Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023)
Edited by Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio

doi: 10.3850/978-981-18-8071-1_P160-cd

Analysis of Efficiency in Response Surface Designs Considering Orthogonality Deviations and Cost Models

Marco Arndt, M.Sc.

Institute of Machine Components, University of Stuttgart, Germany.
E-mail: marco.arndt@ima.uni-stuttgart.de

Dr.-Ing. Martin Dazer

Institute of Machine Components, University of Stuttgart, Germany.
E-mail: martin.dazer@ima.uni-stuttgart.de

Orthogonality in DoE favors non-correlated effects and minimizes their confidence intervals. However, accidental or deliberate deviation from orthogonality is often possible on the one hand, and sometimes even desirable within reliability demonstration testing. Based on an investigation of orthogonality deviations (errors in factor levels) with respect to performance quantities such as a priori power and scaled prediction variance, the result of a study for DoE costs is put into perspective in this paper. In the course of this, the accuracy of regression models in CCDs as well as their likelihoods of detecting them correctly are compared to a cost model, which ultimately gives orthogonality deviations a price tag. Thus, favorable combinations of design modifications regarding orthogonality are measured within a trade-off.

Keywords: design of experiments, reliability engineering, cost modelling, test efficiency, orthogonality deviations, statistical power analysis, response surface methodology

1. Introduction

In order to understand industrial applications holistically in terms of their (linear, polynomial) characteristics, testing is ordinarily required to quantify effects and interactions of influencing variables on target quantities. This applies to technical functionalities, performances and reliability of products, cf. Bertsche (2008). Design of Experiments (*DoE*) and Response Surface Methodology (*RSM*) here address challenges within testing intents in great detail. Designed test plans and combinations of factor settings (controlled inputs) are utilized to entirely investigate significant effects *and* interactions regarding an observation objective (response, output), enabling their analysis and quantification in a statistically validated and highly efficient manner Montgomery (2020). For instance, Central-Composite Designs (*CCDs*) determine testing procedures forming standardized test sequences commonly in use within industrial applications and reliability demonstration. Especially for a case where characteristics and the volatility of the target variable to be investigated are unclear - not to mention the requirements for a probably dedicated *optimal* test design as an alternative - CCDs are commonly used in the transfer of statistics knowledge and in engineering applications. In the course of this a key focus is usually placed on statistical quantities of the test design. In order to guarantee evaluability, *orthogonality* in the design matrix is predominantly favored preventing correlations in the effect determination and minimizing confidence intervals of estimated coefficients (effects). Nevertheless, at times the actual implementation of test designs may favor deviations from orthogonality all by itself, which may result from quite pragmatic reasons of adjustability issues for factors in terms of control engineering, cf. Donev (2004), or for the case of a sequential method, some test runs cannot be realized unexpectedly at all, cf. Box and Wilson (1951); Johnson et al. (2011). Assuming this to be prevented, certain efforts are usually required along with general needs for stochastic modeling, but also conflict with economic efficiency being a top priority within industrial implementation. On the other hand, this may alternatively be assumed beneficial - as testing often generates costs that are only favored economically as long as proof can be provided. Deliberately permissible inaccuracies in instrumentation or omitted test runs may offer a cost advantage which could even be desired, tolerating certain deviations in orthogonality. Effects of such precautions on the model quality of regression equations describing the investigation objective may be partly examined already, cf. Arndt et al. (2022); Mell et al. (2022). Complementary to this a generally applicable systematic approximation quantifying the new overall mandatory budget, not to mention potential savings, is missing so far.

With this work a profound cost model is introduced and consequences of experimental design manipulations are shown in the balance sheet. The basic principles of effect estimation, exemplary options for orthogonality deviation and a breakdown of the considered cost sources are introduced. Finally, the cost model and implemented manipulations are then merged so that a relative budget change can be assigned to deviations in test power and model quality for the estimated fit.

