



PAGE 1 — OFFICIAL PUBLICATION IDENTITY

Experimental Design and Stochastic Capacity Planning for Reliable Production Decisions

Haluk Ali Güneş

OFFICIAL ARTICLE RECORD

Journal	Current Science Georgia
ISSN	2667-9515
Barcode	977266795001
Publication Code	0099-0001-2026-0901
Volume / Issue	Volume 2609, No. 26103001
Year	2026
Page Number	26-51
DOI	DOI-ready metadata field reserved
Status	Approved Publication Record

AUTHOR PORTRAIT /
TOPIC VISUAL

Topic visual placeholder reserved.

Create a professional academic article visual for the topic "Experimental Design and Stochastic Capacity Planning for Reliable Production Decisions", suitable for Current Science Georgia publication visibility, no

AUTHOR INFORMATION

Author	Haluk Ali Güneş
Affiliation	Not specified
Email	Not specified
ORCID	Not specified

VERIFICATION QR



<https://csi.ge/certificate-verification/?verify=0099-0001-2026-0901>





PAGE 2 — OFFICIAL CITATION & JOURNAL INFORMATION

To cite this article: Haluk Ali Güneş, *Current Science*, Volume 2609, No. 26103001, 2026, p. 26-51. - 0099-0001-2026-0901.

JOURNAL INFORMATION

Journal	Current Science Georgia
ISSN	2667-9515
Barcode	977266795001
Publication Code	0099-0001-2026-0901
Page Number	26-51
Verification	https://csi.ge/certificate-verification/?verify=0099-0001-2026-0901

OFFICIAL JOURNAL STATEMENT

Our studies are in a format accredited, approved, and supported by EAALS - European Academic Studies and Laboratory Services. Current Science Georgia offers scholarly works, services, and publications to the international academic community through editorial control, publication preparation, record verification, and scientific publication workflow support.

Current Science Multidisciplinary Academic Journal with Review Panel is a monthly multidisciplinary academic journal with a multi-science peer-review and editorial evaluation workflow.

The journal continues to maintain structured article records, publication archives, verification records, metadata preparation, and journal visibility support for accepted scholarly publications.

Response or rejection time for applications may

EDITORS GROUP

Concessionaire: Tsisana
Kharabadze
Niyaz Bokvadze
Prof. Sabrina Corbi
Prof. Samantha Lewes

APPROVAL SEAL





PRODUCTION STATUS

This article record is marked as an approved publication record by Current Science Georgia.

CSG • VERIFIED • CSG • SECURE • CSG • VERIFIED

CSG • VERIFIED • CSG • SECURE • CSG • VERIFIED



Experimental Design and Stochastic Capacity Planning for Reliable Production Decisions

Haluk Ali Güneş¹

1

ABSTRACT

Production plans can satisfy every constraint in a deterministic model and still be difficult to execute when capacity fluctuates over time. This article develops an integrated approach linking experimental design, stochastic processes, and linear production planning. A structured synthesis of three industrial engineering course manuscripts is combined with a reproducible analytical experiment based on a two product manufacturing problem. Two interventions are examined in a full factorial arrangement: a reduction in assembly time and a reduction in the persistence of daily capacity disturbances. Stationary autoregressive processes represent capacity uncertainty, while allocated chance constraints translate weekly reliability requirements into linear planning limits. Exact Gaussian probabilities are checked against 200,000 simulated weeks. The nominal plan provides a scheduled contribution of 166,000 monetary units but achieves only 49.47 percent joint capacity feasibility. A plan calibrated under an incorrect independence assumption reaches 88.71 percent, below its intended 95 percent guarantee. Among the four intervention combinations evaluated under the same conservative risk allocation, the combined intervention provides the highest net scheduled contribution, 164,066 monetary units, with 96.05 percent model based feasibility. Its advantage over process improvement alone is modest and sensitive to intervention cost, capacity drift, and the assumed persistence reduction. The analysis shows how an operational bottleneck can change the economic value of an experimentally studied factor, creating interaction at the decision level even when physical effects are separately specified. Reliable implementation therefore requires evidence about intervention effects, temporal dependence, and the consequences of infeasibility, together with evaluation on operating periods withheld from model development.

Keywords: experimental design; stochastic processes; production planning; chance constraints; capacity uncertainty

Keywords: *Experimental; Design; Stochastic; Capacity; Planning.*

A production manager who commits the full average capacity of an assembly department has made a stronger promise than the average itself supports. Some weeks will provide fewer usable hours, and a sequence of modest disruptions can be as consequential as a single major failure. Linear programming describes which product mix best uses a specified quantity of resources. Experimental design establishes how controllable changes affect a process. Stochastic modelling describes the distribution and temporal organisation of the resources that remain available. The practical question is how these forms of evidence should be connected when a production commitment must be both economically attractive and reasonably likely to be executable. The distinction between an optimal plan and a reliable commitment gives this article its organising problem (Charnes & Cooper, 1959; Dantzig, 1955).

The problem is developed from three industrial engineering manuscripts by Güneş: engineering experimental design, operations research, and stochastic processes. The first discusses randomisation, replication, blocking, factorial designs, and the assumptions underlying analysis of variance. The second develops linear programming, duality, and sensitivity analysis, including a two product manufacturing example. The third introduces time indexed random variables, Markov models, dependence, and simulation (Güneş, 2026a, 2026b, 2026c). Their common concern is the construction of defensible decisions from simplified representations of an operating system. Read together, they also expose an unresolved connection. Estimated process improvements and temporal uncertainty need to enter the same optimisation model, while the resulting decision must be checked against the conditions under which it will be used.

This connection matters because statistical success and operational value answer different

and optimisation research has formalised the importance of evaluating information through the decision it supports, rather than through prediction accuracy alone (Bertsimas & Kallus, 2020; Elmachoub & Grigas, 2022). Experimental design contributes an additional requirement: the estimated effect of changing a setting should be supported by a design that separates that change from background differences. These requirements become particularly important when the organisation is choosing an intervention, not merely forecasting what happens under its existing policy.

