*Operational Research in Engineering Sciences: Theory and Applications*
*Vol. 6, Issue 2, 2023, pp. 140-159*
*ISSN: 2620-1607*
*eISSN: 2620-1747*
*DOI: https://doi.org/10.31181/oresta/060208*

# A NOVEL ENSEMBLE BASE MODELS SELECTION APPROACH FOR ESTIMATING CREDIBLE RUBBER CONVEYOR BELTING CURE TIMES FROM A SMALL SAMPLE SIZE MRSM DATASET

**Domingo Pavolo[1*], Delson Chikobvu[1], Louis Lagrange[2]**

[1] Department of Mathematical Statistics and Actuarial Sciences, Faculty of Natural and Agricultural Sciences, University of the Free State, P.O. Box 339, Bloemfontein 9300, South Africa
[2] Department of Engineering Sciences, Faculty of Natural and Agricultural Sciences, University of the Free State, P.O. Box 339, Bloemfontein 9300, South Africa

Received: 07 April 2023
Accepted: 08 May 2023
First Online: 09 July 2023

*Research Paper*

***Abstract***: *Multiple response surface methodology (MRSM) is small sample dataset analytics in nature. There are experiences associated with small sample size dataset problems in regression models, model selection and generalisability of models that affect the MRSM solution credibility. In this work, ensembling is used to account for the small sample size problems of an MRSM dataset and to avoid the tradition of simultaneously optimising only single "best" models for each response. A novel ensemble base model(s) selection methodology is used to manage the number of simultaneous optimisation computations. Fifteen model selection criteria are used to vote for the "best" fitting and parsimonious model and, with it, all response models nesting it are included as base models of the ensemble. Simultaneous optimisations, and frequency analysis of solutions and model complexities are performed to arrive at a solution. The ensemble solution is estimated by weighted averaging of simultaneous optimisation solutions using the solution frequencies. When the methodology is applied to the rubber covered conveyor belt problem of* Pavolo and Chikobvu (2022)*, the same estimated credible results are obtained, albeit with a fewer simultaneous optimisations computations. The results also show that the "best" fitting and parsimonious model is an under-fit, and its solution is different from the credible ensemble solution. The approach is recommended for similar small sample size problems.*

**Keywords:** *Multiple response surface methodology, Ensembling, Small data analytics, Simultaneous optimisation, Model selection criteria, Nesting.*

*Corresponding author: dpavolo68@gmail.com (D. Pavolo)
Chikobvu@ufs.ac.za (D. Chikobvu), LagrangeLF@ufs.ac.za (L. Lagrange)

# 1.0 Introduction

## 1.1 The Nature of Multiple Response Surface Methodology

Multiple response surface methodology (MRSM) uses mathematical and statistical tools to simultaneously optimise multiple responses of a process. MRSM is small sample data analytics in nature to alleviate the cost of experimentally generating large data. The current MRSM conceptual framework selects a single "best" model for each response for simultaneous optimisation and this approach experiences credibility problems associated with uncertainties related to regression modelling and model selection as both are data dependant, hence the interest in ensembling. Wang (2008) defined an ensemble as a system of individual models/algorithms working in parallel and whose outputs are combined by a suitable fusion strategy to produce a single answer for a given problem.

In statistics, small sample size datasets are always a challenge to work with in regression modelling, model selection, optimisation and prediction. The credibility of models obtained from regressing small sample size MRSM datasets is suspect and generalisability gets worse the smaller the size of the dataset (Xu and Goodacre, 2018) (Jenkins & Quintana-Ascencio, 2020; Rawlings, Pantula, & Dickey, 1998). Zucchini (2000) mentions that small sample size datasets produce many competing candidate models with a measure of goodness of fit to the small sample size dataset without a clear best model because data will be insufficient to effectively approximate a single "best" response model of the "truth". The problem of model under- and overfitting is therefore common. Ensembling, according to Dietterich (2000), caters for working with multiple good competing models reminiscent of Lieberman's (2009) Common Task Framework. The advantages of ensembling theory are yet to be meaningfully exploited in MRSM. This work employs ensembling in MRSM as a way of accounting for small sample size problems, and as it has been successfully used in machine learning (Ahangi, Langroudi, Yazdanpanah, & Mirroshandel, 2019; Hu, Zhou, Liu, & Tang, 2019; Yang, Hwa Yang, B Zhou, & Y Zomaya, 2010).