2. Test Design Performance Indicators

Within a pragmatic approach, to define appropriate parameters for the evaluation in RSM, the evaluation metrics *regression quality* (cf. Sec. 2.1), *power* (cf. Sec. 2.2) and *prediction quality* (cf. Sec. 2.3) are established as characteristic values for the performance. Therefore, they are briefly described below and implemented as indicators for the goodness of experimental designs.

2.1. *Test Designs and Regression Modeling*

By employing CCDs, linear as well as second-order models can be experimentally assessed and estimated, if it is a matter of describing y as an output based on $i = 1, \dots, n$ observations in a system with main effects, interactions (and quadratic terms), cf. Montgomery (2020).

These can be linearized (multiple linear regression) for $k < n$ regressor variables specified by β_j, $j = 0, 1, \dots, k$ regression coefficients, while leaving ϵ as a random error:

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 \\ \dots + \sum\sum_{i<j=1} \beta_{ij} x_i x_j + \epsilon. \quad (1)$$

For $p = k+1$, in matrix notation, the estimates of β_j remain to be calculated via

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}, \quad (2)$$

where $\boldsymbol{y}$ represents an $n \times 1$ vector of observations, $\boldsymbol{X}$ is an $n \times p$ matrix with the levels of the regressor variables, $\boldsymbol{\beta}$ is a $p \times 1$ vector of regression coefficients and $\boldsymbol{\epsilon}$ is taking into account $n \times 1$ random errors to the following:

$$\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X^T X}\right)^{-1} \boldsymbol{X^T y}. \quad (3)$$

Note that $\boldsymbol{H} = \boldsymbol{X}\left(\boldsymbol{X^T X}\right)^{-1} \boldsymbol{X^T}$ is known as the *hat matrix*, whose properties decisively determine the quality of fitting the observations to a vector of estimated values $\hat{\boldsymbol{y}} = \boldsymbol{Hy}$. For standardised levels (test run settings) of the regressor variables $\boldsymbol{x^T} = [1, x_1, \dots, x_k]$ the fitted regression model then results in

$$\hat{y} = \boldsymbol{x^T} \hat{\boldsymbol{\beta}}. \quad (4)$$

The effectively determined regression coefficients, actually corresponding to estimated effects of factor (regressor variable) level variation, are thus determined at this stage defining the quality of fit. Given that the error $\boldsymbol{\epsilon} \sim$ *NID*$(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$ is normally and independently distributed, here e.g. maximum likelihood estimation (*MLE*) estimates of $\boldsymbol{\beta}$ and ordinary least squares (*OLS*) estimates according to Eq. (3) are equivalent, see Montgomery (2020). In the further analysis the error terms therefore are expected to be *NID*.

So, if the processing of data and the estimation algorithm for the coefficients should detect deviations in effect estimation that originate from the (manipulation associated) setup of the experimental design, design-related modeling quality can be captured.

2.2. *Power of Effect Estimation*

In order to evaluate effects possibly caused by factor-level variation, the principles of hypothesis testing are applied to $\hat{\boldsymbol{\beta}}$. Here appropriate hypotheses are defined as

$$\mathrm{H}_0\colon \forall j : \beta_j = 0 \;\; ; \;\; \mathrm{H}_1\colon \exists j : \beta_j \neq 0. \tag{5}$$

Assuming that a change in y could also occur variance-based without varying the factor levels, H_1 may be incorrectly assumed to be true - the *type-I error* prevails and determines the risk of making a false positive decision. H_0 would therefore have to be rejected if at least one of the k regressor variables x_i contributed a significant effect to the model output y. In contrast, the *type-II* error describes the probability or risk β_p of making a false negative decision to reject an effect (coefficient) that actually and truly exists while adopting H_0. In order to correctly decide this with the tolerance to a significance level α on residual error probability, the observation variances are to be determined by means of ANOVA. Here the mean *e*rror- and *r*egression- *sum of squares* are utilized to calculate the statistic $F_0 = MS_r/MS_e$, resulting in the rejection of H_0 if $F_0 > F_{\alpha,k,n-k-1}$ (alternatively if $p < \alpha$). However, as H_1 is accepted, it remains to be further evaluated by adequacy checking whether the identified *significant* coefficients in modeling generate a satisfactory estimate with respect to the dataset. Observing further the compliment to β_p, the probability results to $\mathrm{P}(reject\ \mathrm{H}_0 \mid \mathrm{H}_1\ is\ true)$ as