The research question is therefore how experimentally interpretable process changes and temporally dependent capacity uncertainty can be incorporated into a transparent production planning model. Three subsidiary questions guide the analysis. How much does ignoring temporal dependence change the reliability of a plan? How do intervention effects change when the active resource constraint changes? What evidence would be needed before the resulting recommendation could guide an actual production commitment? The article addresses these questions through an integrative methodological analysis and a fully specified numerical example. It does not estimate intervention effects from factory records. Its contribution is an auditable connection between the three source topics, with exact calculations that make the consequences of particular assumptions visible.

The analysis retains the small production problem because its simplicity makes the reasoning inspectable. Four intervention combinations, two resource constraints, and a five day horizon are enough to distinguish nominal optimality, probabilistic feasibility, and economic value. The findings show that accounting for dependence can matter more than obtaining an additional decimal place in the nominal solution. They also show that

changes, and a feasible route to field validation.

Methodology

Research design and source synthesis

The study combines a structured conceptual synthesis with an analytical factorial experiment. The source corpus comprises three unpublished doctoral course manuscripts: Engineering Experimental Design, chapters 1 to 10, dated May 2026; Operations Research, chapters 1 to 6, dated March 2026; and Stochastic Processes, dated March 2026. The manuscripts contain 13, 19, and 28 PDF pages respectively, including their title pages. They are treated as the intellectual starting point and as the source of the manufacturing example, rather than as independent empirical studies or peer reviewed evidence. Their textbook discussions motivate the integration, while published methodological sources support the statistical and optimisation claims (Güneş, 2026a, 2026b, 2026c).

Each manuscript was examined against five questions: what can be controlled, what must be estimated, how uncertainty is represented, what counts as a satisfactory decision, and how the model could be checked. This comparison produced the analytical connections summarised in Table 1. The purpose was to identify compatible modelling roles, not to count the prevalence of themes in a population of documents. No interview coding, interrater reliability assessment, systematic review flow diagram, or meta-analysis is claimed. The literature search was purposive, using terms associated with factorial experiments, stochastic programming, chance constraints, autocorrelation, simulation verification, and decision focused learning. Bibliographic details were checked against publisher, journal, institutional, or author repository records. The review consequently supports a methodological

Relevant material

Role in this study

Experimental design

Randomisation, blocking, replication, factorial comparisons

Defines intervention comparisons and the requirements for subsequent field evidence

Operations research

Product mix, resource constraints, duality, sensitivity

Supplies the production example and the interpretation of constrained economic value

Stochastic processes

Time indexed variation, dependence, simulation

Supplies the framework for capacity trajectories and numerical probability checks

Sources: Güneş (2026a, 2026b, 2026c). The mapping is an interpretive synthesis of the course manuscripts.

Production model and decision timing

The numerical starting point is the washing machine and refrigerator example in the operations research manuscript, chapter 2, PDF pages 9 and 10. A washing machine requires one hour of software engineering and two hours of assembly, while a refrigerator requires one hour of each. The weekly resource limits are 90 engineering hours and 120 assembly hours. Refrigerator demand is capped at 50 units; the example imposes no separate demand cap on washing machines. Selling prices less the stated material and additional costs give contributions of 1,600 and 2,200 respectively. These arithmetic inputs are retained, but expressed in monetary

refrigerator demand cap. Continuous quantities are retained initially because the source model is a linear programme; an integer comparison is reported separately. No overtime, emergency outsourcing, cancellation penalty, or adaptive change in the product mix is included. The objective is therefore planned contribution from the scheduled quantities, not expected realised profit under disruption. This distinction is necessary because the model evaluates whether a commitment can be met, but does not price every response when it cannot.

$$\max 1600x_1 + 2200x_2 \quad x_1 + x_2 \leq 90, \quad 2x_1 + x_2 \leq 120 \quad x_1 \geq 0, \quad 0 \leq x_2 \leq 50(1)$$

Factorial intervention structure

Two binary factors define a complete two by two numerical experiment. Factor A represents an assembly process improvement that reduces both assembly coefficients by 10 percent, from 2 and 1 to 1.8 and 0.9 hours. Its weekly equivalent cost is 4,000 MU. Factor B represents a recovery or coordination intervention that reduces the lag one correlation of daily capacity disturbances from 0.6 to 0.2. Its weekly equivalent cost is 1,000 MU. B leaves marginal daily variances and mean capacity unchanged. This is a deliberately narrow representation of an intervention that shortens the persistence of a disturbance. It does not assert that a particular maintenance or management practice achieves that effect in a factory.

The four combinations are D00, D10, D01, and D11, where the first digit identifies A and the second identifies B. The costs are additive and the physical effects are specified separately. Consequently, any nonadditivity in the optimised economic response arises through the planning problem, including changes in its binding constraint. Table 2 distinguishes source inputs from the assumptions introduced for the numerical experiment. Since the experiment uses stipulated parameters and exact optimisation, its

Table 2 Source inputs and assumptions introduced for the numerical experiment

Quantity

Specification

Provenance

Unit contribution

1,600 MU for x_1 ; 2,200 MU for x_2

Source example

Mean capacities

Engineering 90 h; assembly 120 h per week

Source example

Resource requirements

Engineering 1 h each; assembly 2 h and 1 h

Source example

Refrigerator demand

$x_2 \leq 50$ units per week

Source example

Daily marginal standard deviations

Engineering 0.6 h; assembly 1.5 h

Illustrative assumption

Factor A

Assembly coefficients multiplied by 0.9; cost 4,000 MU

Illustrative assumption

Factor B

Daily correlation reduced from 0.6 to 0.2; cost 1,000 MU

All intervention costs are weekly equivalent amounts. No intervention effect was estimated from factory observations.

