Proceedings of the 2017 Winter Simulation Conference
W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.

INCORPORATING ELEMENTS OF A SUSTAINABLE AND DISTRIBUTED GENERATION SYSTEM INTO A PRODUCTION PLANNING MODEL FOR A WAFER FAB

Timm Ziarnetzky
Lars Mönch

Department of Mathematics and Computer Science
University of Hagen
Universitätsstraße 1
Hagen, 58097, GERMANY

Thulasi Kannaian
Jesus Jimenez

Ingram School of Engineering
Texas State University
601 University Drive
San Marcos, TX 78666, USA

ABSTRACT

In this paper, we consider elements of a sustainable and distributed generation system for a wafer fab. Wind turbines (WTs), solar photovoltaics (PVs), a substation with grid access, and a net metering system are included in the generation system. WTs and solar PVs have the highest priority in supplying the daily electricity of the wafer fab. Surplus energy can be returned to the main grid. The objective function of the production planning formulation contains production-related costs, cost for energy from the substation, and penalty costs when a renewable energy penetration is not reached. This cost can be reduced by offering renewable surplus energy to the main grid. The obtained production plans are executed in a simulation environment to compute the expected profit in the face of machine breakdowns, wind power volatility, and uncertain power output of the solar PVs. The approach allows determining an appropriate number of WTs and solar PVs for given demand scenarios. We present results of simulation experiments with the proposed model.

1 INTRODUCTION

Semiconductor manufacturing deals with producing integrated circuits (ICs) on silicon wafers, thin discs made from silicon or gallium arsenide. Four major stages namely wafer fabrication, sort, assembly, and final test (cf. Mönch et al. 2013), are required to manufacture ICs. The wafer fabrication part of the overall manufacturing process is carried out in semiconductor wafer fabrication facilities (wafer fabs). In wafer fabs, the electronic circuits are built up layer-by-layer onto the wafers. There are more than 40 layers for the most advanced technologies. After the wafers are processed in the wafer fab, they are sent to sort where electrical tests are used to identify any defective dices when packaged. After this stage, the probed wafers are transferred to assembly facilities where dices of appropriate quality are put into a package. Packaged dices are sent to test facilities where they are again tested to ensure that only high-quality products are delivered to the customers. Wafer fab and sort are often called front-end, while assembly and test are subsumed under the term back-end. The entire manufacturing process consists of up to 700 process steps and can take up to three months. It is a highly energy-intensive process, annual energy utility bills of up to \$10–20 million for a single wafer fab are common (cf. Villarreal et al. 2013), and are likely to increase in the future due to the expected larger wafer and fab sizes. Therefore, energy conservation efforts are highly desirable in this industry.

Production planning aims to determine release schedules for a single wafer fab, i.e., we are interested in determining which quantity of a certain product has to be launched in a certain period to meet given demand in the future while the profit has to be maximized. In the past, wafer fabs are mainly interested in reaching specific production-related goals. However, there is an increasing need to take into account sustainability issues when running wafer fabs. Up to now there is only a very small body of literature availa-

978-1-5386-3428-8/17/\$31.00 ©2017 IEEE

ble that deals with sustainability issues in semiconductor supply chains (cf. Mönch et al. 2017). This is especially true for integrating production planning with aspects of a sustainable and distributed generation (DG) system. In this paper, we present an integrated model that tries to maximize profit while ensuring that a specific amount of the overall wafer fab load is provided by renewable energy sources such as wind turbines (WTs) or solar photovoltaics (PVs). A deterministic planning approach is proposed that uses discrete-event simulation to assess the release schedules and the design decisions for the amount of renewable energy sources in a stochastic environment. The approach by Villarreal et al. (2013) is extended towards load profiles that are driven by the production planning-related decisions.

The paper is organized as follows. The problem is described and analyzed in Section 2. This includes the discussion of related work. We then present the proposed optimization approach in Section 3. The simulation environment is discussed in Section 4. The results of simulation experiments are presented in Section 5. Conclusions and future research directions are discussed in Section 6.

2 PROBLEM ANALYSIS

2.1 Problem Setting

We are interested in incorporating elements of a sustainable and DG system in production planning formulations for wafer fabs. The generation system must include

- wind turbines (WT)
- solar photovoltaics (PV)
- substation with grid access
- net metering.

WT and solar PV have the highest priority in supplying the daily electricity of the wafer fab. Additional electricity can be hauled from the substation in case of power shortage. If the power provided by WTs and solar PVs is larger than the load, the surplus energy can be returned to the main grid using a net metering system. We have to model the electricity load caused by manufacturing activities in the wafer fab and the power provided by the DG system. Both load and power are stochastic since the wafer fab, the wind power, and the solar radiation are stochastic.