$$power = 1 - \beta_p, \tag{6}$$

and therefore to the chance of correctly detecting an existing effect (coefficient). Even *a priori*, power can be considered as a quality measure favoring successful identification of desired effects.

2.3. *Predicting New Observations*

In industrial experimentation the regression fit of Eq. (4) is certainly utilized to predict new system responses, keeping the precision of the prediction quite crucial: what model quality is achieved for the committed test budget - misestimation? It seems to be relatable that for high-cost specimens of low batch sizes, this is a sensitive aspect.

If $\boldsymbol{x_0}^{\boldsymbol{T}} = [1, x_{01}, \ldots, x_{0k}]$ as a particular investigation point is to be considered respecting *NID*, the prediction value y_0 comes along with a prediction variance in the prognosis of $\hat{y}(\boldsymbol{x}_0) = \boldsymbol{x_0}^{\boldsymbol{T}}\hat{\boldsymbol{\beta}}$. Based on the design size N (total amount of test runs), thus the scaled prediction variance (SPV) results in

$$\begin{aligned}\mathrm{SPV}(\boldsymbol{x}_0) &= \frac{N\mathrm{Var}[\hat{y}(\boldsymbol{x}_0)]}{\sigma^2} \\ &= N\boldsymbol{x_0}^{\boldsymbol{T}}\left(\boldsymbol{X^T X}\right)^{-1}\boldsymbol{x}_0.\end{aligned} \tag{7}$$

Scaling the variance by N and σ^2 allows to evaluate the metric on an per observation basis and scale-free. Moreover, it is quite important for the experimenter that the variance of prediction is reasonably stable within the design space and that its maximum is comparatively small. To achieve this, for the instance of a second-order design, CCDs are arranged *rotatably* so that all n_S axial and n_F factorial runs are on a *sphere*. For a rotatable design this determines the Euclidean distance to the center of the investigation area of $\alpha_D = \sqrt[4]{n_F}$ for axial test runs n_S while the radius of the factorial sphere equals $\alpha_D = \sqrt{k}$, cf. Myers et al. (2016). Further, the prediction variance is stabilized by the amount of center runs n_C. A fraction of the recommended number, for instance 1 instead of 5, multiplies $\mathrm{Var}[\hat{y}(\boldsymbol{x})]$ at the design center while $n_C = 0$ turns $\boldsymbol{X^T X}$ into *singular*, maximizing $\mathrm{Var}[\hat{y}(\boldsymbol{x})]$.

3. Test Design Orthogonality in Industrial Experimentation

Orthogonality as a requirement in experimental design determines the column setup of the design matrix and is given as the sum of products of entries $x_{ij} = \{-\alpha_D, -1, 0, 1, \alpha_D\}$ in any two columns equals zero. As a result, it is ensured that the terms are not correlated and that the effects can be analyzed independently to each other. As indicated in Sec. 1, prevailing orthogonality would imply that $\mathrm{Var}[\hat{y}(\boldsymbol{x}_0)]$ according to Eq. 7, in particular, the confidence intervals of the estimated coefficients through $\mathrm{Var}[\hat{\boldsymbol{\beta}}]$ are minimized. That is when the diagonals of $\boldsymbol{X^T X}$ are maximized while its off-diagonals equal zero. Orthogonality in test design is thus easily achieved, provided that CCDs in use are encoded and implemented as specified with respect to leveling and replication.