There are so many classical model selection (MS) criteria available in literature (Burnham & Anderson, 2002; Claeskens & Hjort, 2008; Schomaker & Heumann, 2020). The selection of a best fitting and parsimonious model would itself suffer from uncertainty because MS criteria do not always agree (Schomaker & Heumann, 2020) and there is also small sample size inefficiency to worry about (Hurvich and Tsai, 1989). Since MRSM involves multiple responses, the credibility of the results is affected by simultaneous optimisation of several single "best" models selected with uncertainty. The selection of single "best" models for each response using classical MS criteria, a tradition in MRSM, is avoided in this work. Empirical studies have shown that ensembling improves on single models in regression problems (Bauer and Kohavi, 1999) (Banfield, Hall, Bowyer, & Kegelmeyer, 2006; Breiman, 2001; Dietterich, 2000; Sohn & Shin, 2007). Practically, predictive performances of single models have been improved by ensembling in various application fields (Kazienko, Lughofer, & Trawiński, 2013; Menahem, Shabtai, Rokach, & Elovici, 2009; Merkwirth et al., 2004; Polikar et al., 2008; Yang et al., 2010).

There are many methodologies for simultaneous optimisation of response models. The desirability function (Costa, Lourenço, & Pereira, 2011; Derringer, 1994) and loss function (Murphy et al., 2005) approaches are popular among MRSM practitioners (Bakhtiarifar, Bashiri, & Amiri, 1999). Other approaches include compromise programming (Govindaluri & Cho, 2007), goal programming \ (Kovach, Cho, & Antony, 2008; Xu, Lin, Tang, & Xie, 2004), inspection of contour plots (Sivertsen et al., 2007), physical programming (Kovach, Cho, & Antony, 2008), Bayesian probability-based (Peterson, Miro-Quesada, & del Castillo, 2009), neural networks (Noorossana, Davanloo Tajbakhsh, & Saghaei, 2009), vectorial optimisation (Ortiz, Sarabia, Herrero, & Sánchez, 2006), weighted Tchebycheff formulations (Dächert, Gorski, & Klamroth, 2012) and modified ε-constraint method (Lee, Kim, & Köksalan, 2012). Though these approaches have their own merits, computational complexity and lack of clear recommendations for proper use are major reasons by which they are of little practical use or appeal to non-statistician practitioners (Costa & Lourenço, 2014). At the same time, they are based on the traditional simultaneous optimisation of single "best" models for each response, which ignores the problems related to small sample size MRSM dataset generated models and the use of data dependant classical MS criteria. Pavolo and Chikobvu (2022) employed an ensembling methodology that totally avoided the use of classical model selection criteria to obtain credible cure time estimates for a rubber covered conveyor belting cure time problem, albeit, after performing a relatively large number of simultaneous optimisations.

The aim of this work is to reduce the number of ensemble base models from that of the multiple simultaneous optimisation ensemble of Pavolo and Chikobvu (2022), yet maintaining credibility in the solution estimate. The novel methodology of (i) selecting a "best" fitting, parsimonious and prediction model using multiple mixed (best fit, parsimonious and prediction) MS criteria, and (ii) adding all possible response models nesting the model to the set of ensemble base models achieved the same credible result with fewer base models to Pavolo and Chikobvu (2022) rubber covered conveyor belting cure time estimation problem using the same small sample size MRSM dataset (Annexure 1). The results suggest that the best fitting and parsimonious model is probably an under-fit and this contributes to the error of optimism.