Stochastic capacity over a five day horizon

For resource j and day t , let e_{jt} be a zero mean capacity deviation measured in hours. Within a simulated week, these deviations follow a stationary first order Gaussian autoregression. The marginal daily standard deviations are 0.6 hours for engineering and 1.5 hours for assembly. The autoregressive coefficient is denoted by ϕ . Independent standard normal innovations drive the two resources, and the initial daily deviation is drawn from its stationary distribution. Resource independence is an assumption of the analytical probability calculation; temporal independence is deliberately not assumed. A simulation replication is one independent five day week, rather than one of five days treated as an independent observation.

$$e_{j,t+1} = \phi e_{j,t} + \sigma_j z_{j,t+1}, z_{j,t+1} \sim N(0, 1), z_{j,t} \sim N(0, 1) \quad (2)$$

Weekly capacity C_j is the mean weekly amount b_j plus the sum of its five daily deviations. For this process, the variance of that sum includes every within-week covariance. Equation 3 follows by expanding the variance of a sum and using $\text{Cov}(e_{jt}, e_{j,t+k}) = \sigma_j^2 \phi^k$. It provides an exact bridge between the process model and the planning horizon. The calculation uses the marginal daily standard deviation σ_j ; the innovation standard deviation is smaller by the factor given in Equation 2. Confusing these two quantities would change the comparison. The stationary autoregressive representation is standard, while the particular parameters here are illustrative (Brockwell & Davis, 2016).

$$C_j = b_j + \sum_{t=1}^5 e_{j,t}, \text{Var}(C_j) = \sigma_j^2 [5 + 2 \sum_{k=1}^4 (5-k)\phi^k] \quad (3)$$

The Gaussian model permits negative capacities

evaluated under an unconditional stationary weekly distribution. If the preceding day's state were observed and used at commitment time, the relevant distribution would instead be conditional on that state. The present calculation does not claim to implement such a forecasting policy.

Reliability constraints and optimisation

The planning requirement is that both resource loads can be met in the same week with probability at least 0.95. Rather than solve the exact joint chance constrained problem, the study assigns a violation allowance of 0.025 to each resource. The union bound then ensures that the probability of at least one violation is no greater than 0.05 when the marginal models are correct. This sufficient condition does not require the resources to be independent. Gaussian marginal capacity makes each allocated constraint equivalent to a linear bound using the 97.5th percentile of the standard normal distribution, $z = 1.959964$ (Charnes & Cooper, 1959; Shapiro et al., 2021).

$$\Pr(L_1 \leq C_1, L_2 \leq C_2) \geq 0.95 \Leftrightarrow L_j \leq b_j - z_{0.975} \sigma_j, j = 1, 2, L_1 = x_1 + x_2, L_2 = (1 - 0.1A)(2x_1 + x_2) \quad (4)$$

The assembly multiplier is $1 - 0.1A$, and implementation cost is $4,000A + 1,000B$. Each design is solved by linear programming with the relevant conservative capacity limits. The reported preferred design is the highest value among these four solutions under the same fixed risk allocation. It is not a claim of optimality over all possible interventions or over every feasible joint chance constrained plan. Under the additional assumption of independence between the two resource processes, the exact joint feasibility probability is the product of the two Gaussian upper-capacity probabilities in Equation 5. This separate evaluation shows how conservative the sufficient constraints are for each selected plan.

$$p(x, A, B) = \Pr(L_j \leq C_j) = \prod_{j=1}^2 \Phi\left(\frac{b_j - L_j \sigma_j}{\sigma_j}\right) \quad (5)$$

reliability rule but calculates weekly variance as though the daily deviations were independent. It is then evaluated under the actual stipulated process with $\phi = 0.6$. The intervention comparisons use the correctly specified dependence for each design. All optimisations were performed with the HiGHS method exposed through SciPy 1.17.0, using Python 3.12.14 and NumPy 2.3.5. The two dimensional problem also permits inspection of active constraints and direct integer enumeration, providing checks on the economic interpretation of the numerical solution.

The analytical probabilities were checked using 200,000 independently generated weeks. NumPy's PCG64 generator used seed 20260930. A standard normal array with dimensions 200,000 by 2 by 5 supplied stationary initial states and subsequent innovations. The same underlying normal draws were reused across policies, preserving a matched comparison without changing any policy's marginal distribution. Independent weeks remained the units used for uncertainty calculations. Common random numbers can improve comparative precision, but their effectiveness depends on the model and comparison; no universal variance reduction guarantee is assumed here (Glasserman & Yao, 1992). The simulation design follows the principle of reporting the generating process, estimand, methods, and performance measure explicitly (Morris et al., 2019).

For each fixed plan, a week passes if both simulated capacities exceed their corresponding loads. The resulting success fraction is accompanied by a 95 percent Wilson interval (Wilson, 1927). These intervals quantify Monte Carlo sampling error, not uncertainty about whether the assumed process describes a real plant. Exact model probabilities remain the principal findings. Additional checks vary the reliability target, enumerate integer product quantities, and hold the selected plan fixed while

Literature Review

Experimental evidence and the meaning of a process effect

Experimental design concerns the evidence generated by a comparison. Random allocation of a treatment supports a different interpretation from observing that a setting happens to coincide with better performance. Blocking addresses known sources of background variation, while replication supplies information about variation across experimental units. A factorial arrangement allows the effect of one factor to be examined at the levels of another. These distinctions are central to the experimental design manuscript and are also reflected in the NIST treatment of randomised blocks (Güneş, 2026a; NIST/SEMATECH, n.d.-c). For production planning, they determine whether a change in a time coefficient or capacity distribution can reasonably be attributed to the proposed intervention.

Box and Wilson's (1951) response surface work connects experiments to the search for favourable operating conditions. Its relevance is the sequential relationship between learning about a response and deciding where to experiment next. A production planning application adds a complication: the economically relevant response may be the value of a constrained system, not the output of a single machine. A locally promising setting can affect the product mix, consume another resource, or alter which constraint binds. The present study extends that observation through a small transparent calculation rather than claiming a new response surface method. It separates the physical response to an intervention from the value that the planning model assigns to that response.

Practical randomisation is often restricted. A

(n.d.-b) explains why split plot structures require an analysis that recognises more than one experimental unit size. This is directly relevant when an assembly improvement is tested alongside an intervention intended to change the persistence of disruptions. Repeated readings from the same configured machine are not independent replications of the configuration. Treating them that way exaggerates precision and can make an uncertain investment appear decisively supported.

