_1 < \delta_{\max}\rho$ the denominator $p(\gamma_2 - \gamma_1)$ is positive. From (1) follows $\gamma_2 - w_1 \leq \gamma_2 - \gamma_1 < \delta_{\max}\rho$ such that $p(\gamma_2 - w_1) \neq 0$ in the integration area and the integrand is well-defined and continuous: $E(X_2)$ has a continuous partial derivative with respect to γ_1.

Because of the jump discontinuity of $\partial p(\gamma_2 - w_1)/\partial \gamma_2$ in $\gamma_2 - w_1 = \delta_{\min}\rho$ Leibniz' rule is not applicable in deriving $\partial E(X_2)/\partial \gamma_2$. The extended Leibniz' rule (Lemma 4, section 6.2) is applicable, though, since $\partial p(\gamma_2 - w_1)/\partial \gamma_2$ is continuous for $\gamma_2 - w_1 \leq \delta_{\min}\rho$ and for $\gamma_2 - w_1 \geq \delta_{\min}\rho$ and despite the discontinuity of $p'(\cdot)$ we get the continuity of

$$\frac{\partial}{\partial \gamma_2} E(X_2) = -E(X_1) \int_{\gamma_1}^{\delta_{\max}\rho} \frac{\varphi(w_1) p'(\gamma_2 - w_1)}{[p(\gamma_2 - w_1)]^2}\, dw_1$$

on C. Since $w_1 > \gamma_2 - \delta_{\min}\rho$ implies $p'(\gamma_2 - w_1) = 0$ we even have

$$\frac{\partial}{\partial \gamma_2} E(X_2) = E(X_1) \int_{\gamma_1}^{\min\{\delta_{\max}\rho,\, \gamma_2 - \delta_{\min}\rho\}} \frac{\varphi(w_1)\varphi(\gamma_2 - w_1)}{[p(\gamma_2 - w_1)]^2}\, dw_1.$$

Thus the partial derivatives of $E(X_2)$ are well-defined and continuous.

Let us now consider the partial derivatives of $E(X_3)$. As above $\partial E(X_3)/\partial \gamma_1$ is seen to exist and be continuous. Next consider $\partial E(X_3)/\partial \gamma_2$. Despite the discontinuity of the derivative of $\max\{\delta_{\min}\rho, \gamma_2 - w_1\}$ for $\gamma_2 = w_1 + \delta_{\min}\rho$ the extended Leibniz' rule is applicable again which gives us continuity of

$$\begin{aligned}
&\frac{\partial}{\partial \gamma_2} \int_{\gamma_1}^{\delta_{\max}\rho} \int_{\max\{\delta_{\min}\rho,\, \gamma_2 - w_1\}}^{\delta_{\max}\rho} \frac{\varphi(w_1)\varphi(w_2)}{p(\gamma_3 - w_1 - w_2)}\, dw_2 dw_1 \\
&= \int_{\gamma_1}^{\delta_{\max}\rho} \frac{\partial}{\partial \gamma_2} \int_{\max\{\delta_{\min}\rho,\, \gamma_2 - w_1\}}^{\delta_{\max}\rho} \frac{\varphi(w_1)\varphi(w_2)}{p(\gamma_3 - w_1 - w_2)}\, dw_2 dw_1 \\
&= -\int_{\gamma_1}^{\delta_{\max}\rho} \frac{\partial}{\partial \gamma_2} \max\{\delta_{\min}\rho, \gamma_2 - w_1\} \cdot \frac{\varphi(w_1)\varphi(\max\{\delta_{\min}\rho, \gamma_2 - w_1\})}{p(\gamma_3 - w_1 - \max\{\delta_{\min}\rho, \gamma_2 - w_1\})}\, dw_1
\end{aligned}$$

Similarly the continuity of the partial derivative of the denominator is seen and $\frac{\partial}{\partial \gamma_2} E(X_3)$ is continuous. Again the derivative can be further simplified:

.