The rest of the paper is organised as follows: Section 2.0 presents the materials and solution methodology in more detail; Section 3.0 presents the results and discussions; Section 4.0 presents the conclusion; and the list of References is in the final Section.

## 2.0 The Materials and Methods

### 2.1 The Rubber Covered Conveyor Belting Problem

Rubber covered conveyor belting is used in agriculture, manufacturing and mining industries for cost effective conveyancing of bulk inputs or products. The major failures of conveyor belting are breakage, component separation and rubber cover excessive wear (Bortnowski, Kawalec, Król, & Ozdoba, 2022; Zimroz & Król, 2009). Under design of conveyor belting break strength or over loading during operation result in breakage. Belting component separation is due to bonding failure. Rubber cover wear is resisted by compound hardness. For every conveyor belt manufactured, break strength, adhesion and cover hardness are measured and checked for conformance with customer specifications and product quality standard requirements. More belting life is obtained by increasing cover hardness and adhesion to prevent early conveyor belting failure. This may result in changes to the specifications of cover and skim compounds (see Figure 1 for a rubber covered conveyor belting construction structure).

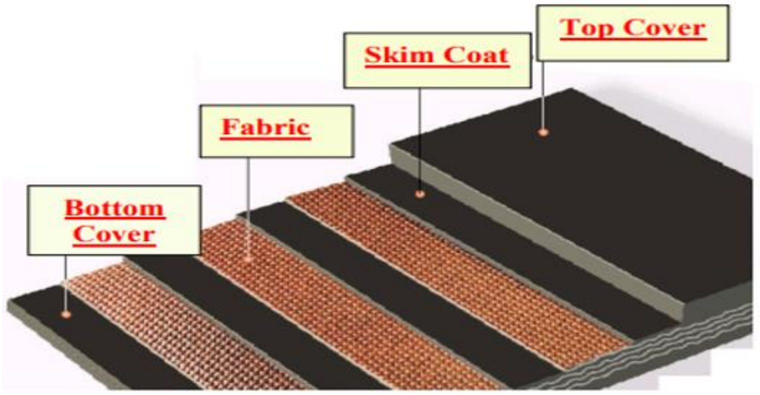


*Figure 1: Components of a rubber covered conveyor belt.*
*Source: image.png (499×305) (uniqueconveyorbelting.com)*

The rubber conveyor belt is constructed at a calendaring machine, then vulcanised at either an intermittent or continuous press. The general process model of the vulcanising process is shown in Figure 2.

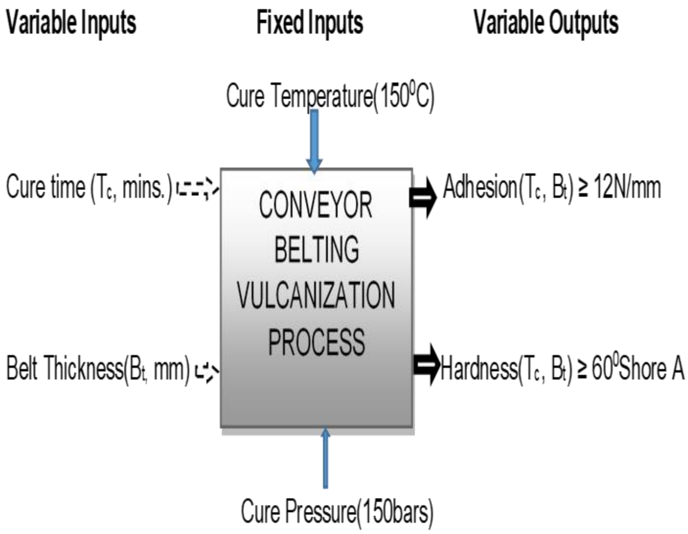


*Figure 2: Vulcanization process model showing the desired minimum quality requirements.*
*Source: Author's own compilation*

