Combined Maintenance Scheduling and Production Optimization

Adriaen Verheyleweghen

Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), Norway. E-mail: adriaen.verheyleweghen@ntnu.no

Himanshu Srivastav

Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology (NTNU), Norway. E-mail: himanshu.srivastav@ntnu.no

Anne Barros

Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology (NTNU), Norway. E-mail: anne.barros@ntnu.no

Johannes Jäschke

Department of Chemical Engineering, Norwegian University of Science and Technology (NTNU), Norway. E-mail: johannes.jaschke@ntnu.no

Optimal operation of complex production and processing plants is important, but challenging to achieve in practice. The reason for this is that decisions in different disciplines, such as design, operations and maintenance, are made independently of each other. This can lead to a large degree of conservativeness. In this paper, we present a unified approach for maintenance- and production planning, which reduces the conservativeness and leads to more economical operation. We model the system using differential equations and then formulate the problem of optimal operation as a numerical optimization problem. The problem is a mathematical program with equilibrium constraints (MPEC), which we solve using off-the-shelf optimization software. Some model approximations were made to make the system numerically tractable. We demonstrate the method on a subsea-inspired case example.

Keywords: Reliability modeling, maintenance scheduling, production optimization

1. Introduction

For certain classes of production systems there exist a trade-off between producing as much as possible, and prematurely wearing the system out. For example, in subsea oil and gas production systems the revenue is directly related with the amount of produced hydrocarbons, while a too high production rate may lead to fast system degradation, with expensive maintenance operations and possible production loss due to downtime.

From an economical point of view, there exists an optimal trade-off between the maintenance cost, the inspection cost and the operational profit. Moreover, when the system has degraded, the operator needs to make a decision relating to what degree the system should be restored. Using commonly employed Monte Carlo methods to obtain the optimal production profile and the optimal maintenance schedule, represents a significant challenge due to the sheer amount of possible scenarios that need to be explored. Numerical optimization seems like an attractive alternative to Monte Carlo methods due to the potential for faster convergence to an optimal solution.

In this paper we propose an integrated approach to operate the system optimally, that is, we propose to integrate the decisions on 1) the system load (how much to produce), 2) when to perform a maintenance operation, and 3) to what degree the system should be restored, in a unified framework.

To demonstrate our approach, we model a subsea oil and gas production system using a four-state Markov chain, where state A represents the new, healthy system, states B and C represent progressively degraded systems, and state D represents the failed, inoperable system. Arrival at the failed state D can be caused by unexpected sudden failure, or due to progressive degradation. The time dependent transition rates are a function of the input usage, thus yielding a multiphase Markov decision process. The production system model is described by a non-linear differential-algebraic equation system (DAE).

Optimal production and operation planning is defined as the case when the sum of the expected value of the revenue minus the inspection cost and

Proceedings of the 29th European Safety and Reliability Conference.
Edited by Michael Beer and Enrico Zio

ISBN: 981-973-0000-00-0 :: doi: 10.3850/981-973-0000-00-0_esrel2019-paper

the maintenance cost, is maximized. As decision variables in the optimization problem we assume the inspection times and the input profile. By inputs, we here mean the operating mode that the plant is running at. For example, the inputs of a compressor could be its throughput and frequency.

At each inspection, all systems found in the failed state D are restored to state A (if we follow the as-good-as-new (AGAN) policy). If the system is not found in state D, the system does not reveal whether its true state is A, B or C, and no maintenance is performed. Consequently, the model becomes a switched differential algebraic model, and the optimization problem can be formulated as a non-smooth, non-linear program. We solve this problem using state of the art methods for non-smooth optimization.

Authors of previously published work on the topic of combined maintenance and production planning typically formulate the problem as a lot-allocation problem (Iravani and Duenyas, 2002; Fitouhi and Nourelfath, 2012; Wolter and Helber, 2016). This often results in a mixed-integer (non-)linear program (MI(N)LP). Our proposed method is different as we do not consider "lots" of products, but rather the case where production can be adjusted in a continuous fashion. We also avoid the use of integer variables by formulating the problem with complementary constraints instead.

The remainder of the paper is structured as follows; in Section 2 we show how a continuous differential equation can be derived for the case of a degrading system with discrete inspection- and maintenance times. In Section 3 it is shown how the derived model can be used in the context of optimization, where the aim is to manipulate inputs and the maintenance times to minimize some objective function. In Section 4, the method is demonstrated on the subsea case example. Finally, concluding remarks and thoughts on future work are given in Section 5.