We differentiate between the planning and the execution level. The planning level determines which quantity of a certain product should be released in which period of the planning horizon to minimize the sum of backlogging, finished goods inventory (FGI) holding, and work in process (WIP) costs. In addition, we introduce a cost term caused by using power provided by the substation. The amount of this energy is given as the difference of the overall load and the amount of power provided by renewable energy sources based on the number of WTs and solar PVs. The following assumptions are made on the planning level:

- The demand is deterministic.
- The energy consumption per lot of a certain product and per period is known.
- Using power provided by the substation in a certain period leads to additional cost.
- If the power provided by renewable energy sources is larger than the load in a given period, the surplus leads to a cost reduction.
- The power is a deterministic value in the planning model that is calculated based on the given number of WTs and solar PVs.

The expected profit is calculated based on release schedules that are executed in the base system, i.e. on the execution level. We consider:

- backlogging cost
- FGI holding cost
- WIP cost
- WT and solar PV equipment installation cost

- WT and solar PV operating and maintenance cost
- cost related to using power produced by the substation
- cost reduction by returning surplus energy to the main grid.

A certain amount of the overall load should be satisfied by renewable energy provided by WTs and PVs. A penalty term is used in the objective function if this constraint is not fulfilled. The load consists of a fixed load that results from running the wafer fab, for instance, for ensuring clean room conditions, and a load that depends on the number of WIP lots of a certain product in a certain period.

2.2 Related Work

We discuss related work with respect to sustainability issues in (semiconductor) supply chains. While there exists a lot of research related to energy consumption-aware scheduling models (cf. Giret et al. 2015 and Gahm et al. 2016 for recent survey papers), it seems that the literature considering medium and long-term planning models that take sustainability issues into account is somehow less developed. Various types of carbon emission constraints are considered in simple planning models for supply chains in Benjaafar et al. (2013). Masmoudi et al. (2015) present a lot sizing model for flow shops where energy constraints and an energy consumption-aware objective function are considered. While the latter two papers deal with short and mid-term planning problems, a multi-objective optimization model for supply network design is proposed by Wang et al. (2011) on the strategic level. Total costs and total carbon emission are considered as criteria. The normalized normal constraint method is used for generating the Pareto frontier. Related survey papers are presented by Duflou et al. (2012) and Biel and Glock (2016).

We are only aware of two papers that deal with semiconductor-specific models that include sustainability aspects. A sustainable and DG system for a wafer fab using simulation optimization to determine an appropriate combination of PVs and WTs for integrating renewable energy into a wafer fab in addition to the main grid under uncertain wind speed and solar irradiance is proposed by Villarreal et al. (2013). A stochastic programming model to deal with contract-based demand requests received by a wafer fab that owns onsite wind and solar generation units is discussed by Santana-Viera et al. (2015). A pay-in-advance scheme is assumed, i.e., the utility company offers a discounted electricity price to the participants during the contract period. Monte-Carlo simulation is applied to solve the resulting stochastic program. Both models assume that the electricity load is an exogenous quantity, i.e., they do not directly link the load to the production activities.

In the present paper, we will overcome this limitation by extending fixed lead time-based production planning formulations for a single wafer fab as proposed by Kacar et al. (2013), (2016) towards sustainability aspects that are taken from Villarreal et al. (2013). The proposed models are similar to the models by Masmoudi et al. (2015) and Benjaafar et al. (2013) with respect to scope and purpose.

3 OPTIMIZATION APPROACH

3.1 Planning Model

We consider a planning horizon of length T divided into discrete periods of equal length. The objective of the model is to determine the amount of each product to be released into the wafer fab so as to minimize the costs caused by these releases. Multiple machine types with finite capacity organized in work centers are considered. A linear programming (LP) formulation based on fixed, exogenous lead times is given as follows:

Sets and indices

G:	set of all products
K:	set of all work centers
t:	period index
g:	product index

k :	work center index
l:	operation index
$O(g)$:	set of all operations of product g
$O(k)$:	set of all operations performed on machines of work center k

Decision variables

Y_{gtl} :	quantity of product g completing its operation l in period t
Y_{gt} :	output of product g in period t from the last operation of its routing
X_{gt} :	quantity of product g released into the first work center in its routing in period t
W_{gt} :	WIP of product g at the end of period t
I_{gt} :	FGI of product g at the end of period t
B_{gt} :	backlog of product g at the end of period t
APS_t :	average amount of power provided by the substation or surplus energy sent back to the main grid in period t
RE_t :	average amount of minimum renewable energy penetration shortage or additional renewable energy exceeding minimum renewable energy penetration in period t

Parameters

h_{gt} :	unit FGI holding cost for product g in period t
b_{gt} :	unit backlogging cost for product g in period t
ω_{gt} :	unit WIP cost for product g in period t
D_{gt} :	demand for product g during period t
C_k :	capacity of work center k in units of time
α_{gl} :	processing time of operation l of product g
$L(g,l)$:	estimated time elapsing from the release of the raw material of product g to the completion of the operation l of product g
nwt:	number of installed WTs
npv :	number of installed solar PVs
e_g :	energy consumption for a single lot of product g per period
$APWT_t$:	average power provided by a single WT in period t
$APPV_t$:	average power provided by a single PV in period t
ce_t :	cost per unit of power taken from the substation in period t
cr_t :	revenue per unit of surplus power returned to the main grid in period t
LF_t :	fixed load, independent from producing chips for providing the clean room environment and for running machines in stand-by mode
λ:	minimum percentage of renewable energy penetration
δ:	unit penalty cost for not reaching the target percentage of renewable energy penetration.