3.1. *Orthogonality Deviations*

Given this context, deviations from orthogonality are to be considered consciously. Being the most accessible methodology in the industry, CCDs are simply part of the application standard. While implementation, the experimenter encounters accruing experimental costs (test budget), the peculiarities of the object system (system characteristics) and the challenges of instrumentation being very likely to occur within test setups (control and regulation). Each of these challenges may lead to issues in implementing true orthogonality, if not to certain orthogonality deviation in terms of typical setting and measurement errors. In this way, an experimenter may face the concerns of: *variance* in test setup adjustments - random disparities due to measurement or control instruments; *shifts* in parameter settings - errors based on a systematic offsets in setup; and *omissions* of test runs/points - as single test runs may unexpectedly not be realizable at all. For their simulative implementation, a nonconforming system behavior can be generated using a polynomial equivalent to Eq. 1, where the representation of an independent (measurement) error, a certain amount of variance, should be taken into account by $\boldsymbol{\epsilon_\beta}, \boldsymbol{\epsilon_{y_i}} \sim NID(\mathbf{0}, \sigma^2\boldsymbol{I})$ - as Donev (2004); Ardakani et al. (2011); Johnson et al. (2011); Arndt et al. (2022); Mell et al. (2022) have previously thought of. Consequently, this creates virtual test results that are attributed to the relevant test designs. The simulation approach upon which this work is based is therefore described next.

3.2. *Statistical Study Approach*

Complementary to a comparison of impacts to test design characteristics on response prognosis (prediction variance, cf. Eq. 7), a simulation for effects by measurement equipment needs to be performed separately. This alone is to capture estimation and measurement errors utilizing metrics such as stated in Sec. 2. Therefore a two-dimensional RSM is chosen, testable virtually as the *Default Model* by orthogonal and non-orthogonal CCD structures. For both a measurement error ϵ_{y_i} and an effect error ϵ_β affecting the coefficient values (effects) β_j, σ is numerically superimposed. As this is exceedingly case-specific, to illustrate one possibility among many and for the sake of comparability these properties are set as stated in Table 1. In line with this, multiple linear regression is then utilized to capture an averaged estimate for the coefficients using OLS, MC = $1e5$ Monte Carlo iterations and non/-orthogonal CCDs, which is intended to represent a numerically averaged general result. As part of the same process the $power$ is calculated as the rate recorded via MC for each $p < \alpha$ case per coefficient. Likewise, for each MC iteration, the cost level as well as the estimated values of the coefficients are calculated and arithmetically averaged. The simulation process thus corresponds to the scheme according to Fig. 1.Corresponding to a standard CCD as shown in Fig. 2, the orthogonal version and derivative orthogonality manipulations according to Section 3.1 capture the deviations in coefficient estimations and in SPV (orthogonal vs. manipulated). An illustration of realistic practical arrangements is provided by following deviations:

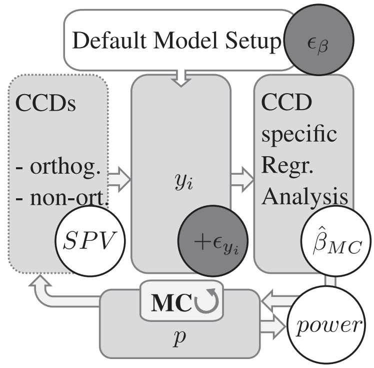

Fig. 1. MC Simulation Study Approach: CCDs (orthogonal, non-orthogonal) analyzed regarding SPV, estimated regression coefficients $\hat{\beta}_{MC}$, deviations $\triangle\beta_j$ and $power$.