$$\frac{\partial}{\partial \gamma_2} \int_{\gamma_1}^{\delta_{\max}\rho} \int_{\max\{\delta_{\min}\rho,\, \gamma_2 - w_1\}}^{\delta_{\max}\rho} \frac{\varphi(w_1)\varphi(w_2)}{p(\gamma_3 - w_1 - w_2)}\, dw_2 dw_1 = -\int_{\gamma_1}^{\min\{\delta_{\max}\rho,\, \gamma_2 - \delta_{\min}\rho\}} \frac{\varphi(w_1)\varphi(\gamma_2 - w_1)}{p(\gamma_3 - \gamma_2)}\, dw_1$$

Let us now look at the partial derivative of $E(X_3)$ with respect to γ_3. Discontinuities of $p'(\cdot)$ are causing problems again. The derivative $p'(\cdot)$ is continuous except for the points $w^i_{\text{targ}} := \delta_{\min}\rho$ and $w^i_{\text{targ}} := \delta_{\max}\rho$ where $p'(\cdot)$ has jump discontinuities. Because of the integration limits in the numerator of $E(X_3)$ we get

$$\gamma_3 - w_1 - w_2 \le \gamma_3 - \gamma_1 - (\gamma_2 - w_1) = (\gamma_3 - \gamma_2) - (\gamma_1 - w_1) < \delta_{\max}\rho \tag{13}$$

from (1). Therefore the upper jump discontinuity is not relevant here, but $\gamma_3 - w_1 - w_2 = \delta_{\min}\rho$ is not excluded. The assumptions of the extended Leibniz' rule are valid with a discontinuity of $p(\gamma_3 - w_1 - w_2)$ in $\gamma_3 - w_1 - w_2 = \delta_{\min}\rho$ which proves continuity of

$$\frac{\partial}{\partial\gamma_3}\int_{\gamma_1}^{\delta_{\max}\rho}\int_{\max\{\delta_{\min}\rho,\,\gamma_2-w_1\}}^{\delta_{\max}\rho}\frac{\varphi(w_1)\varphi(w_2)}{p(\gamma_3-w_1-w_2)}\,dw_2dw_1$$
$$=-\int_{\gamma_1}^{\delta_{\max}\rho}\int_{\max\{\delta_{\min}\rho,\,\gamma_2-w_1\}}^{\delta_{\max}\rho}\frac{\varphi(w_1)\varphi(w_2)p'(\gamma_3-w_1-w_2)}{[p(\gamma_3-w_1-w_2)]^2}\,dw_2dw_1$$

Defining ad hoc a modified integral operator

$$\dot{\int}_a^b f(x)dx := \begin{cases}\int_a^b f(x)dx & \text{if} \quad a<b\\ 0 & \text{otherwise}\end{cases}$$

we can further simplify this:

$$\frac{\partial}{\partial\gamma_3}\int_{\gamma_1}^{\delta_{\max}\rho}\int_{\max\{\delta_{\min}\rho,\,\gamma_2-w_1\}}^{\delta_{\max}\rho}\frac{\varphi(w_1)\varphi(w_2)}{p(\gamma_3-w_1-w_2)}\,dw_2dw_1$$
$$=\int_{\gamma_1}^{\delta_{\max}\rho}\dot{\int}_{\max\{\delta_{\min}\rho,\,\gamma_2-w_1\}}^{\min\{\delta_{\max}\rho,\,\gamma_3-w_1-\delta_{\min}\rho\}}\frac{\varphi(w_1)\varphi(w_2)\varphi(\gamma_3-w_1-w_2)}{[p(\gamma_3-w_1-w_2)]^2}\,dw_2\,dw_1$$

Equation (13) implies that the denominator of the integrand is positive and the integration interval of the inner integral is nonempty.

The differentiability of $E(X_4)$ and $E(X_5)$ is seen in analogy. ❑

A list of the derivatives can be obtained from the authors on request.

6.4 Proof of Lemma 2

Lemma 2 *Let* $\vec{\gamma}^n = (\gamma_1^n, \gamma_2^n, \gamma_3^n, \gamma_4^n)_{n\in\mathbb{N}}$ be a sequence of feasible parameter values. If there exists at least one index i with $\lim_{n\to\infty}(\gamma_i^n - \gamma_{i-1}^n) = \delta_{\max}\rho$ where $\gamma_0^n := 0$ the sequence of the objective function values is unbounded: $\lim_{n\to\infty}\Omega(\gamma_1^n, \gamma_2^n, \gamma_3^n, \gamma_4^n) = \infty$.