Using the abbreviations $x^+ := \max(x,0)$ and $x^- := \min(x,0)$ the model can be formulated as follows:

$$\min \sum_{g\in G}\sum_{t=1}^{T}\left[\omega_{gt}W_{gt} + h_{gt}I_{gt} + b_{gt}B_{gt}\right] + \sum_{t=1}^{T}\left[ce\,APS_t^+ + cr\,APS_t^-\right] + \delta\sum_{t=1}^{T}RE_t^+ \qquad (1)$$

subject to

$$W_{g,t-1} + X_{gt} - Y_{gt} = W_{gt}, \qquad \text{for all } g \in G, t = 1,\ldots,T \qquad (2)$$

$$Y_{gt} + I_{g,t-1} - I_{gt} - B_{g,t-1} + B_{gt} = D_{gt}, \qquad \text{for all } g \in G, t = 1,\ldots,T \tag{3}$$

$$Y_{gtl} = X_{g,t-\lfloor L(g,l) \rfloor}, \qquad \text{for all } g \in G, t = 1,\ldots,T, l \in O(g) \tag{4}$$

$$\sum_{g \in G} \sum_{l \in O(k)} \alpha_{gl} Y_{gtl} \le C_k, \qquad \text{for all } k \in K, t = 1,\ldots,T \tag{5}$$

$$LF_t + \sum_{g \in G} e_g W_{gt} - APWT_t nwt - APPV_t npv = APS_t, \qquad \text{for all } t = 1,\ldots,T \tag{6}$$

$$\lambda\left(LF_t + \sum_{g \in G} e_g W_{gt} \right) - APWT_t nwt - APPV_t npv = RE_t, \qquad \text{for all } t = 1,\ldots,T \tag{7}$$

$$X_{gt}, Y_{gtl}, Y_{gt}, W_{gt}, I_{gt}, B_{gt}, RE_t, APS_t \ge 0, \qquad \text{for all } g \in G, t = 1,\ldots,T, l \in O(g). \tag{8}$$

The objective function (1) to be minimized is the sum of WIP, FGI holding, and backlogging cost over all products and periods. Moreover, additional costs for using energy from the substation and for violating a minimum penetration of renewable energy are considered. At the same time, the cost can be reduced by offering renewable surplus energy to the main grid. WIP variables and WIP balance constraints (2) are included to compute the WIP cost in the objective function. Constraint set (3) represents FGI material balance at the end of the line. Constraints (4) define the relation between the time a lot of product *g* is released into the wafer fab and completing processing at operation *l* of product *g*. As soon as a lot is processed at a given operation, it becomes available to the next operation on its routing. Constraint set (5) ensures that the total time required to process all operations at each work center in a given period *t* does not exceed the time available at that work center. The model assumes that an operation consumes capacity in the period that it is processed. Constraints (6) compute the amount of energy that is taken from the substation or sent back to the main grid. The power shortage or surplus of the required minimum renewable energy penetration is described by constraints (7). Finally, constraints (8) ensure nonnegativity of the decision variables.

Model (1)-(8) incorporates lead time estimates. Let $L(g,l)$ be a lead time estimate for operation l of product g. We compute $L(g,l)$ by the recursion $L(g,l) := L(g,l-1) + FF_g \alpha_{gl}$, for all $g \in G, l \in O(g)$, where $L(g,0) := 0$. Here, FF_g denotes the flow factor of product *g*, defined as the ratio of the average time required for material started into the process to become available as FGI to the sum of the processing times of all its operations. FF_g values are obtained from long simulation runs for a given bottleneck utilization.

3.2 Description of the Execution Level and the Simulation-based Grid Search

The obtained release schedules are executed. To describe this approach, we have to introduce the following additional notations:

τ : micro period index
PWT_τ : realized power provided by a single WT in micro period τ
PPV_τ : realized power provided by a single PV in micro period τ
r_g : unit revenue for product *g*
oWT : operating and maintenance cost per unit of produced power for a single WT per micro period
oPV : operating and maintenance cost per unit of produced power for a single PV per micro period
iWT : fixed installation cost for a single WT unit for the entire planning horizon T
iPV : fixed installation cost for a single PV unit for the entire planning horizon T
PS_τ : amount of power provided by station or surplus energy sent back to the main grid in micro period τ.

