

A Production Model with History Based Random Machine Failures



Stephan Knapp and Simone Göttlich

Abstract In this paper, we introduce a time-continuous production model that enables random machine failures, where the failure probability depends historically on the production itself. This bidirectional relationship between historical failure probabilities and production is mathematically modeled by the theory of piecewise deterministic Markov processes (PDMPs). On this way, the system is rewritten into a Markovian system such that classical results can be applied. In addition, we present a suitable solution, taken from machine reliability theory, to connect past production and the failure rate. Finally, we investigate the behavior of the presented model numerically in examples by considering sample means of relevant quantities and relative frequencies of number of repairs.

1 Modeling Equations

We briefly recall the production network model from [1, 6] first, and according to [4], we present the stochastic extension to a load-dependent production model with machine failures. To keep the notation well-arranged, we consider a production network consisting of a single queue processor unit. We assume a processor, which is represented by an interval $(a, b) \subset \mathbb{R}$, i.e., with length $L = b - a$, where $\rho(x, t)$ describes the density of production goods at $x \in (a, b)$ and time $t \geq 0$. The dynamics of the density, and consequently of the production, is given by the following nonlinear hyperbolic partial differential equation

$$\partial_t \rho(x, t) + \partial_x \min\{v\rho(x, t), c\} = 0, \quad (1)$$

where $c \geq 0$ is the production capacity and $v > 0$ the constant production velocity. In front of the processor a storage, also called queue, is assumed and for an

S. Knapp (✉) · S. Göttlich
University of Mannheim, Mannheim, Germany
e-mail: stknapp@mail.uni-mannheim.de; goettlich@uni-mannheim.de

externally given time-dependent inflow $G_{\text{in}}(t)$ into the production, the queue length q follows the ordinary differential equation

$$\partial_t q(t) = G_{\text{in}}(t) - g_{\text{out}}(t), \tag{2}$$

with

$$g_{\text{out}}(t) = \begin{cases} \min\{G_{\text{in}}(t), c\}, & \text{if } q(t) = 0, \\ c, & \text{if } q(t) > 0. \end{cases}$$

The processor is coupled to the queue by a boundary condition in the form of $\rho(a, t) = \frac{g_{\text{out}}(t)}{v}$ and initial conditions $\rho(x, 0) = \rho_0(x) \in L^1((a, b))$, $q(0) = q_0 \in \mathbb{R}_{\geq 0}$ are prescribed. This deterministic model is well-defined, see, e.g. [1]. The theory of piecewise deterministic Markov processes; see, e.g. [2, 7], has been used to define an appropriate production model with stochastic machine failures in [4], where the probabilities of machine failures depend on the actual workload of the processor. Since this construction only allows for a dependence on the current workload, we can not use the amount of goods produced since the last machine failure as a measure for the next failure. Our new idea lies in adding a variable w governing the workload since the last repair. To do so, we use the time-dependent variable $r(t) \in \{0, 1\}$, and set the capacity as $\mu(t) = r(t)c$ for a maximal capacity $c > 0$. This means that $r(t) = 0 \Rightarrow \mu(t) = 0$ is a down and $r(t) = 1 \Rightarrow \mu(t) = c$ a working processor at time t and we define

$$\text{WIP}(t_0, t_1) = \int_{t_0}^{t_1} \int_a^b \rho(x, t) dx dt$$

as the cumulative work-in-progress of the processor between time t_0 and t_1 . The variable w should therefore satisfy

$$\partial_t w(t) = r(t) \int_a^b \rho(x, t) dx, \quad w(t_0) = w_0 = \int_a^b \rho(x, t_0) dx. \tag{3}$$

Altogether, we define the state space

$$E = \mathbb{R}_{\geq 0} \times \{0, 1\} \times \mathbb{R}_{\geq 0} \times L^1((a, b)),$$

which is a measurable space together with the σ -algebra \mathcal{E} generated by the open sets induced by the metric

$$d((w, r, q, \rho), (\tilde{w}, \tilde{r}, \tilde{q}, \tilde{\rho})) = |w - \tilde{w}| + |r - \tilde{r}| + |q - \tilde{q}| + \|\rho - \tilde{\rho}\|_{L^1((a,b))}.$$