2. Modeling the degrading system

Given a four-state Markov process as shown in Fig. 1. Let $\boldsymbol{\mu}(t) = [\mu_A(t)\ \mu_B(t)\ \mu_C(t)\ \mu_D(t)]^\intercal$ denote the probabilities for being in any state at time t. Assuming time-invariant transition rates λ_a and λ_u between the states, the change in probabilities between two inspections is given by

$$\frac{d\boldsymbol{\mu}}{dt} = \boldsymbol{\Lambda} \cdot \boldsymbol{\mu}(t) \tag{1}$$

$$= (\boldsymbol{\Lambda}_a + \boldsymbol{\Lambda}_u) \cdot \boldsymbol{\mu}(t) \tag{2}$$

where

$$\boldsymbol{\Lambda}_a = \begin{bmatrix} -\lambda_a & 0 & 0 & 0 \\ \lambda_a & -\lambda_a & 0 & 0 \\ 0 & \lambda_a & -\lambda_a & 0 \\ 0 & 0 & \lambda_a & 0 \end{bmatrix} \tag{3}$$

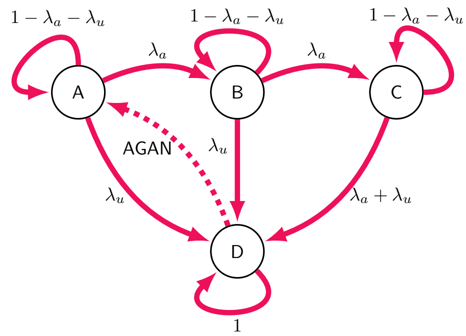

Fig. 1. Markov chain for a system with four discrete degradation states

and

$$\boldsymbol{\Lambda}_u = \begin{bmatrix} -\lambda_u & 0 & 0 & 0 \\ 0 & -\lambda_u & 0 & 0 \\ 0 & 0 & -\lambda_u & 0 \\ \lambda_u & \lambda_u & \lambda_u & 0 \end{bmatrix}. \tag{4}$$

In the above expressions, $\boldsymbol{\Lambda}$ is known as the transition matrix. $\boldsymbol{\Lambda}$ is decomposed into $\boldsymbol{\Lambda}_a$, which is describing the transitions due to aging, and $\boldsymbol{\Lambda}_u$, which is describing the transitions due to unforeseen failures.

Integrating Eq. (1) gives

$$\boldsymbol{\mu}(t) = \exp(\boldsymbol{\Lambda} \cdot (t - t_0))\boldsymbol{\mu}(t_0) \tag{5}$$

2.1. *Probabilities between two inspections*

Upon inspection of the system at time t_1, we reveal if the system is broken down (in state D), or not (either in state A, B or C). This leads to two different cases, depending on the outcome:

Case I:

Upon inspection, we find the system is in state D. Thus, we restore the system by performing maintenance (without time lag). Assuming perfect repairs according to the AGAN policy, the new initial conditions t_1 are

$$\boldsymbol{\mu}^+_{\text{Case I}}(t_1) = [1\ 0\ 0\ 0]^\intercal \tag{6}$$

and

$$\boldsymbol{\mu}_{\text{Case I}}(t > t_1) = \exp\left(\boldsymbol{\Lambda} \cdot (t - t_1)\right) \boldsymbol{\mu}^+_{\text{Case I}}(t_1) \tag{7}$$

Note that we use the notation $\boldsymbol{\mu}^+(t_1)$ to indicate the right-handed limit of $\boldsymbol{\mu}(t_1)$, i.e. directly after the inspection at t_1, and $\boldsymbol{\mu}^-$ to indicate the left-handed limit of $\boldsymbol{\mu}$, i.e. directly before the inspection at t_1. Because $\boldsymbol{\mu}$ is discontinuous at t_1, these two limits will generally not be equal.

Case II:

Upon inspection, we find that the system is not in state D. However, we do not know if the system is in state A, B or C. The new initial conditions at t_1 are:

$$\boldsymbol{\mu}^{+}_{\text{Case II}}(t_1) = \left[\frac{\mu^{-}_{A}(t_1)}{1-\mu^{-}_{D}(t_1)} \ \frac{\mu^{-}_{B}(t_1)}{1-\mu^{-}_{D}(t_1)} \ \frac{\mu^{-}_{C}(t_1)}{1-\mu^{-}_{D}(t_1)} \ 0\right]^{\mathsf{T}} \tag{8}$$

and

$$\boldsymbol{\mu}_{\text{Case II}}(t > t_1) = \exp\left(\boldsymbol{\Lambda}\cdot(t-t_1)\right)\boldsymbol{\mu}^{+}_{\text{Case II}}(t_1) \tag{9}$$

Expressing the probabilities from the perspective of t_0

However, we can not know ahead of time which of the two cases will be observed in the future (non-anticipativity). We must therefore forecast the probabilities into the future by taking the weighted average of both cases.