The tilde symbol is used to indicate that realizations of the decision variables of the planning formulation are considered. The amount of power provided by the substation in a period is computed as follows:

$$LF_\tau + \sum_{g \in G} e_g \tilde{W}_{g\tau} - PWT_\tau nwt - PPV_\tau npv = PS_\tau, \qquad \text{for all } \tau = 1,\ldots,T. \tag{9}$$

The following modified objective function is used to evaluate the outcome of an executed plan:

$$f(nwt,npv) := \sum_{g \in G} \left\{ \sum_{t=1}^{T} \left[r_g \tilde{Y}_{gt} - \omega_{gt} \tilde{W}_{gt} - h_{gt} \tilde{I}_{gt} - b_{gt} \tilde{B}_{gt} \right] - \delta \sum_{\tau=1}^{T} \left(\lambda \left(\sum_{g \in G} e_g \tilde{W}_{g\tau} + LF_\tau \right) - PWT_\tau nwt - PPV_\tau npv \right)^{+} \right\}$$

$$- iWT\, nwt - iPV\, npv - \sum_{\tau=1}^{T} \left[oWT\, PWT_\tau\, nwt + oPV\, PPV_\tau npv \right] - \sum_{\tau=1}^{T} \left[ce\, PS_\tau^{+} + cr PS_\tau^{-} \right]. \tag{10}$$

The first part of the objective function (10) considers revenue. Manufacturing-related costs are then taken into account. The violation of the minimum percentage of renewable energy penetration λ is penalized by δ. Operating and maintenance costs are taken into account in addition to fixed installation costs. Finally, the cost for using energy from the substation or returning surplus energy to the main grid is also considered. A smaller period length is applied for computing the objective function value to ensure that the fine-grained load and weather information can be taken into account. In order to differentiate the smaller periods from the regular ones, we call them micro periods.

Simulation is used to determine the $\tilde{Y}_{gt}, \tilde{W}_{gt}, \tilde{I}_{gt}, \tilde{B}_{gt}, \tilde{W}_{g\tau}, PWT_\tau$, and PPV_τ values for a given number of WTs and PVs. We use a discrete-event simulation model of the wafer fab, a wind power volatility model, and a model of the power output for the solar PVs. The last two models form the weather submodel. The overall infrastructure is shown in Figure 1.

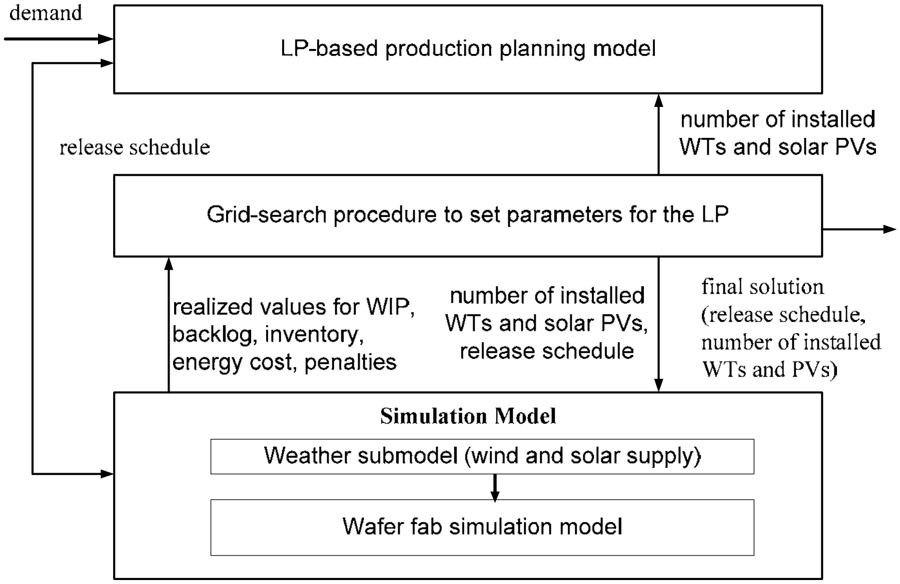

Figure 1: Simulation infrastructure.

The grid-search (GS) procedure, shown in Figure 1, works as described next. We consider the grid $G := \{(nwt,npv) \mid 0 \le nwt \le nwt_{max}, 0 \le npv \le npv_{max}\}$ for a given maximum number of WTs and PVs abbreviated by nwt_{max} and npv_{max}, respectively. The GS procedure is responsible for considering each of the $(nwt_{max}+1)(npv_{max}+1)$ grid points. A chosen grid point is transferred to the LP model and to the simulation model. The corresponding $f(nwt,npv)$ value is determined by executing the release schedule obtained from the LP model in the simulation model.

4 SIMULATION ENVIRONMENT

4.1 Wafer Fab Simulation Submodel

The execution level is represented by a discrete-event simulation of a wafer fab that is derived from the MIMAC I data set of Fowler and Robinson (1995). Lots consisting of 48 wafers are the moving entities in the wafer fab. The processing times at the work centers are deterministic and depend on the number of wafers of a lot or on the lot. Semiconductor characteristics such as unreliable, parallel machines, reentrant flows, sequence-dependent setup times, and batch processing are considered. Here, a batch is a group of lots that are processed at the same time on a single machine. Two products each of them with more than 200 operations are used. The model contains over 200 machines that are organized in around 70 work centers. The number of steppers is adjusted to ensure that they are a planned bottleneck for a product mix of 1:1. An instantaneous material transfer between successive operations on a given route is assumed. The batch processing machines have minimum and maximum batch sizes where only lots of the same product and at the same operation can be batched together. The machines are subject to exponentially distributed machine breakdowns. The First-In-First-Out (FIFO) dispatching rule is used.