Table 1. Coefficients and Standard Deviation for *Default Model* Setup

β_0	β_j	σ
10	10	5

- variances $[0, 10]$, uniform dispersion ranges around each star point - representing control or measurement errors in critical conditions of factor-level combinations;
- shift $[1, \sqrt[4]{n_F}]$, varying a_D of the star points (0,a_D) and ($-a_D$,0), as these should presumably take the lowest and highest value in runtime - representing systematic offset;
- omission of test runs, mapped to the deletion of the star points (0,a_D) and ($-a_D$,0) or to the variation of repetitions $n_C = [0, 8]$ in (0,0) - to illustrate the omission of replications or critical factor combinations/runtimes.

4. Design of Expe-*nses*

Test beeing performed appropriately to design specifications in DOE, a large number of cost factors are incurred that have to be listed in the balance sheet. Even with a fixed number of test runs and replications, these are strongly dependent on - above all - runtimes, especially speaking of end-of-life (*EoL*) reliability testing. Notably, for comparability purposes, it is assumed that such tests are performed sequentially and individually. Nevertheless, a reasonable structuring of the costs is practicable and links between runtime and energy demand can be defined in relation to time-based and power-based cost aspects. Thus, if costs are to be estimated generically for a relation-based assessment and *independent* of the specimen type, Montgomery (2020) may allow a cost analyses based on his multiple examples as follows.

4.1. *Cost Analysis*

According to Sec. 3.2, the *relative* cost change respecting the best performable design is crucial. Therefore, it is consistent to equate an ideal implementation with maximum achievable precision, realization completeness, and protracted runtime/energy as charging the highest conceivable overall costs ($C_{\sum}$) (best instrumentation setup, perfect control equipment, complete runtimes, etc.). However, absolute costs of specimens, regardless of their actual cost amount, do not need to be included in a relative estimate since they are consistent. Thus,$C_{\sum}$ comprises three main domains, each of which can be detailed again:

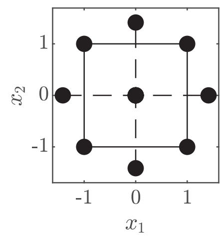

Fig. 2. Setup of a $k = 2$ factor orthogonal CCD with x_1 and x_2, distance of axial runs $\alpha_D = \sqrt[4]{n_F}$.

- size of testing C_S,
  - staff C_O;
  - auxiliary material C_H;
  - energy costs C_E;
  - maintenance cost C_M;
- test equipment C_T;
  - test bench hardware C_B;
  - measurement equipment (quality) C_I;
  - control technology (quality) C_C;
- ambient conditions C_A;
  - laboratory condition costs C_L.

Within the scope of this work, price ranges for low- to high-quality technical test equipment and a representative, exemplary cost assumption for scaling cost factors are determined on the basis of research into the European market (it should be mentioned here that the respective absolute values are insubstantial, exemplary, and interchangeable for relationship-based modeling). At any rate, within this model the most distinctive orthogonality deviation is assigned to the smallest identified hardware costs in each case: low-cost test conditions equal strong deviations from perfect test plans. In order to model their relationships as well, two additional steps are also performed: cost source interactions are captured using links identified by creative methods such as the design structure matrix (*DSM*) to even track a passive cost change in the presence of correlation; it is assumed that low stress levels resulting from the test factors lead to higher run times and thus higher expenses. To account for the latter, each test run is assigned a percentage runtime reduction relative to the most time-consuming run in an oriented

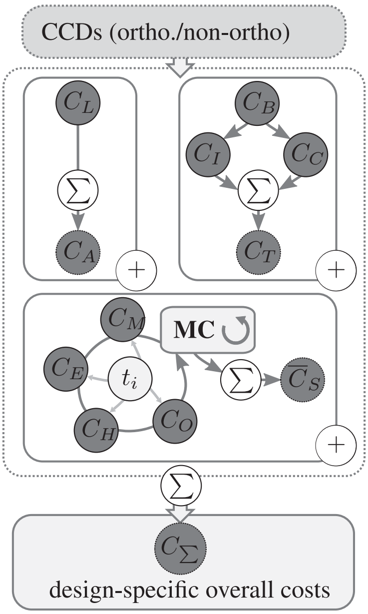

Fig. 3. Cost model to calculate overall costs $C_{\sum}$ as a function of test design implementations (orthogonal, non-orthogonal), identified cost factors and random experimental runtimes t_i.