The two fixed inputs of the vulcanising process are cure temperature ($150^0$C) and cure pressure (150bars). The variable inputs are total rubber thickness (mm) and cure time (minutes). The total rubber thickness is the sum of the top and bottom cover thicknesses and the total skim thickness. The cover and skim compounds have different specifications and cure properties. The skim compound specification determines the adhesion value of the belting components which resists separation. The synthetic fabric determines the breaking strength of the conveyor belting. In a case reported in Pavolo and Chikobvu (2022), the minimum requirements for adhesion and cover hardness are 12N/mm and $60^0$ Shore hardness, respectively. The client wants a table to present to the Shop floor in a work instruction of the form:

*Table 1: Showing the tabular form required in work instruction*

| Rubber Thickness, $R_i$ (mm) | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cure Time, $T_i$ (min.) | $T_7$ | $T_8$ | $T_9$ | . | . | . | . | . | . | . | . | . | . | $T_{20}$ |

Source: Pavolo and Chikobvu (2021)

In Table 1, for a rubber thickness $R_i = i(mm)$, the cure time is $T_i$. The cure time estimation problem is a multiple response problem with two variable inputs, rubber thickness and cure time, and two responses, adhesion and cover hardness. The cure times for each of the given conveyor belting total rubber thicknesses were required to be determined for the vulcanisation process.

### 2.2 Modelling the problem

The cure time estimation problem is an intractable (np-hard) constrained optimisation problem of the form:

$\boldsymbol{Minimise}\ T_i = f(R_i)$, (1)
Given
$Adhesion\ (T_i, R_i) \geq 12N/mm$, (2)
$Hardness(T_i, R_i)\ \geq 60\ Shore\ hardness$, (3)
with $R_i$ as it is given in Table 1.

To solve the above problem requires that the two constraint relationships of adhesion and hardness with cure time and rubber thickness be established first. This is done by first running an MRSM experiment, to obtain a small sample size MRSM dataset which is then used in regression modelling to obtain the relationships below.

$$Adhesion\ = \beta_0 + \beta_1 T_i + \beta_2 R_i + \ \beta_{12}(T_i * R_i) + \beta_{11}{T_i}^2 + \beta_{22}{R_i}^2 + \varepsilon, \quad (4)$$
$$Cover\ hardness\ = \beta_0 + \beta_1 T_i + \beta_2 R_i + \ \beta_{12}(T_i * R_i) + \beta_{11}{T_i}^2 + \beta_{22}{R_i}^2 + \varepsilon, \quad (5)$$
Or simply using model notation $[T_i, R_i, (T_i * R_i), {T_i}^2, {R_i}^2]$

where $\beta_0$is the intercept and $\beta_1$, $\beta_2$, $\beta_{12}$, $\beta_{11}$, and $\beta_{22}$ are estimates of parameters for cure time ($T_i$), and rubber thickness ($R_i$), the cure time by rubber thickness interaction term ($T_i * R_i$), and the second order terms (${T_i}^2$) and (${R_i}^2$), respectively. The two response models are then simultaneously optimised to obtain the minimum integral value of $T_i$ that simultaneously give an adhesion of 12N/mm and cover hardness of $60^0$ Shore A hardness for each rubber thickness of Table 1.

### 2.2 The theory of ensembling

The main theory behind ensembling is directed towards bias-variance-covariance decomposition. Geman, Bienenstock, and Doursat (1992) performed a bias-variance decomposition of a single regression model as shown:

$$E\{[\hat{f}_i - \mathrm{E}(f)]\}^2 = \ [\mathrm{E}(\hat{f}_i) - \mathrm{E}(f)]^2 \ + \mathrm{E}\{[\hat{f}_i - \mathrm{E}(\hat{f}_i)]\}^2. \quad (6)$$
where $\hat{f}_i$ is the fitted value and $\mathrm{E}(f)$ is the expected value of the function $f$. This decomposition can be reduced to
$$MSE(f) = (\mathrm{bias}(f))^2 + \mathrm{var}(f). \quad (7)$$