4.2 Simulation Submodel for Wind and Solar Supply

The simulation submodel of the wafer fab's DG system is adapted from Villarreal *et al.* (2013). The model comprises detailed representations of the wind power generation, solar PV power generation, a grid-connected substation and a net-metering system. The sequence of events is shown in Figure 2.

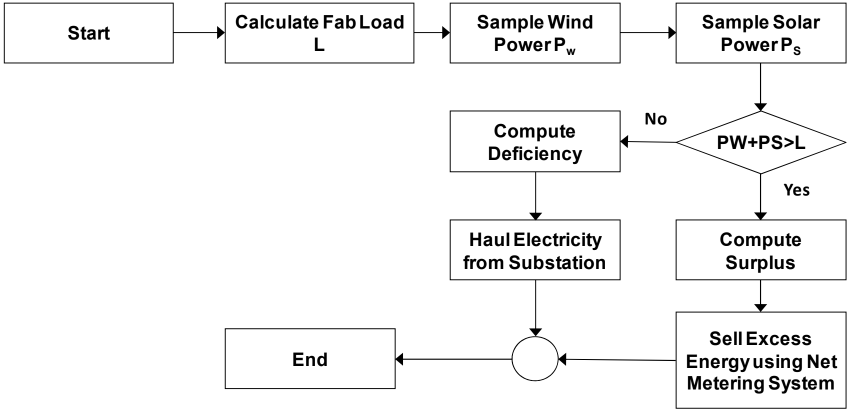

Figure 2: An overview of wafer fab's distributed generation system simulation sub-model.

The fab load L is determined based on the requirements of the wafer fab sub-model. The wind power (P_W) and the solar PV power (P_S) are estimated and aggregated. If the total power generated by the fab's renewable energy resources is greater than the load, then the excess energy is sent to the grid using the net metering system. On the other hand, if the total power is less than the load, then the shortage of energy is hauled from the substation.

The wind power is generated from the power function:

$$P_W(\upsilon_t) := \begin{cases} 0, & 0 \le \upsilon_t < \upsilon_c \\ 0.5\eta_W \rho A_W \upsilon_t^3, & \upsilon_c \le \upsilon_t < \upsilon_r \\ P_{max}, & \upsilon_r \le \upsilon_t < \upsilon_s \\ 0, & \upsilon_s \le \upsilon_t \end{cases}. \tag{11}$$

The wind speed υ_t, sampled from a Weibull distribution, is the argument. The parameters of this distribution depend on the geographical location of the facility. In (11), ρ denotes the air density, A_W is the area covered by the turbine blades, $\eta_W = 0.5926$ denotes a conversion rate while P_{max} is the maximum power capacity of the WT. The WT has the operating conditions standby (first condition in (12)), nonlinear (second condition), constant (third condition), and cut-off power (last condition). These operating conditions are described by the WT's cut-in speed (υ_c), the rated speed (υ_r), and the cut-off speed (υ_s).

The solar PV is generated according to the function $P_S(I_t, M) = M\eta_S A_S I_t(1-0.005(T_0-25))$. This function depends on a location-specific variable, the solar irradiance I_t. The irradiance is the solar radiation received by a solar panel under clear sky conditions, and it changes depending on the position of the sun throughout the day, the day of the year and the tilt angle of the solar panel. Weather patterns such as the partly cloudy day or a cloudy day reduce the value of I_t. In the simulation, the weather conditions predominating at the place where the fab site is located are simulated using a discrete random variable M. Moreover, η_S denotes a solar panel conversion rate which is set between 10-15%, A_s denotes the panel area, whereas T_o is the PV operating temperature.

5 COMPUTATIONAL EXPERIMENTS

5.1 Design of Experiments

We expect the performance of the production planning approach to depend on the bottleneck utilization. Normally distributed demand is determined for each product such that bottleneck utilization levels of 70% and 90% are reached. A coefficient of variation (CV) of 0.1 is used. Three independent demand instances are generated for each demand scenario. A planning horizon of 52 weeks is considered. The period length is one week. The approach of Leachman (2001) is applied to avoid horizon effects, i.e., three additional periods are used since the average cycle time is three weeks. The last three regular periods serve as frozen interval and constrain the releases of each product g to be equal in these periods. The demand of the additional periods is set to be equal to the average demand over the frozen interval. Long simulation runs for given bottleneck utilization levels are performed to determine initial WIP distributions in order to avoid initialization effects.