Optimisation under uncertainty and the reliability requirement

Linear programming gives a precise answer to a precisely stated resource allocation problem. The answer is conditional on the coefficients, capacities, and admissible actions supplied to it. Dantzig (1955) explicitly distinguished decisions made before uncertainty is resolved from actions available in later stages. This distinction remains essential in contemporary stochastic programming (Shapiro et al., 2021). If overtime or outsourcing is available after a disruption, a model can trade the planned contribution against expected corrective costs. If the central requirement is that an advance commitment fit ordinary capacity with a specified probability, a chance constraint provides a different formulation. The two approaches represent different operational agreements, even when they concern the same plant.

Chance constrained programming makes the permissible probability of a violation part of the mathematical problem (Charnes & Cooper, 1959). A nominal resource limit cannot supply this information by itself. The probability pertains to a defined event, such as satisfying both weekly resource requirements, and depends on the underlying capacity distribution. Individual and joint constraints therefore need to be distinguished. Requiring each of two resources to be adequate with probability 0.95 does not

claim of exact joint optimality.

Robust optimisation supplies another response to uncertain coefficients. Bertsimas and Sim (2004) show how protection against data variation can be adjusted while preserving tractability in important linear settings. Their formulation differs from the Gaussian probability model used here: an uncertainty set and its protection parameters are not automatically empirical failure probabilities. Distributionally robust optimisation moves the uncertainty to the distribution itself. Mohajerin Esfahani and Kuhn (2018) construct tractable models using Wasserstein neighbourhoods of an empirical distribution. These approaches become relevant when the distribution or its parameters are poorly known. They do not remove the need to decide which deviations are credible, how protection is calibrated, or whether the observations represent the operating regime of interest.

Dependence and the planning horizon

Temporal dependence changes how variability accumulates. For independent observations, variances add without covariance terms. Under positive serial correlation, neighbouring deviations tend to reinforce one another, increasing the variance of their sum. A sequence of small capacity losses can therefore make a weekly plan less reliable than a daily standard deviation suggests. Time series analysis formalises the distinction between marginal behaviour, covariance structure, and the distribution conditional on observed history (Brockwell & Davis, 2016). NIST/SEMATECH (n.d.-a) also identifies autocorrelation as a reason that procedures relying on independent observations can give misleading uncertainty assessments. The issue here is the consequence for a capacity commitment, rather than the significance of an individual regression coefficient.

The stochastic processes manuscript provides the

processes, or a larger state model would be warranted by different operating mechanisms, not by a general preference for mathematical complexity. A process model should be chosen for the decisions and observations it must support. In particular, a model for gradual day to day deviations and a model for rare prolonged shutdowns make different claims about the probability of an adverse week.

Decision quality and computational evaluation

Bertsimas and Kallus (2020) connect predictive information to operational prescriptions, while Elmachtoub and Grigas (2022) develop prediction losses informed by the downstream optimisation problem. The broader review by Mandi et al. (2024) situates decision focused learning as an integration of estimation and constrained decision making. These contributions support the choice to evaluate process information through scheduled contribution and feasibility. They do not imply that every production problem needs a learned optimisation pipeline. For the small model examined here, explicit coefficients and analytical probabilities make the decisive assumptions easier to inspect. More elaborate estimation would need to demonstrate value against that transparent baseline.

Simulation introduces its own distinction between variability in the represented system and uncertainty about a fitted response. Ankenman et al. (2010) make both components explicit in stochastic kriging, which is useful when a simulator is too costly to evaluate throughout a design space. The current model needs no surrogate because its four cases can be solved directly. The distinction nevertheless matters. Increasing the number of simulated weeks reduces numerical sampling error, while it leaves a misspecified intervention effect or serial correlation unchanged. Morris et al. (2019) similarly emphasise precision and clarity in the design and reporting of simulation studies. A large

alternative policies are evaluated using noisy estimates. Smith and Winkler (2006) explain how selecting the most favourable estimate can produce systematic disappointment, even when the individual estimates have reasonable properties. Verification and validation address complementary parts of this problem: correct implementation of a model does not establish its adequacy for its intended real world use (Sargent, 2013). The present calculation avoids simulation based selection by solving the four stipulated cases analytically, then using simulation as a check. The remaining question is more consequential for deployment: whether the process changes and capacity behaviour assumed by the analysis can be sustained in operation. The arguments return to this issue through an explicit field validation proposal.

Aim

This section should be reviewed during editorial preparation. If the original manuscript does not provide this section separately, the editorial office should decide whether the existing article content already covers it or whether author-approved expansion is required.

Findings

Dependence changes the capacity allowance

The five day variance multiplier is 5 under independence, 13.0832 when $\phi = 0.6$, and 6.8752 when $\phi = 0.2$. At $\phi = 0.6$, the weekly standard deviation is therefore about 1.618 times the independence value, despite an unchanged marginal daily standard deviation. Engineering capacity has a weekly standard deviation of 2.170 hours, and assembly capacity has one of 5.426 hours. Under the recovery intervention, these fall to 1.573 and 3.933 hours. The change is entirely

that can be committed under the allocated chance constraints. Without an intervention, the planning limits are 85.746 engineering hours and 109.366 assembly hours. Under B, they become 86.917 and 112.291 hours. An analyst assuming independent days would instead use 87.370 and 113.426 hours. That last pair is more permissive than the limits justified even by the stipulated reduction in persistence. The ordering illustrates why a reliability label attached to an optimisation model is insufficient: the safety allowance depends on how uncertainty was aggregated to the decision horizon.

Economic performance and capacity feasibility

The nominal plan schedules 35 washing machines and 50 refrigerators, earning 166,000 MU of scheduled contribution. It exactly uses mean assembly capacity. Under the specified Gaussian process, assembly capacity exceeds its mean with probability one half; engineering capacity provides only a further restriction. Exact joint feasibility is consequently 49.47 percent. The independence comparator schedules 31.713 washing machines and 50 refrigerators, reducing contribution to 160,741 MU. It does introduce a safety allowance, but its actual model based reliability is only 88.71 percent when evaluated with $\phi = 0.6$. Accounting for uncertainty while ignoring its temporal structure therefore leaves a substantial gap below the intended 95 percent requirement.