manner and iterated over MC to an average. A representative maximum runtime is specified exemplary. This implies that, for instance, ($-a_D$,0), (0,$-a_D$) and (-1,-1) are in a time intensive range, whereas (a_D,0), (0,a_D) and (1,1) represent short runners. Everything in between has a σ-random EoL runtime that impacts C_E, C_M, C_H and C_O. This particularly takes into account that the omission of a long-runner also realizes greater savings in C_S, respectively energy, staff and maintenance costs. Together with the cost structure, this is shown in Fig. 3. With respect to the orthogonality deviation mentioned in Sec. 3.2, thus:

- maximum variance for star runs corresponds to the least expensive equipment represented in C_T and C_A;
- shifts of factor combinations to low stress ranges correspond to increased runtime and higher operating costs in C_S;
- the omission of test runs corresponds to the omission of single C_S with a redistribution of per-run costs by keeping same fixed costs C_T.

As a result, it is therefore possible to calculate the overall cost of the test design implementation for each orthogonality deviation and to relate these to $C_{\sum}$ of ideal orthogonality.

5. Balancing Sheet on Orthogonality

According to the performance parameters specified in Sec. 2 a relative cost change is respectively compared to the change in percentage points (pp) of $power$, $\hat{\beta}$ and SPV$|_{r=1}$ for each orthogonality deviation. A tolerance limit of $5pp$ is taken into account in all evaluations; smaller values are thus not considered or displayed graphically. In the first instance, the influence of variance to star points can be analyzed comparing Fig. 4. Based on present cost modeling, the results show $C_{\sum}$ savings of $5pp$. This can be explained with the underlying cost structure of the test instruments. From a qualitative perspective, low gray cost bars and flat deviation curves are thus very desirable for the present and other diagrams. While the SPV$|_{r=1}$ exemplified in the sphere with Eucl. distance ($r = 1$) remains constant, only the $power$ values for the coefficient of β_0 show a deviation of approx. $7pp$ on average. *Key findings:* Under the assumed instrument costs, toleration of a scatter range $> 5\%$ for star runs decreases $C_{\sum}$ by approximately $> 5\%$, while the accuracy in the prediction as well as the certainty for effect detection remains largely unaffected.

Considering the replication amount of the center point n_C (Fig. 5), the SPV$|_{r=1}$ is minimized at $n_C = 3$. $n_C < 5$ decreases the estimated costs

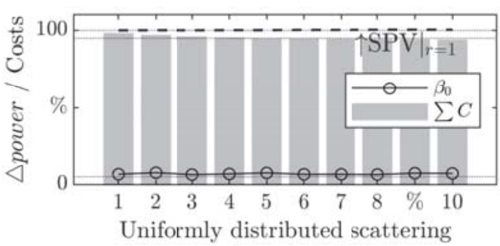

Fig. 4. Performance Indicators and overall costs subject to varying global, uniformly distributed scattering of star points: Percentage deviation $\triangle power$ of β_0; $C_{\sum}$; design-based SPV$|_{r=1}$ - normalized to $k = 2$, $n_C = 5$ CCD.