A maximum single WT and PV capacity of one MW is assumed. The average wind speed and its standard deviation are based on information from the National Climate Data Center (2017) and set to 3.5 m/s and 0.3 m/s, respectively. Cut-in, rated, and cut-off speed are 2.5 m/s, 10.0 m/s, and 25 m/s, respectively. The number of sunny days is 135 while the PV efficiency and skin temperature are 22.5% and 45 °C, respectively. The fixed load for providing the clean room environment, powering the recirculation air fans, and supplying ultrapure water and pure gases is 60% of the total wafer fab load (cf. Villareal et al. 2013). A scaled-down total wafer load of 3.8 MW (cf., for example, Hu and Chuah 2003, Quisenberry and Fenstermarker 2008) is considered. The energy consumption for a single lot is derived from Patton and Wiese (1999) and the average number of completed lots per year in the simulation model ensuring a 40% power usage by manufacturing activities. The fixed installation cost for a single WT and PV is \$825,000 and \$1,000,000 while the annual operating and maintenance cost is \$60,000 and 0.5% of the installation cost, respectively. A payoff period of 12 years is assumed. For the sake of simplicity, the annual interest rate is omitted.

Three levels of minimum percentage of renewable energy penetration are investigated. The unit penalty cost for not reaching the target percentage δ is \$0.01 per Wh. A revenue cr_t per unit of surplus power returned to the main grid of \$0.10 per kWh is used (cf. Flores-Espino 2015) while the cost per unit of power taken from the substation ce_t is \$0.15 per kWh (cf. National Public Radio 2017). Ten independent simulation replications are performed for each demand instance to obtain statistically valid results. The design of experiments is summarized in Table 1. The GS requires $121 \cdot 10$ simulation runs per factor combination that results in a total amount of 21,780 simulation runs. Manufacturing-related unit costs of \$3,500, \$1,500, and \$5,000 per week for WIP, FGI holding, and backlogging, respectively, are considered. The revenue per lot is $40,000. The set of lots from the planning level to be released in a given period is distributed uniformly over that period. A length of one hour is specified for the micro periods.

Table 1: Design of experiments.

Factor	Level	Count
Mean utilization	low, high	2
Minimum percentage of renewable energy penetration λ	0.20, 0.50, 0.70	3
Number of installed WTs (grid point coordinate)	$nwt \in \{0,\ldots,10\}$	11
Number of installed solar PVs (grid point coordinate)	$npv \in \{0,,\ldots,10\}$	11
Independent demand scenarios		3
Independent simulation replications		10

The infrastructure shown in Figure 1 is implemented in the C++ programming language while the simulation engine AutoSched AP is used. The average computing time for a single instance of the model (1)-(8) is up to one minute on a computer with 3.40 GHz Intel Core™ i7-2600 CPU and 16GB RAM.

5.2 Computational Results

The expected percentage of renewable energy penetration, the reduction of the expected profit (10), and the expected additional energy costs according to the second line of objective function (10) are used as performance measures. Average values over all demand scenarios are determined. The results of the simulation experiments are summarized in Table 2.

Table 2: Results of the simulation experiments.

λ	Mean utilization	nwt	npv	Renewable energy (%)	Profit reduction (%)		Additional energy cost (%)
					local	global	
0.20	low	0.33	0.67	21.93	1.90	1.90	7.86
	high	0.00	1.00	20.57	1.92	3.51	7.63
0.50	low	2.00	0.00	39.77	3.71	8.81	7.71
	high	2.00	1.00	54.47	-0.04	9.34	16.03
0.70	low	0.67	2.33	49.00	4.49	12.95	19.82
	high	2.00	2.00	62.13	-2.69	15.44	19.58

The first and second column of Table 2 describes the demand scenario that is characterized by the desired minimum percentage of renewable energy penetration and the mean bottleneck utilization. For each of the grid points $(nwt, npv) \in G \setminus (0,0)$ the configuration with the highest realized profit is determined. The averages over all demand scenarios are shown in the third and fourth column of Table 2. The fifth column summarizes the realized average renewable energy penetration of the best configurations. Both, the results of the configuration $(nwt, npv) = (0,0)$ with $\lambda = 0$ and the results of the configuration $(0,0)$ for the predefined minimum renewable energy penetration of Table 1 can serve as reference solution to measure the realized profit improvement. The first setting is called global reference solution while we refer to the latter one as local reference solution. The sixth and seventh columns represent the realized average profit improvement over the local and global reference solution, respectively. Finally, the expected additional energy costs over the local reference solution are summarized in the last column.

The total number of installed renewable energy resources and the realized percentage of renewable energy penetration increase with the desired minimum percentage of renewable energy penetration. However, the realized percentage of renewable energy penetration of the best configuration with respect to profit improvement matches the desired minimum percentage of renewable energy penetration only in the case of low λ values. The violation of renewable energy penetration is penalized at the planning level.

The penalties lead to an adaption of the release schedules increasing with the mean bottleneck utilization and the desired minimum percentage of renewable energy penetration and decreases with the number of installed renewable energy resources. The local reference solution is characterized by strong modifications of the release schedules to avoid penalties when the λ values and mean bottleneck utilization levels are large. The poor performance of the local reference solution results in profit improvements by an appropriate number of installed WTs and PVs. However, the comparison of the realized profit over the global reference solution with $\lambda = 0$ shows the profit reduction when taking sustainability issues into account.