Table 3 reports all four process-aware plans. Doo satisfies the reliability requirement with 29.683 washing machines and a net contribution of 157,493 MU. D11 provides the greatest net scheduled contribution among the evaluated designs, 164,066 MU, after its 5,000 MU weekly equivalent intervention cost. It schedules 36.917 washing machines and 50 refrigerators, with exact joint feasibility of 96.05 percent. The contribution gain over Doo is 6,574 MU, or approximately 4.17 percent. This is a comparison between two model-

analytical capacity feasibility

Policy

Washing machines

Refrigerators

Net contribution MU

Joint feasibility

Nominal

35.000

50

166,000

49.47%

Independence

31.713

50

160,741

88.71%

Doo

29.683

50

157,493

97.50%

D10

35.746

50

163,194

95.09%

97.50%

D11

36.917

50

164,066

96.05%

D00: neither intervention; D10: A only; D01: B only; D11: A and B. Independence is the deliberately misspecified daily-independence comparator. Feasibility is evaluated under the stipulated process, not a field estimate.

Figure 1 Planned contribution and exact joint capacity feasibility. The dashed line marks the 95 percent requirement. IID fit denotes the daily-independence comparator. Source: calculations from the specified model.

The actual joint feasibility of D00 and D01 is approximately 97.50 percent because assembly is the binding allocated constraint and engineering has considerable additional slack. D10, by comparison, achieves 95.09 percent because both resource loads are close to their respective probability limits. D11 lies between these cases at 96.05 percent. Thus the common 95 percent requirement does not produce identical realised model probabilities. The results reflect a sufficient constraint system with a fixed risk allocation, not direct maximisation over the exact joint feasibility region.

Factorial contrasts and the active resource

The economic response to A depends on B. Adding A to D00 increases net scheduled contribution by 5,701 MU, while adding it to D01 increases contribution by 5,233 MU. Adding B increases contribution by 1,340 MU without A and by 872 MU with A. The difference in differences is -468 MU. This negative interaction contrasts with the addition costs and opportunity specified

allocation decision.

The dual values make this change visible. In D00 and D01, an additional hour of the conservative assembly limit has a local value of 800 MU, while additional engineering capacity has no local value. In D10 and D11, the engineering limit has a local value of 1,600 MU per hour, while assembly has slack. These are marginal values within the current linear programming basis. D10's assembly slack is only 0.022 hours, so its apparent switch to an engineering bottleneck is particularly close to another change in the active constraints. The numerical precision of a dual value should not be mistaken for a wide range of practical validity.

Simulation checks and integer commitments

The 200,000-week calculation agrees with the exact probabilities at the reported Monte Carlo precision. For D11, 192,110 simulated weeks satisfy both requirements, giving 96.055 percent and a 95 percent Wilson interval of 95.969 to 96.139 percent. Table 4 shows the corresponding checks for all six policies. Every analytical probability lies within its reported interval. For D10, the interval extends slightly below 95 percent even though the exact probability is above it. This illustrates the distinction between a known probability under a stipulated model and uncertainty arising from a finite numerical sample. It is not evidence that the analytical constraint has failed.

Table 4 Monte Carlo verification using 200,000 independent simulated weeks

Policy

Exact probability %

Simulation estimate %

95% Wilson interval %

Nominal

88.712

88.682

88.542–88.820

D00

97.500

97.490

97.421–97.558

D10

95.086

95.076

94.980–95.170

D01

97.500

97.466

97.396–97.533

D11

96.052

96.055

95.969–96.139

Intervals describe simulation sampling error conditional on the specified model. They are not confidence intervals for actual factory performance.

Enumerating feasible whole unit quantities preserves the design ordering at the 95 percent target. The best integer commitments are 29, 35, 31, and 36 washing machines for D00, D10, D01, and D11 respectively, with 50 refrigerators in every case. Their net contributions are 156,400, 162,000, 158,600, and 162,600 MU. D11's advantage over D10 falls from 872 MU in the

Reliability sensitivity and model stress

Across target joint reliability levels of 80, 90, 95, 97.5, and 99 percent, D11 remains the highest value design among the four under equal risk allocation. Raising the target from 95 to 99 percent reduces its net continuous contribution from 164,066 to 162,516 MU, a difference of 1,550 MU. Actual model probabilities are greater than or equal to the target because the allocated constraints are sufficient. This comparison describes the price of a stricter commitment within the model; it does not establish which target a production organisation should adopt.

Holding the 95 percent D11 plan fixed exposes stronger sensitivity to model error. If the actual persistence remains 0.6 rather than falling to 0.2, joint feasibility drops to 86.93 percent. Increasing both daily standard deviations by 25 percent gives 90.29 percent feasibility. Reducing both mean weekly capacities by two hours gives 71.83 percent. Combining these changes gives 54.57 percent. Table 5 and Figure 2 show why intervention effectiveness and capacity calibration matter more than the small Monte Carlo interval around the original probability.

Table 5 Stress evaluation of the fixed D11 commitment

Actual process scenario

Lag one correlation

SD multiplier

Mean shift h

Joint feasibility

Specified process

0.2

1.00

1.00

0

86.93%

Daily SD plus 25 percent

0.2

1.25

0

90.29%

Mean capacity minus 2 hours

0.2

1.00

-2

71.83%

Combined stress

0.6

1.25

-2

54.57%

The production quantities are held at 36.9165 washing machines and 50 refrigerators. SD multipliers and mean shifts apply to both resources. No reoptimisation is performed in these stress cases.

Figure 2 Feasibility of the fixed D11 plan as actual temporal persistence changes. Daily marginal variances and mean capacities remain fixed. Source: Equation 5 with the variance in Equation 3.

Arguments

Experimental effects must match the operational

weekly capacity. Testing only whether average machine cycle time falls would leave important parts of the decision unresolved. A faster operation might shift work to engineering, increase rework, or consume time elsewhere. An experiment designed for production planning should record the quantities that enter the resource constraints, together with outcomes that could make the apparent improvement misleading. This follows from the model's accounting structure rather than from a preference for collecting as many variables as possible. A measurement is useful when it can change a coefficient, a constraint, or a stated consequence of the decision.