Since we construct a piecewise deterministic Markov process, we define the deterministic dynamics between jump times as

$$\Phi_{st} : E \rightarrow E, \quad (w_0, r_0, q_0, \rho_0) \mapsto (w(t), r(t), q(t), \rho(t)),$$

i.e., Φ_{st} is the solution to Eqs. (1), (2), (3), and $r(t) = r_0$ with initial conditions $(w_0, r_0, q_0, \rho_0) \in E$. In between the jump-times, where the capacity changes, we have a capacity, which is given by cr_0 and independent of time. This allows us to apply the theory of the deterministic model (1)–(2) to obtain continuity properties of Φ . To characterize the stochastic part, we introduce

$$\psi(t, y) = \lambda_{r,r}(t, w), \quad \eta(t, y, B) = \frac{\lambda_{r,(1-r)}(t, w)}{\psi(t, y)} \epsilon_{(r w, (1-r), q, \rho)}(B)$$

for every $y = (w, r, q, \rho) \in E$ and $B \in \mathcal{E}$, where $\lambda_{i,j}(t, w)$ describes the transition rate from capacity i to j at time t and actual workload w , $i, j \in \{0, 1\}$ and ϵ_x is the Dirac measure with unit mass in x . The function ψ is the total intensity determining whether a jump occurs, or not, and the function η describes the probability distribution of the systems jump given the system changes at time t . For example, given the state $y = (w, 1, q, \rho)$ at the time of a jump, the system jumps to $(w, 0, q, \rho)$ and, vice versa, given the state $y = (w, 0, q, \rho)$ the system jumps to $(0, 1, q, \rho)$, i.e., the workload has been “reset”. The open question is whether this model can be represented by a piecewise deterministic Markov process. Following [4], it is straightforward to show

Theorem 1 *Let $\lambda_{i,j} : [0, T] \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ be uniformly bounded, continuous and satisfy $\lambda_{i,i} = \lambda_{i,i-1}$ for $i \in \{0, 1\}$. Then for all initial data $x_0 \in E$ there exists a Markov process*

$$X = ((w(t), r(t), q(t), \rho(r)), t \in [0, T]) \subset E$$

on some probability space (Ω, \mathcal{A}, P) , satisfying

1. $X(0) = x_0$ P -almost surely,
2. for every $t \in (0, T)$, $(w, r, q, \rho) \in E$ and $j \in \{0, 1\}$, it holds that

$$P(r(t + \Delta t) = j | X(t) = (w, r, q, \rho)) = (1 - \Delta t \lambda_{r,r}(t, w)) \mathbb{1}_r(j) + \Delta t \lambda_{r,(1-r)}(t, w) \mathbb{1}_{1-r}(j) + o(\Delta t),$$

3. there exists a P -null set $\mathcal{N} \in \mathcal{A}$ such that for every $\omega \in \Omega \setminus \mathcal{N}$, there exist times $T_0 = 0 \leq T_1 \leq \dots \leq T_M = T$ such that for every $k = 0, \dots, M - 1$, $X(t) = \Phi_{T_k,t}(X(T_k))$ for $t \in [T_k, T_{k+1})$ with capacity $\mu(r(T_k, \omega))$, i.e., X behaves deterministic between jump times.

The main and new ingredient is the mapping $t \mapsto w(t)$, which is a continuous mapping since $t \mapsto \rho(t)$ is continuous.

2 Computational Results

Due to the fact that solutions to (1) move with non-negative velocities only, we can use the first order left-sided upwind scheme for a numerical approximation of the density ρ . Furthermore, we use the explicit Euler scheme to approximate the queue length q given by (2) and w given by (3), where we use a rectangular rule for the integration. This yields an approximation of the deterministic dynamics between the jump times. The simulation of the jump times is done with the thinning algorithm presented in [4]. Its basic idea is to use the uniform bound on the rate functions and generate exponentially distributed times with high intensity, representing the times between jumps, and thin these times during the numerical simulation of the whole system with an appropriate acceptance rejection procedure.

The choice of the rate functions $\lambda_{i,j}(t, w)$ is a crucial point in numerical examples. Here, we make use of the choice in [9] and set for $\theta_1, \theta_2 > 0$ the rate function as

$$\lambda_{1,0}(t, w) = \lambda_{1,0}^{\min} + (\lambda_{1,0}^{\max} - \lambda_{1,0}^{\min})(1 - e^{-(\theta_1 w)^{\theta_2}}),$$

which is a scaled version of the cumulative distribution function of a Weibull distribution, i.e., $F(t) = 1 - e^{-(\theta_1 t)^{\theta_2}}$. The classical interpretation of t in the latter expression is the lifetime of a machine and $F(t)$ is the probability that a failure happens after time t , see, e.g. [8]. In our case we use the variable w , which measures the amount of goods produced since the last repair happened. Therefore, if $w = 0$, then $\lambda_{1,0}(t, 0) = \lambda_{1,0}^{\min}$, which corresponds to the minimal failure rate and $\lim_{w \rightarrow \infty} \lambda_{1,0}(t, w) = \lambda_{1,0}^{\max}$. The function $\lambda_{1,0}(t, w)$ is monotonically increasing in w and incorporates the idea of an increasing failure rate depending on past workloads. On the other hand, we assume $\lambda_{0,1}(t, w) = \lambda_{0,1}$ because repair times do not depend on the amount of goods produced.