A high wafer fab load, i.e. a large number of WIP lots, results in significant energy and manufacturing costs. The energy cost increases with the number of installed WTs and PVs because of fixed installation and operating and maintenance costs. However, profit improvements can be observed in case of high load situations for medium and high λ indicating a poor local reference solution. The impact of the energy cost increases with the load of the wafer fab and the percentage of renewable energy penetration. On the one hand, additional energy cost and profit reductions over the global reference solution can be observed for using renewable energy resources. On the other hand, an increased renewable energy penetration mitigates the carbon footprint.

6 CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

In this paper, we discussed the incorporation of elements of a sustainable and DG system into a mid-term production planning formulation for a single wafer fab. The production planning formulation is based on fixed, exogenous lead times. We applied a GS approach to determine an appropriate number of WTs and PVs. We obtained release schedules and the decisions made for sizing the renewable energy sources were assessed using a wafer fab simulation model and a simulation model for the weather conditions. The experiments demonstrated that it is reasonable to combine production-related decisions and decisions with respect to the design of a DG system.

There are several directions for future research. First of all, more simulation experiments with different demand and weather scenarios are necessary to assess the taken approach. The time-consuming GS can be avoided by using a metaheuristic-based search method similar to the approach proposed by Ziarnetzky and Mönch (2016). Moreover, applying the number of WTs and PVs obtained from the GS in production planning models that are assessed in a rolling horizon environment using the simulation infrastructure of Posnignon and Mönch (2014) seems to be interesting too. In this setting it is possible to use updated demand information. While we proposed a mid-term planning approach for a single wafer fab in the present paper, we believe that an embedding of the design of the DG system should take place on a more strategic level. Therefore, the deterministic network design model proposed by Stray et al. (2006) and its stochastic counterpart by Rastogi et al. (2011) might serve as a starting point for future research. The proposed model might use discrete-event simulation on the supply chain level for assessing the performance of the proposed planning decisions. Here, we believe that is worth to enrich the supply chain simulation testbed proposed by Ewen et al. (2017) by sustainability aspects.

ACKNOWLEDGMENTS

The authors would like to thank Hans Ehm and Thomas Ponsignon from Infineon Technologies AG for helpful discussions on sustainability issues in semiconductor supply chains. The research is funded in parts by a research grant from the University of Hagen within the MaXFab project. The authors gratefully acknowledge this financial support.

REFERENCES

Benjaafar, S., Y. Li, and M. Daskin. 2013. “Carbon Footprint and the Management of Supply Chains: Insights From Simple Models.” *IEEE Transactions on Automation Science and Engineering* 10(1):99–116.

Biel, K., and C. H. Glock. 2016. “Systematic Literature Review of Decision Support Models for Energy-efficient Production Planning.” *Computers & Industrial Engineering* 110:243–259.

Chien, C., S. Dauzère-Pérès, H. Ehm, J. Fowler, Z. Jiang, S. Krishnaswamy, L. Mönch, and R. Uzsoy. 2011. “Modeling and Analysis of Semiconductor Manufacturing in a Shrinking World: Challenges and Successes.” *European Journal of Industrial Engineering* 5(3):254-271.

Duflou,, J. R., J. W. Sutherland, D. Dornfeld, C. Herrmann, J. Jeswiet, S. Kara, M. Hauschild, and K. Kellens. 2012. “Towards Energy and Resource Efficient Manufacturing: A Processes and Systems Approach.” *CIRP Annals - Manufacturing Technology* 61:587–609.

Ewen, H., L. Mönch, H. Ehm, T. Ponsignon, J. Fowler, and L. Forstner. 2017. “A Testbed for Simulating Semiconductor Supply Chains.” *IEEE Transactions on Semiconductor Manufacturing* 30(3):293-305.

Flores-Espino, F. 2015. “Compensation for Distributed Solar: A Survey of Options to Preserve Stakeholder Value.” *National Renewable Energy Laboratory, Tech. Rep. NREL/TP-6A20-62371*.

Fowler, J. W., and J. Robinson. 1995. “Measurement and Improvement of Manufacturing Capacity (MIMAC) Final Report.” *Technology Transfer #95062861A-TR, SEMATECH*.

Gahm, C., F. Denz, M. Dirr, and A. Tuma. 2016. “Energy-efficient Scheduling in Manufacturing Companies: A Review and Research Framework.” *European Journal of Operational Research* 248:744–757.

Giret, A., D. Trentesaux, and V. Prabhu. 2015. “Sustainability in Manufacturing Operations Scheduling: A State of the Art Review.” *Journal of Manufacturing Systems* 37(1):126–140.

Hu, S., and Y. Chuah. 2003. “Power Consumption of Semiconductor Fabs in Taiwan.” *Energy* 28(8):895–907.

Kacar, N. B., L. Mönch, and R. Uzsoy. 2013. “Planning Wafer Starts using Nonlinear Clearing Functions: a Large-Scale Experiment.“ *IEEE Transactions on Semiconductor Manufacturing* 26(4):602-612.

Kacar, N. B., L. Mönch, and R. Uzsoy. 2016. “Modeling Cycle Times in Production Planning Models for Wafer Fabrication.” *IEEE Transactions on Semiconductor Manufacturing* 29(2):153-167.