$$\begin{aligned}\boldsymbol{\mu}^{+}(t_1) &= \boldsymbol{\mu}^{+}_{\text{Case I}}(t_1)\cdot\mu^{-}_{D}(t_1) && (10)\\ &+ \boldsymbol{\mu}^{+}_{\text{Case II}}(t_1)\cdot\left(1-\mu^{-}_{D}(t_1)\right) && (11)\\ &= \begin{bmatrix}\mu^{-}_{A}(t_1)+\mu^{-}_{D}(t_1)\\ \mu^{-}_{B}(t_1)\\ \mu^{-}_{C}(t_1)\\ 0\end{bmatrix} && (12)\end{aligned}$$

Or, in matrix notation:

$$\begin{aligned}\boldsymbol{\mu}^{+}(t_1) &= \boldsymbol{M}\cdot\exp\left(\boldsymbol{\Lambda}_0\left(t_1-t_0\right)\right)\cdot\boldsymbol{\mu}(t_0) && (13)\\ &= \boldsymbol{M}\cdot\boldsymbol{\mu}^{-}(t_1) && (14)\end{aligned}$$

where we further decompose $\boldsymbol{M}$ as

$$\boldsymbol{M} = (\boldsymbol{I} + \boldsymbol{R}\boldsymbol{S}^{\mathsf{T}}) \tag{15}$$

where $\boldsymbol{I}$ is the identity matrix, $\boldsymbol{R}$ is the repair matrix, and $\boldsymbol{S}$ is a selection matrix. The selection matrix chooses the failed state D.

$$\boldsymbol{S} = [0\ 0\ 0\ 1]^{\mathsf{T}} \tag{16}$$

For AGAN repairs, we have

$$\boldsymbol{R} = [1\ 0\ 0\ {-1}]^{\mathsf{T}}\,. \tag{17}$$

In Section 2.2, we will come back to why the decomposition of $\boldsymbol{M}$ into $\boldsymbol{R}$ and $\boldsymbol{S}$ is useful.

The evolution of the probabilities for case 1 and 2, and the weighted average of the two cases is illustrated in Figure 2.

Generalization

Using the general expression for the probabilities between two maintenance stops from Eq. (5), we can express the probabilities at any time t as the piece-wise function

$$\boldsymbol{\mu}(t) = \begin{cases}\exp\left(\boldsymbol{\Lambda}\cdot(t-t_0)\right)\cdot\boldsymbol{\mu}^{+}(t_0) & \text{if} \quad t < t_1\\ \exp\left(\boldsymbol{\Lambda}\cdot(t-t_1)\right)\cdot\boldsymbol{\mu}^{+}(t_1) & \text{if} \quad t > t_1\end{cases} \tag{18}$$

where

$$\boldsymbol{\mu}^{+}(t_0) = \boldsymbol{\mu}(t_0) = \boldsymbol{\mu}_0 = [1\ 0\ 0\ 0]^{\mathsf{T}} \tag{19}$$

is the specified initial condition.

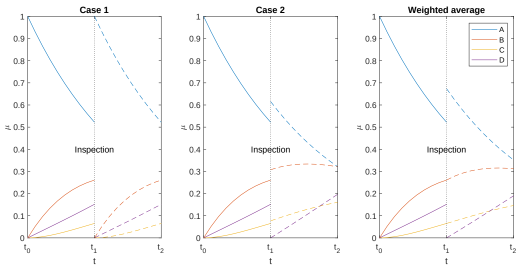

Fig. 2. Illustration of evolution of the probabilities μ before and after inspection. Left: system is found to be in the failed state D upon inspection, and is repaired without time lag. Middle: system is found to be in working order upon inspection, and no repairs are performed. Right: weighted average of the two previous cases.

Taking it one step further, we can have an arbitrary amount of inspections and maintenances, k, between two times t_0 and t_f and express the probabilities as

$$\boldsymbol{\mu}(t) = \begin{cases}\exp\left(\boldsymbol{\Lambda}\cdot(t-t_0)\right)\cdot\boldsymbol{\mu}^{+}(t_0)\\ \quad \text{if} \quad t_0 < t < t_1\\ \exp\left(\boldsymbol{\Lambda}\cdot(t-t_1)\right)\cdot\boldsymbol{\mu}^{+}(t_1)\\ \quad \text{if} \quad t_1 < t < t_2\\ \cdots\\ \exp\left(\boldsymbol{\Lambda}\cdot(t-t_k)\right)\cdot\boldsymbol{\mu}^{+}(t_k)\\ \quad \text{if} \quad t_k < t < t_f\end{cases} \tag{20}$$

where

$$\begin{aligned}\boldsymbol{\mu}^{+}(t_i) &= \boldsymbol{M}\cdot\boldsymbol{\mu}^{-}(t_i) && (21)\\ \boldsymbol{\mu}^{-}(t_i) &= \exp\left(\boldsymbol{\Lambda}\cdot(t_i-t_{i-1})\right)\cdot\boldsymbol{\mu}^{+}(t_{i-1}) && (22)\end{aligned}$$