In the following, we examine the presented model using numerical examples. Here, we assume a production velocity of $v = 1$, the interval $a = 0, b = 1$, and the capacity is given as $\mu(t) = 2r(t)$. We use a spatial discretization with step-size $\Delta x = 10^{-1}$ and a temporal step-size that satisfies the Courant-Friedrichs-Lewy condition, which reads as $\Delta t \leq \Delta x$ for the chosen parameters. The simulation results are based on samples of the stochastic process X and we use the classical Monte-Carlo estimator to evaluate moments or probabilities of the samples. We used a sample size of 10^5 for all following results.

We analyze the expected queue length, capacity and the distribution of the number of repairs within a time horizon $[0, 50]$ for two different constant inflow profiles. We denote by $G_{\text{in}}^1(t) \equiv 0.5$ and by $G_{\text{in}}^2(t) \equiv 1.5$ as inflow profiles and use the parameters

$$\lambda_{0,1}(t, w) = \frac{1}{0.5}, \quad \lambda_{1,0}^{\min} = \frac{1}{10}, \quad \lambda_{1,0}^{\max} = \frac{1}{0.5}, \quad \theta_1 = \frac{1}{10}, \quad \theta_2 = 5.$$

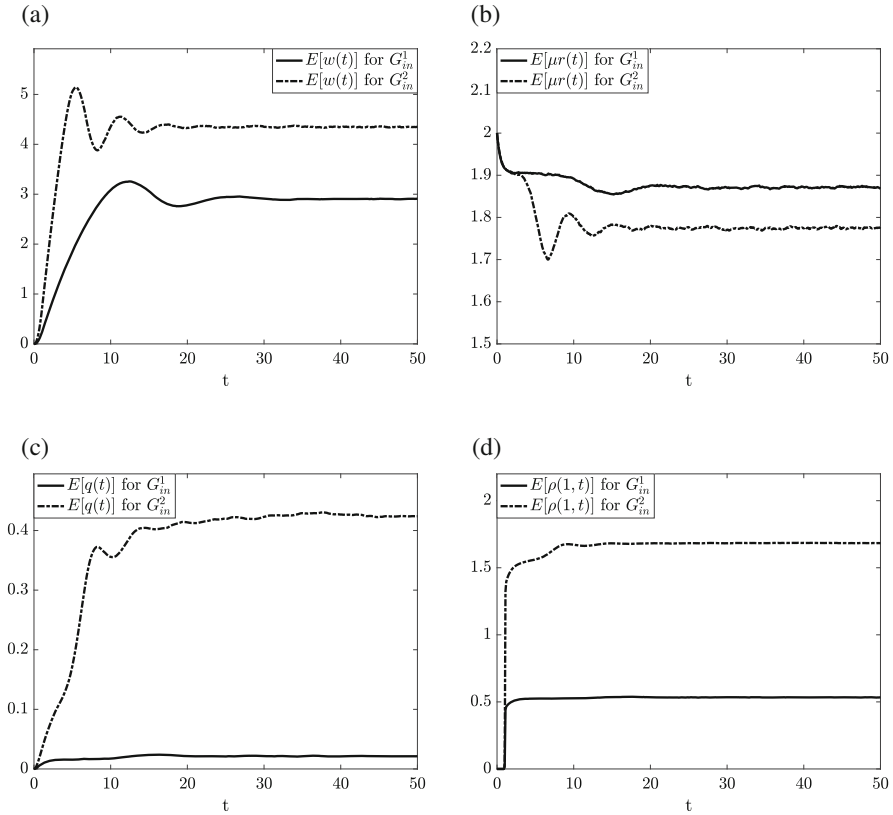


Fig. 1 First order moments of w , the capacity, queue-length and density. (a) Expected w . (b) Expected capacity $\mu(t)$. (c) Expected queue-length $q(t)$. (d) Expected density at $x = 1$