Leachman, R. 2001. “Semiconductor Production Planning.” *Handbook of Applied Optimization*, P. M. Pardalos and M. G. C. Resende, Eds. New York, NY, USA: Oxford Univ. Press:746-762.

Masmoudi, O., A. Yalaoui, Y. Ouazene, and H. Chehade. 2015. “Lot-sizing in Flow-Shop with Energy Consideration for Sustainable Manufacturing Systems.” *IFAC-PapersOnLine* 48(3):727–732.

Mönch, L., J. W. Fowler, and S. J. Mason. 2013. *Production Planning and Control for Semiconductor Wafer Fabrication Facilities: Modeling, Analysis, and Systems*. New York: Springer.

Mönch, L., R. Uzsoy, and J. W. Fowler. 2017. “A Survey of Semiconductor Supply Chain Models Part I: Semiconductor Supply Chains, Strategic Network Design, and Supply Chain Simulation.” *International Journal of Production Research*. Submitted for publication.

National Climate Data Center. 2017. [Online]. Available: http://www.ncdc.noaa.cov/cdo-web/search, Accessed May 9, 2017.

National Public Radio. 2017. [Online]. Available: http://www.npr.org/sections/money/2011/10/27/-141766341/the-price-of-electricity-in-your-state, Accessed May 9, 2017.

Patton, R., and S. Wiese. 1999. “Worldwide Fab energy Survey Report.” *Technology Transfer #99023669B-ENG, SEMATECH*.

Ponsignon, T., and L. Mönch. 2014. “Simulation-based Performance Assessment of Master Planning Approaches in Semiconductor Manufacturing.“ *Omega* 46:21-35.

Quisenberry, J., and K. Fenstermarker. 2008. “Summary Facilities Energy Consumption in 200 mm and 300 mm Fabs.” *Technology Transfer #08024920A-TR, SEMATECH*.

Rastogi, A., J. Fowler, W. Carlyle, O. Araz, A. Maltz, and B. Büke. 2011. "Supply Network Capacity Planning for Semiconductor Manufacturing with Uncertain Demand and Correlation in Demand Considerations." *International Journal of Production Economics* 134(2):322-332.

Santana-Viera, V., J. Jimenez, T. Jin, and J. Espiritu. 2015. "Implementing Factory Demand Response via Onsite Renewable Energy: A Design-of-Experiment Approach." *International Journal of Production Research* 53(23):7034-7048.

Stray, J., J. Fowler, W. Carlyle, and A. Rastogi. 2006. "Enterprise-Wide Strategic and Logistics Planning for Semiconductor Manufacturing." *IEEE Transactions on Semiconductor Manufacturing* 19(2):259-268.

Villarreal, S., J. Jimenez, T. Jin, and M. Cabrera-Ríos. 2013. "Designing a Sustainable and Distributed Generation System for Semiconductor Wafer Fabs." *IEEE Transactions on Automation Science and Engineering* 10(1):16–26.

Wang, F., X. Lai, and N. Shi. 2011. "A Multi-Objective Optimization for Green Supply Chain Network Design." *Decision Support Systems* 51(2):262-269.

Ziarnetzky, T., and L. Mönch. 2016. "Simulation-Based Optimization for Integrated Production Planning and Capacity Expansion Decisions." In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, 2992-3003. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

AUTHOR BIOGRAPHIES

TIMM ZIARNETZKY is a Ph.D. student at the Chair of Enterprise-wide Software Systems, University of Hagen. He received M.S. degree in Mathematics from the Technical University Dortmund, Germany. His research interests include production planning, production control, and simulation-based optimization. He can be reached by email at Timm.Ziarnetzky@fernuni-hagen.de.

THULASI KANNAIAN is a M.S. student in Industrial Engineering at the Ingram School of Engineering, Texas State University. He received his B.S. degree in Aeronautical Engineering from Hindustan University, India. His research interests include modelling and analysis of manufacturing systems, discrete-event simulation, design of experiments, supply chain, and quality. His email address is tkk12@txstate.edu.

JESUS JIMENEZ is an Associate Professor in the Ingram School of Engineering at Texas State University. He received his B.S. and M.S. in Industrial Engineering from the University of Texas at El Paso, and his Ph.D. in Industrial Engineering from Arizona State University. His research interests are in simulation modeling and analysis of manufacturing systems, data-intensive analysis and simulation, and sustainable lean manufacturing. His email address is jesus.jimenez@txstate.edu.

LARS MÖNCH is Full Professor of Computer Science at the Department of Mathematics and Computer Science, University of Hagen where he heads the Chair of Enterprise-wide Software Systems. He holds M.S. and Ph.D. degrees in Mathematics from the University of Göttingen, Germany. After his Ph.D., he obtained a habilitation degree in Information Systems from Technical University of Ilmenau, Germany. His research and teaching interests are in information systems for production and logistics, simulation, scheduling, and production planning. He can be reached by email at Lars.Moench@fernuni-hagen.de.