2.2. *Differentiating to get the differential model*

Equation (1) described the evolution of the state $\boldsymbol{\mu}$ between two inspection times. However, as we have shown, we can express the state at any given time by the piece-wise model from equation (20). If we differentiate (20), we get

$$\frac{d\boldsymbol{\mu}}{dt} = \boldsymbol{\Lambda}\cdot\boldsymbol{\mu}(t) + \boldsymbol{R}\boldsymbol{S}^{\mathsf{T}}\boldsymbol{\mu}(t)\cdot\left(\sum_{i=1}^{k}\delta(t-t_i)\right) \tag{23}$$

where δ is the Dirac delta function.

Let us now introduce the variable

$$\boldsymbol{r}(t) = \boldsymbol{S}^{\mathsf{T}}\boldsymbol{\mu}(t)\cdot\left(\sum_{i=1}^{k}\delta(t-t_i)\right) \tag{24}$$

to obtain the form

$$\frac{d\boldsymbol{\mu}}{dt} = \boldsymbol{\Lambda}\cdot\boldsymbol{\mu}(t) + \boldsymbol{R}\cdot\boldsymbol{r}(t). \tag{25}$$

We choose to work with this form of the reliability model as it allows us to distinguish between the degradation of the system due to aging and unforeseen failures (first term of Eq. (25)), and the maintenance of the system (second term of Eq. (25)). Our aim is to do numerical optimization of the maintenance times, which can be achieved by optimizing the breakpoints t_i of the function $\boldsymbol{r}(t)$. More on this in Section 3.

Furthermore, we can easily change the maintenance strategy from as-good-as-new (AGAN) to as-bad-as-old (ABAO) by changing $\boldsymbol{R}$

$$\boldsymbol{R}_{\text{AGAN}} = [1\ 0\ 0\ -1]^{\intercal} \tag{26}$$

$$\boldsymbol{R}_{\text{ABAO}} = [0\ 0\ 1\ -1]^{\intercal} \tag{27}$$

An illustration of how $\boldsymbol{\mu}(t)$ changes as a function of $\boldsymbol{r}(t)$ is shown in Fig. 3. Note that $\boldsymbol{r}(t)$ is a sum of Dirac functions. In addition, we show the integral $\int_0^{t_f} \boldsymbol{r}(t)$, which is proportional to the maintenance cost.

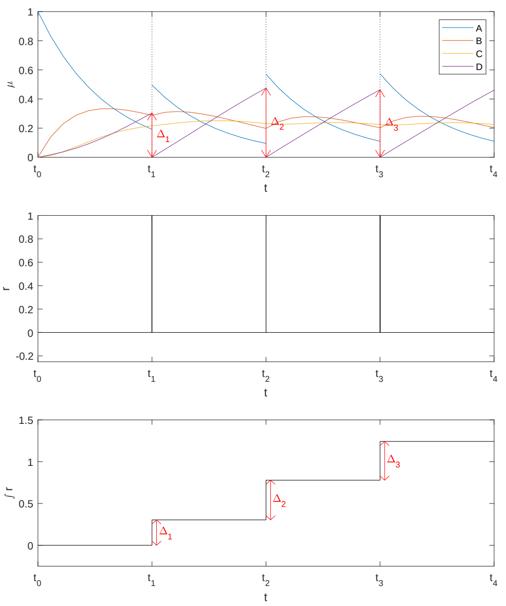

Fig. 3. Illustration of how $\boldsymbol{r}(t)$ influences $\boldsymbol{\mu}(t)$. The cumulative maintenance cost is proportional to the integral of $\boldsymbol{r}$, shown in the bottom plot, while the inspection cost is proportional to number of spikes (three, in this case)

2.3. *Modeling the effect of inputs*

The second contribution of this paper is the inclusion of the effect of inputs $\boldsymbol{u}(t)$, which influence the degradation rate of the system. Thus by changing $\boldsymbol{u}(t)$, we can actively steer the rate of degradation of the system. This is very useful, as it allows us to optimize the performance of the system by co-optimizing $\boldsymbol{u}(t)$ and $\boldsymbol{r}(t)$. We model this behavior by letting $\boldsymbol{\Lambda}_a$ be a function of the inputs and time, instead of a constant matrix like before. $\boldsymbol{\Lambda}_u$ remains constant, as we assume that unexpected failures cannot be influenced by changing the inputs. The differential model now is

$$\frac{d\boldsymbol{\mu}}{dt} = (\boldsymbol{\Lambda}_a(\boldsymbol{u}, t) + \boldsymbol{\Lambda}_u) \cdot \boldsymbol{\mu}(t) + \boldsymbol{R} \cdot \boldsymbol{r}(t). \tag{28}$$