The experimental unit also needs to follow the mechanism of the intervention. If A is installed once on a production line, thousands of subsequent units provide repeated observations under one installation, not thousands of independent installations. If B is intended to improve recovery from disruptions, a single good week does not identify its effect on persistence. Multiple operating periods under each condition are required, with treatment assignment and background conditions recorded. Blocks can address identifiable differences between lines, shifts, or periods, provided the chosen arrangement leaves the effects of interest estimable. The source manuscript's attention to experimental and observational units becomes operationally consequential at this point (Güneş, 2026a).

A full factorial field study would expose each intervention at both levels of the other, but the four labels alone do not guarantee credible evidence. Carryover may occur if an installation cannot be removed cleanly, training persists after the nominal treatment period, or a recovery policy changes subsequent behaviour. A crossover arrangement then requires an explicit account of washout and residual effects. When reversal is

in a regression.

The current numerical experiment deliberately stops short of this causal claim. It shows what follows if A and B have the stipulated effects. A field experiment would instead estimate those effects and their uncertainty. Maintaining that boundary makes the calculation more useful: it specifies the evidence needed to support the recommendation and identifies the quantities to which the decision is sensitive. A persuasive operational case would therefore include both an estimate of the assembly time change and a defensible account of the change in temporal dependence. Without them, the preferred combination is a candidate for testing, not an established engineering improvement.

Persistence can have economic value without changing the daily mean

Factor B illustrates a form of improvement that a mean comparison can miss. Under the assumed stationary model, the distribution of a single day's capacity deviation is identical before and after the intervention. The difference lies in how one day's deviation relates to the next. Faster recovery prevents a negative deviation from persisting as strongly through the week, while the lower positive persistence also limits strings of favourable deviations. The weekly distribution becomes narrower around the same mean. When a plan is constrained by a lower capacity quantile, that narrowing increases the amount that can be committed at a given reliability level.

This interpretation suggests concrete measurements. A recovery intervention should be evaluated using time ordered records of usable hours, the onset and duration of disruptions, and the timing of restorative actions. Aggregated totals alone may conceal whether an apparent improvement came from higher mean availability, reduced marginal variation, or weaker persistence. Conversely, an estimated lag one

a maintenance team restores capacity.

The horizon matters as well. The experiment uses five days because the original allocation problem is weekly. A daily commitment would place different weight on persistence, while a monthly commitment would accumulate a different set of covariances. There is no universal conversion from a daily coefficient of variation to a reliable amount of capacity at every planning horizon. The variance formula in Equation 3 makes this dependence explicit. For a longer horizon, the terms and their weights must change; for a forecast conditional on the current state, the means and covariance matrix must be recalculated. Reusing a weekly safety allowance in another decision would silently change the event whose probability was controlled.

An intervention that reduces persistence may therefore deserve investment even if a standard test finds no change in the daily average. That proposition does not establish the profitability of B in every setting. Its value depends on the reliability target, the active resource, and the cost of achieving and maintaining the effect. In this example, B's contribution gain after A is only 872 MU net of its assumed cost. Evidence that the process remains stable matters precisely because the advantage is limited. The model directs attention to the duration and recurrence of disturbances, where a mean-only analysis would provide an incomplete account of the production decision.

Bottleneck changes explain the economic interaction

The negative interaction between A and B is a property of the optimised response. Without A, assembly is scarce at the relevant planning limit, so reducing its uncertainty releases capacity with a local value of 800 MU per hour. Once assembly times fall, engineering becomes the limiting resource. B still improves both capacity limits, but

constraint is currently binding.

This is why a factorial study needs a clearly named response variable. Interaction in cycle time, interaction in weekly available hours, and interaction in optimal contribution are different quantities. The -468 MU contrast concerns the last of these. Calling it evidence of a physical incompatibility between assembly improvement and recovery management would be incorrect. It is instead evidence, within the stipulated example, that separately estimated benefits cannot simply be added after each has been valued against a different baseline. Investment proposals often make precisely this aggregation error when departments calculate savings independently and the organisation then sums them without reoptimising the joint plan.

Dual values provide a useful explanation but require care near a basis change. In D10, the assembly slack of approximately 0.02248 hours permits only about 0.01249 additional engineering hours before the implied extra washing machines exhaust that slack. Thus the local engineering value of 1,600 MU per hour cannot be extended to a full hour without resolving the model. D11 has more assembly slack, yet its corresponding range is still limited. The sensitivity discussion in the operations research manuscript is therefore essential to interpreting a shadow price, rather than an optional addition after obtaining the solution (Güneş, 2026c).

The same calculation offers a direct cost threshold. Before deducting B's cost, adding it to A increases scheduled contribution by 1,872 MU in the continuous model. If its weekly equivalent cost exceeded that amount while all physical assumptions remained fixed, D10 would dominate D11 on this objective. Under whole unit commitments, the corresponding break-even cost is 1,600 MU. These thresholds are more decision relevant than an unqualified statement that the combined intervention is best. They show the

failed week

The joint feasibility probability answers a specific question: can both departments supply the scheduled resource requirements during the week? It does not answer how late a shipment will be, how costly an unmet commitment becomes, or whether the organisation has an effective recovery action. A plan with a low frequency of very costly failures can be less attractive than one with more frequent small overruns. The present chance constraint is appropriate for examining the credibility of an advance capacity commitment, but its probability is not a complete economic description of disrupted production. The meaning of the target must remain attached to the event that has actually been modelled.

Nor does weekly aggregate feasibility establish an executable schedule. Engineering work may have to precede assembly, certain tasks may require particular workers, and a resource surplus late in the week may not compensate for a shortage before a due date. The teaching model deliberately aggregates those details. Its probability is therefore a necessary capacity screening measure for a more detailed operational model, not a general on-time-delivery probability. Where sequencing matters, the accepted product mix should be passed to a scheduling model or a discrete event simulation. If that second analysis rejects the mix, the capacity planning model needs an additional restriction or a revised aggregation rule.