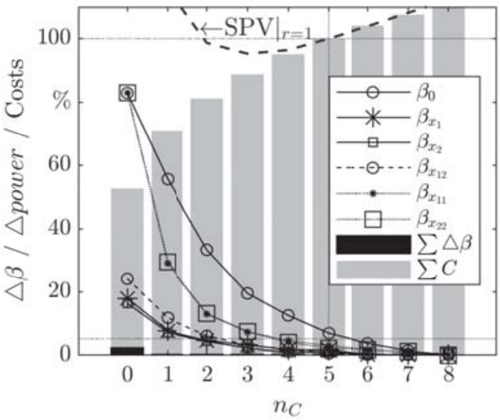

Fig. 5. Performance Indicators and overall costs subject to varying n_C: Percentage deviation $\triangle power$ per coefficient; percentage deviation of cumulative coefficient amount $\triangle\beta = |\hat{\beta}_{MC} - \beta_j|$ (*black bars*); $C_{\sum}$; design-based SPV$|_{r=1}$ (SPV$|_{n_C=0}$ = NaN) - normalized to $k = 2$ CCD, $n_C = 5$ as default.

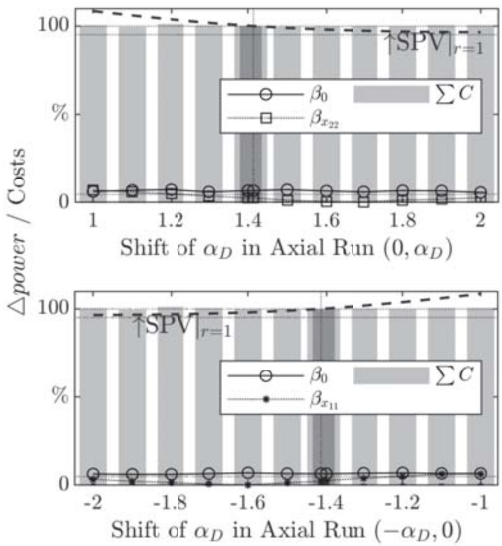

Fig. 6. Performance Indicators and overall costs subject to varying α_D shift in axial runs: Percentage deviation $\triangle power$ per coefficient; $C_{\sum}$; design-based SPV$|_{r=1}$ - normalized to $k = 2$, $n_C = 5$ CCD.

by $> 5pp$ compared to the reference, although the $power$ of the average to the system response β_0 decreases by $> 12pp$, too. Concurrently, $n_C > 5$ steadily enhances $C_{\sum}$ starting from 104% while not improving the indicators in any way. With $n_C \leq 3$, the $power$ of quadratic terms also deviates to the point where, if center points are completely omitted ($n_C = 0$), up to $83pp$. Here even the absolute value of the coefficient estimates β_j in total starts to deviate increasingly (*black bars*). *Key findings:* The realization of $n_C = 3$ is quite reasonable, considering here: up to $12pp$ $C_{\sum}$ may be saved, SPV is minimized, the error probability does not increase more than $20pp$ for β_0; attempting $\leq$ 5% $power$ deviation on the other hand results in an unreasonable increase in cost for $n_C > 5$.

Compared to this, the axial shifts of the star points $(0, \alpha_D)$, $(-\alpha_D, 0)$ do not appear to strongly affect SPV$|_{r=1}$, nor $C_{\sum}$, nor the coefficient deviation, cf. Fig 6. Accordingly, the overlaid MC average has a more dominant effect with respect to runtime costs, keeping $C_{\sum}$ nearly constant. It is only the $power$ values of the intercept and the square parts that vary occasionally more than $5pp$, which is why only these are displayed. One should also note that with orthogonality deviations at selected star points, the performance indicators behave complementary to each other as expected. In any case, deviations in the absolute value of the model coefficients are not found here. *Key findings:* Star point shifts are generally manageable and optimize the SPV$|_{r=1}$ with increasing proximity toward the forecast point.