Note that if $\boldsymbol{\Lambda}_a(\boldsymbol{u}, t)$ is a piece-wise constant function, meaning we can write it as

$$\boldsymbol{\Lambda}_a(\boldsymbol{u}, t) = \begin{cases} \boldsymbol{\Lambda}_{a,1} & \text{if} \quad t_0 < t < t_1 \\ \boldsymbol{\Lambda}_{a,2} & \text{if} \quad t_1 < t < t_2 \\ \dots & \\ \boldsymbol{\Lambda}_{a,k} & \text{if} \quad t_k < t < t_f \end{cases}, \tag{29}$$

the system becomes a Multiphase Markov process. However, we do not require this assumption, and are free to choose whatever form of $\boldsymbol{\Lambda}_a(\boldsymbol{u}, t)$ we need.

3. Formulating the optimization problem

In order to find the optimal combined production and maintenance strategy, we first formulate an optimization problem in terms of an objective function and constraints. [a]

$$\min_{\boldsymbol{u},\boldsymbol{r}} \quad \int_{t_0}^{t_f} \Big(- f(t, \boldsymbol{\mu}, \boldsymbol{u}) \Big) dt$$

$$+ f_i(t, \boldsymbol{\mu}, \boldsymbol{r}) + f_m(t, \boldsymbol{\mu}, \boldsymbol{r}) \tag{30a}$$

$$\text{s.t.} \quad \frac{d\boldsymbol{\mu}}{dt} = \boldsymbol{\Lambda}(t, \boldsymbol{u}) \cdot \boldsymbol{\mu}(t) + \boldsymbol{R} \cdot \boldsymbol{r}(t) \tag{30b}$$

$$\boldsymbol{r}(t) = \boldsymbol{S}^{\intercal} \boldsymbol{\mu}(t) \cdot \left(\sum_{i=1}^{k} \delta(t - t_i) \right) \tag{30c}$$

$$0 \leq \boldsymbol{\mu} \leq 1 \tag{30d}$$

$$\sum_{i \in \{A,B,C,D\}} \mu_i = 1 \tag{30e}$$

$$0 \leq \boldsymbol{r} \leq \infty \tag{30f}$$

$$\boldsymbol{u}_{\min} \leq \boldsymbol{u} \leq \boldsymbol{u}_{\max} \tag{30g}$$

In the above optimization problem, f denotes some economical objective which is to be maximized (typically profit or production), f_i denotes

[a] Note that the constraint in Eq. (30e) is implied by Eq. (30b), but we include it for completeness.

the inspection cost, which is typically proportional to the number of inspections k,

$$f_i(t, \boldsymbol{\mu}, \boldsymbol{u}) \propto k, \tag{31}$$

and f_m denotes the maintenance cost, which is assumed to be proportional to the integral of $\boldsymbol{r}(t)$

$$f_m(t, \boldsymbol{\mu}, \boldsymbol{u}) \propto \int_{t_0}^{t_f} \boldsymbol{r}(t)dt. \tag{32}$$

3.1. *Problem re-formulation for numerical optimization*

The optimization problem from Eqs. (30a)-(30g) can be solved in a multitude of ways. A common approach is to approximate the dynamic problem by a static non-linear programming (NLP) problem through the use of so-called direct methods, where direct multiple shooting and direct collocation are common approaches (Biegler, 2010). In this work, we use the direct collocation approach.

One issue with (30) is in $\boldsymbol{r}(t)$. Since it is the summation of Dirac functions, it is unbounded as shown in Eq. (30f). An alternative to using the formulation adopted in this paper is to formulate the problem as a mixed integer problem, as was done in Ashayeri et al. (1996). Due to the nonlinear nature of the problem, we have to solve a mixed-integer non-linear programming (MINLP) problem, as done in e.g. And and Grossmann (1998); Georgiadis and Papageorgiou (2000). These MINLP problems are however known to be very difficult to solve in the general case, despite recent progress in algorithmic development.

Instead, we approximate $\boldsymbol{r}(t)$ using Boxcar functions as

$$\boldsymbol{r}(t) \approx \tilde{\boldsymbol{r}}(t) \tag{33}$$

$$\tilde{\boldsymbol{r}}(t) = \sum_{i=1}^{k} \text{Boxcar}(t) \tag{34}$$

$$= \sum_{i=1}^{k} h_i \Big(\mathrm{H}(t - t_i) - \mathrm{H}(t - t_i - \epsilon_i) \Big) \tag{35}$$

where H is the Heaviside function, h_i is the height and ϵ_i is the width of each "box".