The omission of recourse also limits the interpretation of contribution. The nominal plan's 166,000 MU is not an expected profit forecast. Neither is D11's 164,066 MU. Both figures value scheduled quantities before the economic consequences of infeasibility have been introduced. A two-stage extension could select the product commitment first and allow overtime, subcontracting, delayed completion, or cancellation after capacity is observed. The costs

The risk target must consequently come from an operational requirement, contractual obligation, or explicit management preference. Ninety-five percent is an illustrative modelling choice here. A plant serving time critical demand might choose a stricter event and lower violation allowance. A plant with inexpensive buffers might accept more variation while pricing the consequences. Comparing several reliability levels is useful because it reveals how much scheduled contribution the organisation gives up for additional protection. It cannot determine the correct social or commercial tolerance for a missed commitment without further information. These comparisons support a decision; choosing the tolerance remains part of specifying the problem.

Conservative protection should be distinguished from optimal protection

Equal allocation of the violation budget is easy to explain. Each resource is assigned a 2.5 percent allowance, and their sum cannot exceed the 5 percent joint allowance. That transparency is a strength of the demonstration. It is also a restriction. When engineering has large spare capacity, assigning it the same nominal allowance as a tightly constrained assembly department may protect the overall plan more than required. This is why Doo achieves approximately 97.5 percent rather than merely 95 percent model feasibility. The resulting lost contribution is a property of the chosen sufficient formulation, not an unavoidable cost of every 95 percent reliable policy.

An improved allocation could optimise separate risk allowances while maintaining their sum at or below the joint limit. Alternatively, under a defensible joint distribution, the exact probability constraint could be imposed directly. Either change alters the feasible region and may alter the preferred intervention. The present comparison deliberately keeps the allocation rule common

and could distinguish modelling conservatism from physical resource scarcity.

Robust and distributionally robust formulations would address other shortcomings, especially uncertainty about the tails or the distribution itself. Their protection should be described in the language of the formulation, including the uncertainty set or distributional neighbourhood and the evidence used to calibrate it (Bertsimas & Sim, 2004; Mohajerin Esfahani & Kuhn, 2018). Calling every cautious plan robust would obscure these differences. The appropriate comparison is whether each method provides useful protection for the same defined operational event at an acceptable economic cost, evaluated on conditions beyond those used to select the policy.

Numerical precision does not resolve model uncertainty

The simulation check gives a narrow interval around D11's specified probability. That precision is reassuring about the implementation and the finite simulation sample. It says nothing by itself about the true effect of the recovery intervention. When persistence returns to 0.6 in the stress analysis, reliability falls by more than nine percentage points, far exceeding the Monte Carlo interval width. A two hour reduction in both mean weekly capacities is even more damaging. These are changes in the represented system, not random numerical fluctuations. Their consequences cannot be removed by simulating another million weeks from the original assumptions.

Three uncertainty questions must therefore remain separate in interpretation. Random operating variation concerns which week occurs under a given process. Parameter uncertainty concerns how accurately that process and the intervention effects have been estimated. Structural uncertainty concerns whether the chosen process and constraints represent the

not justify treating all the others as known.

Validation should preserve the temporal structure that the model claims to capture. Randomly shuffling daily observations between development and evaluation sets can place neighbouring observations from the same disruption into both, creating an unduly favourable impression of performance. Later operating periods, complete disruption episodes, or separated blocks should be reserved according to the intended use. The evaluation should examine realised resource shortfalls and the stability of intervention effects, not merely the average fit of the capacity series. The particular separation rule depends on the process memory and on how frequently the policy would be revised in practice.

When many candidate designs are investigated, the evaluation period must also be protected from repeated selection. Otherwise, the preferred policy may be the one that benefited most from favourable estimation errors. Smith and Winkler's (2006) analysis of the optimiser's curse provides the relevant warning. A sound sequence would specify the decision criterion, estimate candidate effects, select the policy, and then evaluate it using information not already consumed by that selection. Operational monitoring would follow after adoption. Those stages need not be administratively elaborate, but they must be distinguishable if the reported performance is to be credible.

A practical route from the example to a field study

A field application could begin with the two resource records already implied by the example: usable engineering hours and usable assembly hours, recorded by day together with the product mix and major interruptions. The first task would be to reconcile those hours with actual completed work. Planned staffing hours may include meetings, setup, training, or unavailable equipment, whereas the model requires hours

The intervention experiment would then estimate A and B at the organisational scale at which they can be assigned. Blocks could balance persistent differences between lines or periods, and the protocol would record concurrent changes that could confound the comparison. The study should specify its primary decision outcome in advance, while retaining the intermediate physical outcomes needed to explain it. A sample size calculation would use plausible variability, dependence, and a minimum economically relevant effect; the number of recorded units alone would not supply the number of independent replications. These requirements follow the relationship between experimental structure and analysis emphasised by NIST/SEMATECH (n.d.-b).

After estimation, the same four design options could be reoptimised with their measured coefficients and costs. Parameter uncertainty could be explored through repeated fits or appropriately structured resampling, and whole unit constraints should be included if commitments cannot be divided. An independent operating period would then assess whether the selected plan respects its stated capacity reliability. A detailed schedule would test sequencing and due dates. Where frequent corrective actions are needed, their measured costs should motivate the recourse extension rather than remain hidden outside the model. This sequence would turn the numerical example into an empirical study without changing its central research question.

The contribution of the current work is accordingly methodological and conditional. It provides a small model in which the effects of experimental assumptions, temporal dependence, and resource scarcity can be followed from inputs to decisions. It does not establish the intervention parameters as industry benchmarks, nor does it show that a Gaussian autoregression is generally the best capacity model. The value of the

Argument

This section should be reviewed during editorial preparation. If the original manuscript does not provide this section separately, the editorial office should decide whether the existing article content already covers it or whether author-approved expansion is required.

Conclusion

Experimental design, stochastic processes, and operations research address connected parts of the same production decision. Experiments can identify how controllable settings change resource requirements and operating behaviour. A stochastic process model determines how capacity uncertainty accumulates over the commitment horizon. Optimisation then translates those quantities into a product mix under explicit economic and reliability requirements. The analysis shows that the connection between these stages must be preserved if an attractive numerical plan is to have a defensible operational interpretation.