Ueda and Nakano (1996) then produced the bias-variance-covariance decomposition of an ensemble. In this decomposition it is assumed that:

$$f_{ens} = \frac{1}{k} \times \sum_{i=1}^{k} (\hat{f}_i), \qquad (8)$$

where $f_{ens}$ is the ensemble mean square error and $k$ is the number of base models in the ensemble. Then

$$E\{[f_{ens} - E(f)]\}^2 = Bias^2 + (\frac{1}{k}) \times Variance + (1 - \frac{1}{k}) \times Covariance, \qquad (9)$$

where

$$Bias = \frac{1}{k} \times \sum_{i=1}^{k} [\hat{f}_i - \mathrm{E(f)}]\,, \qquad (10)$$

$$Variance = \frac{1}{k} \times \sum_{i=1}^{k} \mathrm{E}[\hat{f}_i - \mathrm{E}(\hat{f}_i)]^2, \qquad (11)$$

$$Covariance \frac{1}{k(k-1)} \times \sum_{i=1}^{k} . \left\{ \sum_{j=1,(i\neq j)}^{k} \mathrm{E}\left[\hat{f}_i - \mathrm{E}(\hat{f}_i)\right]\left[\hat{f}_j - \mathrm{E}(\hat{f}_j)\right] \right\}. \qquad (12)$$

This suggests that an increase in prediction performance can be expected if it is possible to design an ensemble with low-correlated individual models, low bias and low variance. The stochastic discrimination theory (Kleinberg, 1990), the margin theory (Schapire et al., 1998), and the strength-correlation theory Bernard, Heutte, and Adam (2010); (Breiman, 2001), have all been proved to agree with the bias-variance-covariance decomposition theory (Reu et al., 2016). Hansen and Salamon (1990) concluded that a necessary and sufficient condition for an ensemble of models to be more accurate than any of its constituent individual members is if the base models are both accurate and diverse.

Bias – Variance Trade-off Characteristic

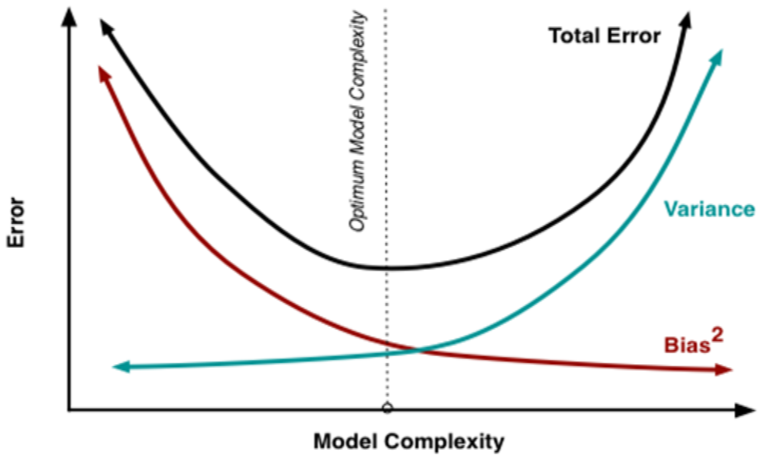


*Figure 3: The variation of bias and variance with the model complexity.*
*Source: http://scott.fortmann-roe.com/docs/BiasVariance.html*

The bias – variance trade-off characteristic, Figure 3, shows that with under-fitted models the bias is high and falls asymptotically with model complexity whilst variance is low and increases with complexity (Emmert-Streib & Dehmer, 2019). Overfitting, is then the addition of more predictors to the optimum complexity model resulting in the increase of variance (diversity) but continued reduction of bias (improvement in accuracy) in accordance with equation (2).

In the same line of argument, an ensemble with a set of base models including the optimum complexity model plus selected overfitted models nesting it should give an accurate predicted solution with averaged minimum or no bias and variance.