An illustration of this approximation is shown in the middle plot in Fig. 4. Note that the approximation for $\boldsymbol{\mu}$ is good if ϵ is sufficiently small, and that the approximation gives the same cumulative maintenance cost $\int \tilde{\boldsymbol{r}}(t)dt$. Furthermore, we observe that $\boldsymbol{\mu}$ is now continuous (although nonsmooth), which makes the optimization problem easier to solve.

Another numerical issue is posed by the inspection cost from Eq. (31). In the original

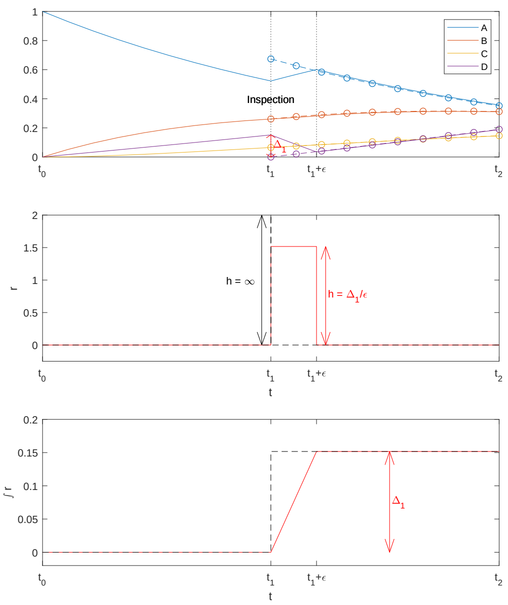

Fig. 4. Illustration of how $\boldsymbol{r}(t)$ (dashed line, circles) can be approximated by $\tilde{\boldsymbol{r}}(t)$ (solid line) to obtain a continuous $\boldsymbol{\mu}$.

formulation with Dirac functions, one might be tempted to find k as

$$k = \int_{t_0}^{t_f} \Big(\sum_{i=1}^{k} \delta(t - t_i) \Big) dt \tag{36}$$

$$= \int_{t_0}^{t_f} \Big(\big(\boldsymbol{S}^\intercal \boldsymbol{\mu}(t)\big)^{-1} \cdot \boldsymbol{r}(t) \Big) dt, \tag{37}$$

but for this to work, we must assert that

$$\mu_D(t_i) = \mu_D^-(t_i) \tag{38}$$

to avoid division by zero. Such a condition might be difficult to impose numerically.

Instead, we propose to solve the problem by introducing the additional variable $\boldsymbol{y}$, which we use to formulate additional constraints:

$$0 \leq (1 - \boldsymbol{y}) \perp \tilde{\boldsymbol{r}} \geq 0 \tag{39}$$

$$0 \leq \boldsymbol{y} \leq 1 \tag{40}$$

Here, the $\perp$ operator indicates complementary, i.e. we require that at all times either $\tilde{\boldsymbol{r}}$ or $(1 - \boldsymbol{y})$ or both are zero. The inspection cost can then be written as

$$f_i(t, \boldsymbol{\mu}, \boldsymbol{u}) \propto \frac{\boldsymbol{y}}{\epsilon} \tag{41}$$

In order to minimize the cumulative inspection cost, $\boldsymbol{y}(t)$ will be a function that is either at its lower bound (zero) when no inspection is performed, or at its upper bound (one) when inspection is performed. By integrating we get

$$k = \int_{t_0}^{t_f} \Big(\sum_{i=1}^{k} \delta(t - t_i) \Big) dt \approx \int_{t_0}^{t_f} \frac{\boldsymbol{y}}{\epsilon} dt \qquad (42)$$

4. Case study

As a case example, we consider a system inspired by subsea oil and gas production. Subsea technology is key to satisfying the energy demands of tomorrow, due to the intermittent nature of renewables and the continued need for petroleum products also in a green society. Reliability is a major issue for subsea installations, as maintenance interventions are very costly. Consequently, it is important to optimize both production from the subsea installation, as well as the maintenance interventions.

Assume that production from the subsea installation actively degrades critical components such as pumps, valves and heat exchangers. The transition rates can therefore be assumed to be proportional to the inputs $\boldsymbol{u}$ that we apply to the system. In our case, $\boldsymbol{u}(t)$ represents the production rate of oil and gas. A higher production rate will give more immediate profit, but also increased degradation.

Poor instrumentation and a lot of measurement uncertainty mean that a system may not be properly diagnosed to have failed without inspection. An example of this could be the failure of a single well going to a manifold with several other wells. The failure of the single well may be masked by the large variability in production of the other wells. Well tests (which can be thought of as inspections) are thus required to reveal the state of the single well.