In the illustrative manufacturing problem, the nominal optimum achieves only 49.47 percent joint capacity feasibility. Introducing a safety allowance under an incorrect daily independence assumption raises that probability to 88.71 percent, still below the intended 95 percent requirement. Correctly representing temporal

scheduled contribution, 164,066 MU, at 96.05 percent model feasibility. Estimates from 200,000 simulated weeks agree with the exact probabilities within Monte Carlo uncertainty, and the preferred design remains unchanged when whole unit quantities are imposed.

The recommendation remains conditional. Its advantage over assembly improvement alone is 872 MU in the continuous model and 600 MU under integer commitments. Cost, persistence, and mean capacity changes can readily erode that margin or invalidate the reliability claim. The factorial analysis also shows that economic interaction can arise from a changed bottleneck even when intervention costs and physical effects are specified separately. A treatment effect measured at one machine cannot therefore be assigned a system value without considering the complete resource allocation problem.

The next empirical step is a factorial or appropriately restricted field design with time ordered capacity records, explicit intervention costs, and evaluation on operating periods withheld from development. The event being protected should remain clear: weekly resource adequacy is different from schedule feasibility and from realised profitability after corrective action. Within that boundary, the proposed integration supplies a reproducible basis for asking which evidence would change a production decision, how much protection it supports, and when the model should be revised. Reliable planning depends on this connection between evidence and commitment as much as on the optimisation algorithm used to compute the plan.



interpretation, and final approval of the submitted work unless otherwise stated.

Conflict of Interest

The author declares no conflict of interest unless otherwise stated in the manuscript or correspondence.

Ethics Statement

The author is responsible for ethical compliance, consent, institutional permission, and research integrity where applicable.

Data Availability

Data availability information was not separately provided unless stated by the author.

Correspondence & ORCID

Correspondence details and ORCID information are displayed where supplied by the author.

References

- Ankenman, B., Nelson, B. L., & Staum, J. (2010). Stochastic kriging for simulation metamodeling. *Operations Research*, 58(2), 371–382. <https://doi.org/10.1287/opre.1090.0754>
- Bertsimas, D., & Kallus, N. (2020). From predictive to prescriptive analytics. *Management Science*, 66(3), 1025–1044. <https://doi.org/10.1287/mnsc.2018.3253>
- Bertsimas, D., & Sim, M. (2004). The price of robustness. *Operations Research*, 52(1), 35–53. <https://doi.org/10.1287/opre.1030.0065>
- Box, G. E. P., & Wilson, K. B. (1951). On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society Series B Methodological*, 13(1), 1–38. <https://doi.org/10.1111/j.2517-6161.1951.tb00067.x>
- Brockwell, P. J., & Davis, R. A. (2016). *Introduction to time series and forecasting* (3rd ed.). Springer. <https://doi.org/10.1007/978-3-319-29854-2>
- Charnes, A., & Cooper, W. W. (1959). Chance-constrained programming. *Management Science*, 6(1), 73–79. <https://doi.org/10.1287/mnsc.6.1.73>
- Dantzig, G. B. (1955). Linear programming under uncertainty. *Management Science*, 1(3–4), 197–206. <https://doi.org/10.1287/mnsc.1.3-4.197>
- Elmachtoub, A. N., & Grigas, P. (2022). Smart “predict, then optimize”. *Management Science*, 68(1), 9–26. <https://doi.org/10.1287/mnsc.2020.3922>
- Glasserman, P., & Yao, D. D. (1992). Some guidelines and guarantees for common random numbers. *Management Science*, 38(6), 884–908. <https://doi.org/10.1287/mnsc.38.6.884>
- Goos, P., & Jones, B. (2011). *Optimal design of experiments: A case study approach*. Wiley. <https://doi.org/10.1002/9781119974017>
- Güneş, H. A. (2026a, May). *Mühendislik deney tasarımı* [Engineering experimental design, chapters 1–10; unpublished doctoral course assignment]. St Clements University, Doctoral Programme in Industrial





course assignment]. St Clements University, Doctoral Programme in Industrial Engineering.

Mandi, J., Kotary, J., Berden, S., Mulamba, M., Bucarey, V., Guns, T., & Fioretto, F. (2024). Decision-focused learning: Foundations, state of the art, benchmark and future opportunities. *Journal of Artificial Intelligence Research*, 80, 1623–1701. <https://doi.org/10.1613/jair.1.15320>

Mohajerin Esfahani, P., & Kuhn, D. (2018). Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations. *Mathematical Programming*, 171, 115–166. <https://doi.org/10.1007/s10107-017-1172-1>

Morris, T. P., White, I. R., & Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. *Statistics in Medicine*, 38(11), 2074–2102. <https://doi.org/10.1002/sim.8086>

NIST/SEMATECH. (n.d.-a). Consequences of non-randomness. In e-Handbook of statistical methods, section 1.2.5.1. <https://www.itl.nist.gov/div898/handbook/eda/section2/eda251.htm>

NIST/SEMATECH. (n.d.-b). How can I account for nested variation restricted randomization? In e-Handbook of statistical methods, section 5.5.5. <https://www.itl.nist.gov/div898/handbook/pri/section5/pri55.htm>

NIST/SEMATECH. (n.d.-c). Randomized block designs. In e-Handbook of statistical methods, section 5.3.3.2. <https://www.itl.nist.gov/div898/handbook/pri/section3/pri332.htm>

Sargent, R. G. (2013). Verification and validation of simulation models. *Journal of Simulation*, 7(1), 12–24. <https://doi.org/10.1057/jos.2012.20>

Shapiro, A., Dentcheva, D., & Ruszczyński, A. (2021). *Lectures on stochastic programming: Modeling and theory* (3rd ed.). Society for Industrial and Applied Mathematics. <https://doi.org/10.1137/1.9781611976595>

Smith, J. E., & Winkler, R. L. (2006). The optimizer's curse: Skepticism and postdecision surprise in decision analysis. *Management Science*, 52(3), 311–322. <https://doi.org/10.1287/mnsc.1050.0451>

Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, 22(158), 209–212. <https://doi.org/10.1080/01621459.1927.10502953>





approved within the Current Science Georgia production workflow.



Authorized Signatory
Current Science Georgia Editorial Office

CSG • VERIFIED • CSG • SECURE • CSG • VERIFIED

CSG • VERIFIED • CSG • SECURE • CSG • VERIFIED