In this work a novel method of pruning potential competing response models is employed in which (i) fifteen classical MS criteria (see Annexure 2) are used to vote for the best fitting and parsimonious adhesion response model from the all possible regression modelling set (see Annexure 3), then (ii) adding all other models that have this response model nested in them to the set of base models of the ensemble. This reduces the compliment of ensemble base models drastically from the multiple simultaneous optimisation ensemble of Pavolo and Chikobvu (2022), shown in Annexure 4, implying a fewer number of simultaneous optimisations performed to obtain the same credible cure time estimates.

The solution methodology is summarised in Figure 4.

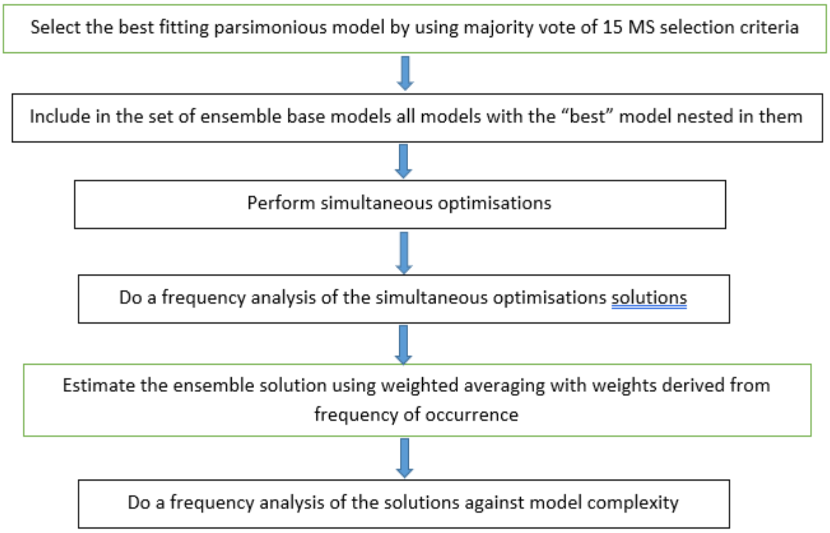


*Figure 4: The flow chart for the solution methodology*

A frequency analysis of the solutions against model complexity is performed in the analysis of under- and overfitting.

## 3.0 Results

The results of the best fitting parsimonious adhesion response model selection are given in Table 2. The fifteen classical model selection criteria are shown in the first column of the table. The six adhesion response models that had at least one MS criterion selection are shown in the first row in summary notation. Each column of the response models has the corresponding criterion values. The selection criteria values are shown in red in each column. For example, in the second row, the classical prediction MS criteria $R^2$ (pr.) selected the adhesion response model $[(T_i * R_i), {R_i}^2]$ as best. The VOTES row shows the total selections for each response model. The model with the highest selections is the adhesion response model $[(T_i * R_i), {R_i}^2]$, with ten votes and thus the best fitting and parsimonious model, having been chosen by the MS criteria $R^2$ (pr.), Cp-k, PRESS, AIC, BIC, $AIC_c$, $KIC_c$, KIC, MKIC and TIC.

The adhesion response model $[(T_i * R_i), {R_i}^2]$ is shown in its expanded form as equation (6). This summary and expanded concept is applicable to all models used in this paper.