Proof We look at the cases $i = 1, 2$ etc. successively.

(i) $i = 1$: From $\lim_{n\to\infty}\gamma_1^n = \delta_{\max}\rho$ follows directly

$$\lim_{n\to\infty} E(X_1) = \lim_{n\to\infty} \frac{1}{\Phi(\delta_{\max}\rho) - \Phi(\gamma_1^n)} = \frac{1}{\Phi(\delta_{\max}\rho) - \lim\limits_{n\to\infty} \Phi(\gamma_1^n)} = \infty.$$

Thus, the expectation value $E(X_1)$ is unbounded and so is the objective function.

(ii) $i = 2$: Without loss of generality we may assume $\lim\limits_{n\to\infty} \gamma_1^n \neq \delta_{\max}\rho$ since otherwise we could argue as in case (i). In the following we therefore assume $\lim\limits_{n\to\infty} \gamma_1^n = \gamma_1{}^* < \delta_{\max}\rho$ and $\gamma_2{}^* - \gamma_1{}^* = \delta_{\max}\rho$. The function $\Phi(x)$ is Lipschitz continuous with respect to the constant $L = 1/(\sigma_W\sqrt{2\pi})$, i.e. $|\Phi(x_1) - \Phi(x_2)| \le L|x_1 - x_2|$. With the auxiliary variables $M := \int_{\delta_{\min}\rho}^{\delta_{\max}\rho} \varphi(w)dw$, $m := \min\{\varphi(w) | w \in [\delta_{\min}\rho, \delta_{\max}\rho]\}$ we get:

$$\begin{aligned} E(X_2) &= \frac{\int\limits_{\gamma_1^n}^{\delta_{\max}\rho} \frac{\varphi(w_1)}{p(\gamma_2^n - w_1)}\, dw_1}{\int\limits_{\gamma_1^n}^{\delta_{\max}\rho} \varphi(w_1)\, dw_1} \ge \frac{m}{M} \int\limits_{\gamma_1^n}^{\delta_{\max}\rho} \frac{1}{p(\gamma_2^n - w_1)}\, dw_1 \\ &= \frac{m}{M} \int\limits_{\gamma_1^n}^{\delta_{\max}\rho} \frac{1}{\Phi(\delta_{\max}\rho) - \Phi(\max\{\delta_{\min}\rho, \gamma_2^n - w_1\})} dw_1 \\ &\ge \frac{m}{M} \int\limits_{\gamma_1^n}^{\delta_{\max}\rho} \frac{\sigma_W\sqrt{2\pi}}{\delta_{\max}\rho - \max\{\delta_{\min}\rho, \gamma_2^n - w_1\}}\, dw_1 \end{aligned} \tag{14}$$

The last inequality in (14) results from the Lipschitz continuity of $\Phi(x)$. Splitting the integration interval at $\min\{\delta_{\max}\rho, \gamma_2^n - \delta_{\min}\rho\}$ allows for an explicit solution of the integrals:

$$E(X_2) \geq \frac{m}{M}\left(\int_{\gamma_1^n}^{\min\{\delta_{\max}\rho,\, \gamma_2^n - \delta_{\min}\rho\}} \frac{\sigma_W\sqrt{2\pi}}{\delta_{\max}\rho - \max\{\delta_{\min}\rho,\, \gamma_2^n - w_1\}}\, dw_1 \right.$$

$$\left. + \int_{\min\{\delta_{\max}\rho,\, \gamma_2^n - \delta_{\min}\rho\}}^{\delta_{\max}\rho} \frac{\sigma_W\sqrt{2\pi}}{\delta_{\max}\rho - \max\{\delta_{\min}\rho,\, \gamma_2^n - w_1\}}\, dw_1 \right)$$

$$= \frac{m\sigma_W\sqrt{2\pi}}{M}\left(\ln(\delta_{\max}\rho - (\gamma_2^n - w_1))\Big|_{\gamma_1^n}^{\min\{\delta_{\max}\rho,\, \gamma_2^n - \delta_{\min}\rho\}} \right. \tag{15}$$