In the last instance, the effects of omitting the previously addressed star points are investigated (Fig. 7). In each case, a substantial decrease in $C_{\sum}$ is observed, which may be attributed to the assumed distribution of run times and thus test duration costs (C_S). The SPV$|_{r=1}$ increases by $17pp$ to $38pp$, while especially the $power$ values of the quadratic coefficients $\beta_{x_{11},x_{22}}$ strongly decrease and that of the intercept β_0 varies the least. As expected, the $power$ values behave in a complementary manner between the two omissions $(0, \alpha_D)$ and $(-\alpha_D, 0)$, but an absolute deviation of the coefficient values $\triangle\beta$ is not observed here either. *Key findings:* While omitting (particularly expensive) starpoints may have a strong saving

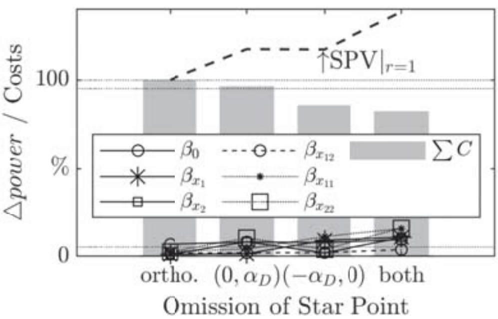

Fig. 7. Performance Indicators and overall costs subject omission of axial runs: Percentage deviation $\triangle power$ per coefficient; $C_{\sum}$; design-based SPV$|_{r=1}$ - normalized to $k = 2$, $n_C = 5$ CCD.

effect, one must be aware that the effect detection for quadratic terms can be affected strongly and the variance in the prediction is shifted.

5.1. *General Findings*

From the accounting in Sec. 5, the following insights emerge under given assumptions about the system and cost modeling for the specific given model: a variance in star points of up to 10% has no noticeable effect on the performance indicators and can be almost freely accepted; when implementing center point replications, a sharp balance must be made between cost effort and model quality: the power and model quality may decrease, but appealing cost benefits could result; the manner of star point realization must be evaluated depending on the desired model terms.

6. Summary and Conclusion

With this work, effects of orthogonality deviations in CCDs are presented and compared from statistical as well as economic points of view. Therefore, a study approach virtually creating and testing of a prescribed system with manipulated CCDs as well as a cost model is presented. The results of this study are generalized from the average of Monte Carlo iterations and compared. This allows manipulated orthogonality at CCDs (default: $k = 2$, $n_C = 5$) to be given a quality tag via the presented performance indicators as well as a price tag via the cost model. Although the results deliberately depend on the individual assumptions (evaluation point of the SPV$|_{r=1}$ as a metric, variances, predetermined effects, cost-structure and -dimensions), the results thus show plausibly evaluable phenomena and allow a targeted assessment of more efficient experimental designs for industrial applications.

References

Ardakani, M. K., D. Das, S. S. Wulf, and T. J. Robinson (2011). Estimation in second-order models with errors in the factor levels. *Commun. in Statistics 40*(9), 1573–1590.

Arndt, M., P. Mell, and M. Dazer (2022). Generic effects of deviations from test design orthogonality on test power and regression modelling of central-composite designs. *Proceedings of the Probabilistic Safety Assessment and Management Conference 2022 PSAM 16*.

Bertsche, B. (2008). *Reliability in Automotive and Mechanical Engineering: Determination of Component and System Reliability* (1st ed.). VDI-Buch. Springer.

Box, G. E. P. and K. B. Wilson (1951). On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society. Series B (Methodological) 13*(1), 1–45.

Donev, A. N. (2004). Design of experiments in the presence of errors in factor levels. *Statistical Planning and Inference 126*(2), 569–585.

Johnson, R. T., D. C. Montgomery, and B. A. Jones (2011). An expository paper on optimal design. *Quality Engineering 23*(3), 287–301.

Mell, P., M. Arndt, and M. Dazer (2022). Non-orthogonality in test design: practical relevance of the theoretical concept in terms of regression quality and test plan efficiency. *Proceedings of the Probabilistic Safety Assessment and Management Conference 2022 PSAM 16*.

Montgomery, D. C. (2020). *Design and Analysis of Experiments* (10th ed.). Wiley.

Myers, R. H., D. C. Montgomery, and C. M. Anderson-Cook (2016). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments* (4 ed.). Hoboken, NJ, USA: John Wiley & Sons, Inc.