4.1. *Objective function*

The objective is to maximize the average production from the well over the lifetime of the field, while simultaneously minimizing the inspection and maintenance costs. This economical objective can be written as

$$\min_{\boldsymbol{u}(t), \tilde{\boldsymbol{r}}(t)} \int_{0}^{t_f} \Big(\frac{-f + f_m + f_i}{(1+d)^t} \Big) dt \qquad (43a)$$

where

$$f(t) = \boldsymbol{c}_p^{\mathsf{T}} \cdot \boldsymbol{\mu}(t) \cdot \boldsymbol{u}(t) \qquad (43b)$$

$$f_m(t) = \boldsymbol{c}_m^{\mathsf{T}} \cdot \tilde{\boldsymbol{r}}(t) \qquad (43c)$$

$$f_i(t) = \boldsymbol{c}_i^{\mathsf{T}} \cdot \boldsymbol{y}(t). \qquad (43d)$$

Here, $\boldsymbol{c}_p$ is the productivity in each state, $\boldsymbol{c}_m$ is the maintenance cost, $\boldsymbol{c}_i$ is the inspection cost. The entire economic objective is discounted by a factor d to reflect the decreasing value of future income streams compared to present income streams. Note that the objective is non-linear due to the bi-linear term in $f(t)$.

4.2. *Constraints*

The objective function is optimized subject to the following constraints

$$s.t. \quad \frac{d\boldsymbol{\mu}}{dl} = \boldsymbol{\Lambda}(t, \boldsymbol{u})\boldsymbol{\mu}(t) + \boldsymbol{R}\tilde{\boldsymbol{r}}(t) \qquad (43e)$$

$$\boldsymbol{\Lambda}(t, \boldsymbol{u}) = \boldsymbol{\Lambda_a} \cdot \boldsymbol{u}(t) + \boldsymbol{\Lambda}_u \qquad (43f)$$

$$\boldsymbol{\mu}(0) = \boldsymbol{\mu}_0 \qquad (43g)$$

$$0.1 \leq \boldsymbol{u} \leq 1.0 \qquad (43h)$$

$$0 \leq \tilde{\boldsymbol{r}} \leq \frac{1}{\epsilon_{\min}} \qquad (43i)$$

$$\epsilon_{\min} \leq \boldsymbol{\epsilon} \leq \epsilon_{\max} \qquad (43j)$$

$$0 \leq \boldsymbol{\mu} \leq 1 \qquad (43k)$$

$$0 \leq (1 - \boldsymbol{y}) \perp \tilde{\boldsymbol{r}} \geq 0 \qquad (43l)$$

$$0 \leq \boldsymbol{y} \leq 1 \qquad (43m)$$

Note that the transition matrix $\boldsymbol{\Lambda}(t, \boldsymbol{u})$ is linear in $\boldsymbol{u}$. Since we require $\boldsymbol{u}(t)$ to be a piece-wise constant function, $\boldsymbol{\Lambda}(t, \boldsymbol{u})$ is also a piece-wise function. Consequently we are dealing with a Multiphase Markov process, as discussed in Section 2.3.

The parameters for the problem are summarized in Table 1.

Table 1. Parameters used for the optimization

Parameter	Description	Value
λ_u	Sudden failure transition rate	10^{-4}
λ_a	Base aging transition rate	10^{-2}
d	Discount rate	.001
$\boldsymbol{c}_p$	Productivities in each state	$[28\ 21\ 14\ 2.8]^{\mathsf{T}}$
$\boldsymbol{c}_m$	Maintenance cost	300
$\boldsymbol{c}_i$	Inspection cost	30
t_f	Final time	200 weeks

4.3. *NLP formulation*

Multiple approaches exist to solve dynamic problems like (43), but we choose to use orthogonal collocation on finite elements. The original dynamic problem is reformulated as an NLP, which can be solved using standard non-linear optimization algorithms. We will not go into details about how to discretize the problem, see e.g. Biegler (2010) for a summary.

The resulting NLP is implemented in MATLAB using Casadi 3.4.1 (Andersson et al., 2018). The interior-point solver IPOPT 3.12.3 (Wächter and Biegler, 2006) is used to solve the optimization problem.

4.4. *Solution strategy*

Our problem is non-convex and local solvers such as IPOPT will consequently only find local solutions. In other words, we cannot guarantee global optimality of the solution. In order to ensure global optimality, global solvers such as BARON (Sahinidis, 2017) have to be used. Global solvers come with some drawbacks, such as being computationally intractable for large problems.

To remedy this, we use a multi-start approach where the problem is repeatedly re-optimized with different initial guesses. After a certain number of optimizations (1000 in this case), the best local minimum is returned. Fewer than 1000 optimization runs could suffice, but as they are computationally cheap, we choose to run 1000 to ensure that a good solution is obtained.