$$\left. + \frac{\delta_{\max}\rho - \min\{\delta_{\max}\rho,\, \gamma_2^n - \delta_{\min}\rho\}}{\delta_{\max}\rho - \delta_{\min}\rho} \right)$$

$$= \frac{m\sigma_W\sqrt{2\pi}}{M}(\ln(\delta_{\max}\rho - (\gamma_2^n - \min\{\delta_{\max}\rho,\, \gamma_2^n - \delta_{\min}\rho\})) - \ln(\delta_{\max}\rho - (\gamma_2^n - \gamma_1^n)))$$

$$+ \frac{m\sigma_W\sqrt{2\pi}}{M} \cdot \frac{\delta_{\max}\rho - \min\{\delta_{\max}\rho,\, \gamma_2^n - \delta_{\min}\rho\}}{\delta_{\max}\rho - \delta_{\min}\rho}$$

Because of our assumption $\lim_{n\to\infty}(\gamma_2^n - \gamma_1^n) = \delta_{\max}\rho$ there exists for every $\varepsilon > 0$ an index $n_0 = n_0(\varepsilon)$, such that $\delta_{\max}\rho - (\gamma_2^n - \gamma_1^n) \leq \varepsilon$ for all $n \geq n_0$. Inserting this in (15) gives us:

$$E(X_2) \geq \frac{m\sigma_W\sqrt{2\pi}}{M}(\ln(\delta_{\max}\rho - (\gamma_2^n - \min\{\delta_{\max}\rho,\, \gamma_2^n - \delta_{\min}\rho\})) - \ln(\varepsilon))$$

$$+ \frac{m\sigma_W\sqrt{2\pi}}{M} \cdot \frac{\delta_{\max}\rho - \min\{\delta_{\max}\rho,\, \gamma_2^n - \delta_{\min}\rho\}}{\delta_{\max}\rho - \delta_{\min}\rho} =: e(\varepsilon)$$

Since $\varepsilon > 0$ can be chosen arbitrarily small we have a divergent minorant and thus

$$\lim_{\substack{\gamma_2^n - \gamma_1^n \to \delta_{\max}\rho \\ n\to\infty}} E(X_2) \geq \lim_{\substack{\varepsilon \to 0^+ \\ n \geq n(\varepsilon)}} e(\varepsilon) = \infty.$$

(iii) $i \geq 3$: Without loss of generality we assume $\lim_{n\to\infty}\gamma_1^n \neq \delta_{\max}\rho$ and $\lim_{n\to\infty}(\gamma_k^n - \gamma_{k-1}^n) \neq \delta_{\max}\rho,\ k = 2, \ldots, i-1$; otherwise case (i), case (ii) or case (iii) for $i-1$ would apply. Construction of a divergent minorant is similar to case (ii). Since the observed weights $(w_1, \ldots, w_{i-1})$ of layers already accepted by the algorithm are realizations of an $(i-1)$-dimensional random variable $(W_1, \ldots, W_{i-1})$ we can define a further random variable W_Σ as the sum $W_\Sigma := W_1 + \ldots + W_{i-1}$. The random variable $W_\Sigma = W_\Sigma(\gamma_1^n, \ldots, \gamma_{i-1}^n)$ describes the weight of $i-1$ stacked layers. Let $\varphi_\Sigma(w) = \varphi_\Sigma(w, \gamma_1^n, \ldots, \gamma_{i-1}^n)$ denote the density of W_Σ. Because of the assumption $\lim_{n\to\infty}\gamma_1^n \neq \delta_{\max}\rho$, $\lim_{n\to\infty}(\gamma_k^n - \gamma_{k-1}^n) \neq \delta_{\max}\rho$,