In Fig. 1, first order moment estimations are shown. In detail, Fig. 1a shows the expected value of the variable w , Fig. 1b the expected capacity, Fig. 1c the expected queue length and Fig. 1d the expected density at the end of the processor. The dynamics is quite interesting: the expected capacity decreases approximately until time $t = 6$ for the second inflow, then increases and decreases again. Indeed, the mean time to failure is given by $\Gamma(1 + \frac{1}{\theta_2})\theta_1^{-1}$, see e.g. [8]. If w corresponds to the lifetime in our model, we see that an intact system with constant inflow G_{in} is more likely to fail around time $\Gamma(1 + \frac{1}{\theta_2})(\theta_1 G_{in})^{-1}$. In our case, this leads to time 18.4 for the first and time 6.1 for the second inflow profile, which is close to the times at which the shape of the expected graphs in Fig. 1 changes. We observe these characteristic times also in the other graphs in Fig. 1. In contrast to the models presented in [3–5], where quantities monotonically converge, we obtain an oscillatory behavior of the quantities for constant inputs. The oscillatory effects are natural and caused by the history we incorporate in w . This means, the first machine failures are likely around time 18.4(6.1), the second around 36.8(12.2) and so on. At the same time

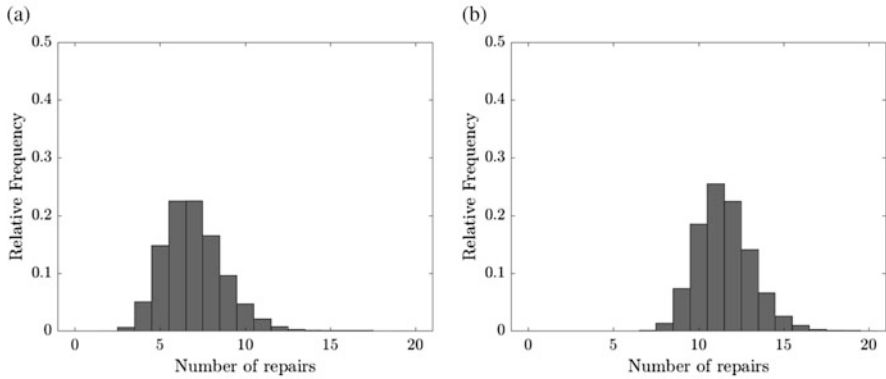


Fig. 2 Distribution of the number of repairs within $[0, 50]$. **(a)** Inflow G_{in}^1 . **(b)** Inflow G_{in}^2

the failures, which occur between these likely times, smooth this effect out as time evolves and the quantities converge.

Figure 2 shows the distribution of the number of repairs within the time horizon $[0, 50]$ and emphasizes the impact of the chosen inflow on the reliability of the processor. In Fig. 2a the case of G_{in}^1 is shown, where mostly 5–9 repairs have been done. The situation for inflow profile G_{in}^2 is different, where 9–14 repairs during the time horizon are more likely.

To conclude, we deduced a production model with random machine failures including failure probabilities depending on the workload of the machine since the last repair occurred. The extension of the model to complex production networks is straightforward, see, e.g. [4]. Simulation results showed a big impact of the history on expected workload, capacity, queue length and density. These effects are not negligible for production planning and control and must be taken into account.

Acknowledgements This work has been financially supported by the BMBF project ENets (05M18VMA).

References

1. D’Apice, C., Göttlich, S., Herty, M., Piccoli, B.: Modeling, Simulation, and Optimization of Supply Chains. SIAM, Philadelphia (2010)
2. Davis, M.H.A.: Piecewise-deterministic Markov processes: a general class of nondiffusion stochastic models. *J. Roy. Stat. Soc. Ser. B* **46**, 353–388 (1984)
3. Göttlich, S., Knapp, S.: Semi-Markovian capacities in production network models. *Discrete Contin. Dyn. Syst. Ser. B* **22**, 3235–3258 (2017)
4. Göttlich, S., Knapp, S.: Load-dependent machine failures in production network models (2018). arXiv:1806.03091v1
5. Göttlich, S., Martin, S., Sickenberger, T.: Time-continuous production networks with random breakdowns. *Netw. Heterog. Media* **6**, 695–714 (2011)

6. Göttlich, S., Herty, M., Klar, A.: Network models for supply chains. *Commun. Math. Sci.* **3**, 545–559 (2005)
7. Jacobsen, M.: *Point Process Theory and Applications*. Birkhäuser, Boston (2006)
8. Jiang, R., Murthy, D.N.P.: A study of Weibull shape parameter: properties and significance. *Reliab. Eng. Syst. Saf.* **96**, 1619–1626 (2011)
9. Rivera-Gómez, H., Montaña-Arango, O., Corona-Armenta, J.R., Garnica-González, J., Hernández-Gress, E.S., Barragán-Vite, I.: Production and maintenance planning for a deteriorating system with operation-dependent defectives. *Appl. Sci.* **8**, 165 (2018)