4.5. *Solution*

The optimized production and maintenance strategy is shown in Fig. 5. As can be seen, the optimal strategy is to operate at maximum $\boldsymbol{u}$ all the time (a typical property of almost-linear optimization problems), while inspections / maintenance is performed at $t = 88$, 127 and 160 weeks. The objective function value is 3924 M\$. Note that the first inspection is quite late. The reason for this is that the probabilities of being in the degraded states are initially low. As a result, performing inspections and maintenance at an early stage in the race is sub-optimal.

In an actual implementation, one would re-optimize the problem upon obtaining new information about the system state (such as after an inspection). This is known as model-predictive control (MPC) or rolling horizon optimization (Biegler, 2010). However, the optimization problem itself remains the same with only the initial conditions (43g) changing. We therefore chose to skip the closed-loop results in the interest of time and space.

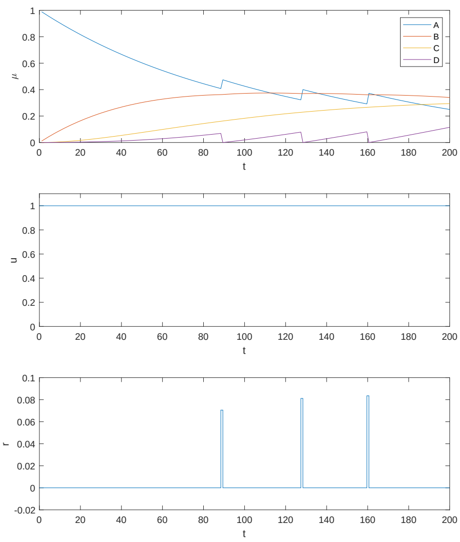

Fig. 5. Optimal solution of the case example

5. Conclusions and future work

In this paper, we have introduced a method for simultaneous production- and maintenance optimization. The problem was motivated by a Markov-chain representation of a degrading production system, which would have been difficult to optimize using the traditional Monte Carlo based approach. We reformulated the problem as a algebraic-differential equation system, which we solved using a non-linear optimization approach. While not all problems can be solved like this, we showed how for the specific problem at hand, the problem could be cast into a form which could be solved using off-the-shelf solvers. The concept of input-dependent transition rates can easily be included in the framework. Some approximations were introduced to make the problem numerically tractable. The method was demonstrated on a case example inspired by subsea oil and gas production.

Possible future research directions include:

- Detailed comparison to Monte Carlo-based methods for optimization.
- Inclusion of more maintenance strategies by modification of $\boldsymbol{R}$ and optimization of the trade-off between the different maintenance strategies.
- A distributionally robust problem formulation to safeguard the solution against uncertainties in the transition rates.
- A multi-step approach to include the value of future information in the open-loop optimization problem.
- Looking at the case where maintenance is not instantaneous, i.e. when there is lag-time.
- Analysis of the losed-loop performance.
- A more complex case study with multiple simultaneously degrading units

Acknowledgments

This work is funded by the SUBPRO center for research based innovation, `www.ntnu.edu/subpro`. Johannes Jäschke and Anne Barros acknowledge support by DNV-GL.

References

And, V. J. and I. E. Grossmann (1998). Cyclic scheduling of continuous parallel-process units with decaying performance. *AIChE Journal 44*(7), 1623–1636.

Andersson, J. A. E., J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl (In Press, 2018). CasADi – A software framework for nonlinear optimization and optimal control. *Mathematical Programming Computation*.

Ashayeri, J., A. Teelen, and W. Selenj (1996). A production and maintenance planning model for the process industry. *International journal of production research 34*(12), 3311–3326.

Biegler, L. T. (2010). *Nonlinear programming: concepts, algorithms, and applications to chemical processes*, Volume 10. Siam.

Fitouhi, M.-C. and M. Nourelfath (2012). Integrating noncyclical preventive maintenance scheduling and production planning for a single machine. *International Journal of Production Economics 136*(2), 344–351.

Georgiadis, M. and L. Papageorgiou (2000). Optimal energy and cleaning management in heat exchanger networks under fouling. *Chemical Engineering Research and Design 78*(2), 168–179.

Iravani, S. M. and I. Duenyas (2002). Integrated maintenance and production control of a deteriorating production system. *Iie Transactions 34*(5), 423–435.

Sahinidis, N. V. (2017). *BARON 17.8.9: Global Optimization of Mixed-Integer Nonlinear Programs,* User's Manual.

Wächter, A. and L. T. Biegler (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical programming 106*(1), 25–57.

Wolter, A. and S. Helber (2016). Simultaneous production and maintenance planning for a single capacitated resource facing both a dynamic demand and intensive wear and tear. *Central European Journal of Operations Research 24*(3), 489–513.