$k = 2, \ldots, i-1$, and the properties of any probability density function we have $\varphi_\Sigma(w) \geq 0$, $\int_{-\infty}^{\infty} \varphi_\Sigma(w)dw = 1$ and there exists for every $\varepsilon > 0$ an index $n_0 = n_0(\varepsilon)$, such that

$$\left| \min\{\varphi_\Sigma(w, \gamma_1^n, \ldots, \gamma_{i-1}^n) \middle| w \in [\gamma_{i-1}, (i-1)\delta_{\max}\rho]\} \right.$$
$$\left. - \min\{\varphi_\Sigma(w, \gamma_1{}^*, \ldots, \gamma_{i-1}{}^*) \middle| w \in [\gamma_{i-1}{}^*, (i-1)\delta_{\max}\rho]\} \right| \leq \varepsilon$$

for all $n \geq n_0$. With the auxiliary variables

$$M_n = \int_{\gamma_{i-1}^n}^{(i-1)\delta_{\max}\rho} \varphi_\Sigma(w)dw \leq 1 =: M$$

$$2m := \min\{\varphi_\Sigma(w, \gamma_1{}^*, \ldots, \gamma_{i-1}{}^*) \middle| w \in [\gamma_{i-1}{}^*, (i-1)\delta_{\max}\rho]\} > 0$$

$$\varepsilon := m$$

$E(X_i)$ can be reduced for almost all indices $n \in \mathbb{N}$ to the following expression:

$$\begin{aligned} E(X_i) &= \frac{\int_{\gamma_{i-1}^n}^{(i-1)\delta_{\max}\rho} \frac{\varphi_\Sigma(w)}{p(\gamma_i^n - w)}\, dw}{\int_{\gamma_{i-1}^n}^{(i-1)\delta_{\max}\rho} \varphi_\Sigma(w)\, dw} \geq \frac{m}{M} \int_{\gamma_{i-1}^n}^{(i-1)\delta_{\max}\rho} \frac{1}{p(\gamma_i^n - w)}\, dw \\ &\geq \frac{m}{M} \int_{\gamma_{i-1}^n}^{(i-1)\delta_{\max}\rho} \frac{\sigma_G\sqrt{2\pi}}{\delta_{\max}\rho - \max\{\delta_{\min}\rho, \gamma_i^n - w\}}\, dw \end{aligned} \tag{16}$$

The last inequality in (16) results from the Lipschitz continuity of $\Phi(x)$. Splitting the integration interval at $\min\{(i-1)\delta_{\max}\rho, \gamma_i^n - \delta_{\min}\rho\}$ allows for an explicit solution of the integrals:

$$
\begin{aligned}
E(X_i) \;\geq\; & \frac{m}{M}\left(\int_{\gamma_{i-1}^n}^{\min\{(i-1)\delta_{\max}\rho,\, \gamma_i^n - \delta_{\min}\rho\}} \frac{\sigma_W\sqrt{2\pi}}{\delta_{\max}\rho - \max\{\delta_{\min}\rho,\, \gamma_i^n - w\}}\, dw \right. \\
& \left. + \int_{\min\{(i-1)\delta_{\max}\rho,\, \gamma_i^n - \delta_{\min}\rho\}}^{(i-1)\delta_{\max}\rho} \frac{\sigma_W\sqrt{2\pi}}{\delta_{\max}\rho - \max\{\delta_{\min}\rho,\, \gamma_i^n - w\}}\, dw \right) \\
= \; & \frac{m\sigma_W\sqrt{2\pi}}{M}\left(\ln(\delta_{\max}\rho - (\gamma_i^n - w))\Big|_{\gamma_{i-1}^n}^{\min\{(i-1)\delta_{\max}\rho,\, \gamma_i^n - \delta_{\min}\rho\}} \right. \\
& \left. + \frac{(i-1)\delta_{\max}\rho - \min\{(i-1)\delta_{\max}\rho,\, \gamma_i^n - \delta_{\min}\rho\}}{\delta_{\max}\rho - \delta_{\min}\rho} \right) \\
= \; & \frac{m\sigma_W\sqrt{2\pi}}{M}\ln(\delta_{\max}\rho - (\gamma_i^n - \min\{(i-1)\delta_{\max}\rho,\, \gamma_i^n - \delta_{\min}\rho\})) \\
& - \frac{m\sigma_W\sqrt{2\pi}}{M}\ln(\delta_{\max}\rho - (\gamma_i^n - \gamma_{i-1}^n)) \\
& + \frac{m\sigma_W\sqrt{2\pi}}{M} \cdot \frac{(i-1)\delta_{\max}\rho - \min\{(i-1)\delta_{\max}\rho,\, \gamma_i^n - \delta_{\min}\rho\}}{\delta_{\max}\rho - \delta_{\min}\rho}
\end{aligned}
\tag{17}
$$

In analogy to case (ii) it is easy to verify that the term $-\ln(\delta_{\max}\rho - (\gamma_i^n - \gamma_{i-1}^n))$ in (17) is unbounded for $\lim_{n\to\infty}(\gamma_i^n - \gamma_{i-1}^n) = \delta_{\max}\rho$. Thus, the expectation value is unbounded, i.e.

$$
\lim_{\substack{n\to\infty \\ \gamma_i^n - \gamma_{i-1}^n \to \delta_{\max}\rho}} E(X_i) = \infty.
$$

To summarize we can say that the objective function is unbounded, if $\lim_{n\to\infty}\gamma_1^n = \delta_{\max}\rho$ or $\lim_{n\to\infty}(\gamma_i^n - \gamma_{i-1}^n) = \delta_{\max}\rho$ for a suitable chosen index $i \in \{2, 3, 4, 5\}$ holds. ❑

7 Conclusions

We analyzed how with a minimum of raw material input brown coal bundles of a given minimum weight can be produced. In a first step we modeled the relevant part of the production facility and formulated the problem as a mathematical optimization problem. Next we analyzed the objective function for differentiability and proved existence of minimum points. Finally, the problem could be solved numerically. A sensitivity analysis showed that optimal parameters may depend strongly on material and system parameters. Therefore frequent readjustments are necessary if material and system parameters change. As an alternative approach we discussed a weight control via the mean thickness μ_L and

without weighing each layer of a bundle. The proposed actions lead to a significant improvement of the facilities overall performance. This would not have been possible without an in-depth mathematical analysis.

8 References

[1] F.S. ACTON. *"Numerical Methods that Work"*, The Mathematical Association of America, Washington, 1990

[2] J. BIEDERMANN, L. CROMME, K. WEBER, S. WÜRLL, F. REHN, L. RHEDE. "*Modellierung von Verpackungsprozessen für Braunkohlenbriketts*", Braunkohle - Surface Mining, 49. Jahrgang, Nr. 4, 8/1997

[3] J. BIEDERMANN, L. CROMME, L. RHEDE, K. WEBER, S. WÜRLL. "Simulation and optimization of a production process in brown coal industry". In A. Sydow, editor, Proceedings of the 15th IMACS World Congress on Scientific Computation, Modelling and Applied Mathematics, volume 5: Systems Engineering, pages 707–712. Wissenschaft & Technik Verlag, Berlin, August 1997.

[4] R. COOLS, D. LAURIE, L. PLUYM. *"Cubpack++: A C++ package for automatic two-dimensional cubature",* Report TW 220, Dept. of Computer Science, K.U.Leuven, 1994.

[5] L. CROMME. "An Extended Leibniz' Rule and Applications.", Nonlinear Analysis, to appear.

[6] G.M. FICHTENHOLZ. "Differential- und Integralrechnung II", VEB Deutscher Verlag der Wissenschaften Berlin, 1972.

[7] R. HOOKE, T.A. JEEVES. *"Direct Search Solution of Numerical and Statistical Problems"*, Journal of the ACM 8, pp. 212-229, 1961.

[8] J.A. NELDER, R. MEAD. *"A simplex method for function minimization"*, Computer Journal 7, pp. 308-313, 1965.

[9] W.H. PRESS et.al. *"Numerical Recipes in C"*, 2nd Edition, Cambridge University Press, 1992

[10] J. WERNER. *"Numerische Mathematik 1"*, Vieweg, 1992.

[11] IMSL®, Fortran and C application development tools from Visual Numerics, Manual "Fortran subroutines for mathematical applications", vol. 1, pp. 615-620, 1994.

Authors addresses: Ludwig Cromme, Universitätsplatz 3-4, D–03044 Cottbus,

Germany; email: cromme@math.tu-cottbus.de

Stephan Würll, Lufthansa Systems Berlin GmbH, Fritschestr. 28-29, D–10585 Berlin,

Germany; email: Stephan.Wuerll@LHSystems.com