$$Adhesion = 10.497 + 0.0203(T_i * R_i) - 0.0258{R_i}^2 \quad (13)$$

*Table 2: Voting for the best fitting and parsimonious model*

| MODEL | $[T_i, R_i, (T_i * R_i), {T_i}^2, {R_i}^2]$ | $[T_i, (T_i * R_i), {T_i}^2, {R_i}^2]$ | $[R_i, (T_i * R_i), {R_i}^2]$ | $[(T_i * R_i), {R_i}^2]$ | $[T_i, R_i, {R_i}^2]$ | $[T_i, (T_i * R_i)]$ |
|---|---|---|---|---|---|---|
| $R^2$ (pr.) | 26.5 | 51.5 | 49.9 | 65.4 | 49.4 | 52 |
| Adeq. pr. | 10.4 | 4.1 | 5.7 | 1.8 | 11.9 | 5.4 |
| Cp-k | 1.0 | 0.1 | 0 | 0 | 2.1 | 0 |
| PRESS | 88.1 | 59.3 | 81.3 | 42.1 | 62.2 | 58.6 |
| AIC | 11.7 | 9.8 | 11.6 | 9.8 | 13 | 17.3 |
| BIC | 15.1 | 12.6 | 13.9 | 11.5 | 15.3 | 19 |
| AICc | 20.2 | 14.8 | 14.3 | 11 | 15.7 | 18.5 |
| APCp | 2.9 | 2.4 | 2.6 | 2.9 | 2.6 | 4.0 |
| SBC | 1.9 | 1.7 | 4.9 | 4.8 | 6.1 | 11.9 |
| HQc | 1.1 | 1.1 | 4.3 | 4.5 | 5.6 | 11.5 |
| KICc | 87.3 | 64.4 | 51.2 | 38.1 | 52.6 | 45.6 |
| HQ | 0.5 | 0.5 | 4.2 | 4.2 | 5.6 | 11.7 |
| KIC | 20.7 | 17.8 | 18.6 | 15.8 | 20 | 23.3 |
| MKIC | 18.2 | 12 | 9.1 | 5.4 | 9.8 | 10.7 |
| TIC | 13.7 | 11.8 | 13.6 | 11.8 | 15 | 19.3 |
| VOTES | 2 | 6 | 1 | 10 | 1 | 1 |

The same best fitting and parsimonious model is presented in Table 3 in another summarised form.

*Table 3: Best fitting and parsimonious response model*

| MODEL | $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\beta}_{12}$ | $\hat{\beta}_{11}$ | $\hat{\beta}_{22}$ |
|---|---|---|---|---|---|---|
| $[(T_i * R_i), {R_i}^2]$ | 10.49700 | | | 0.02025 | | -0.02579 |

Table 4 shows the seven adhesion response models, with the best fitting and parsimonious model in the first row, that nest the adhesion response model $[(T_i * R_i), {R_i}^2]$ as shown with the yellow shading in the summarised form similar to Table 3. It is important to note that the columns with the coefficients of the intercept, the $(T_i * R_i)$ and the ${R_i}^2$ terms have values for all the models. This is the confirmation for the nesting.

*Table 4: Response models nesting the "best" fitting and parsimonious model*

| MODEL | $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\beta}_{12}$ | $\hat{\beta}_{11}$ | $\hat{\beta}_{22}$ |
|---|---|---|---|---|---|---|
| $[(T_i * R_i), {R_i}^2]$ | 10.49700 | | | 0.02025 | | -0.02579 |
| $[T_i, (T_i * R_i), {R_i}^2]$ | 9.14000 | 0.09100 | | 0.01559 | | -0.02242 |
| $[R_i, (T_i * R_i), {R_i}^2]$ | 11.08000 | | -0.09800 | 0.02078 | | -0.02309 |
| $[(T_i * R_i), {T_i}^2, {R_i}^2]$ | 10.39000 | | | 0.01890 | 0.00054 | -0.02485 |
| $[T_i, R_i, (T_i * R_i), {R_i}^2]$ | 8.61000 | 0.10700 | 0.04700 | 0.01450 | | -0.02309 |
| $[T_i, (T_i * R_i), {T_i}^2, {R_i}^2]$ | 1.95000 | 0.75900 | | 0.01676 | -0.01491 | -0.02336 |
| $[R_i, (T_i * R_i), {T_i}^2, {R_i}^2]$ | 11.21000 | | -0.11300 | 0.02150 | -0.00026 | -0.02317 |
| $[T_i, R_i, (T_i * R_i), {T_i}^2, {R_i}^2]$ | 0.74000 | 0.80400 | 0.09700 | 0.01450 | -0.01510 | -0.02476 |

Cure time estimations were obtained by simultaneous optimisation of the adhesion response models with the hardness response model $[T_i, R_i, (T_i * R_i), {T_i}^2]$, the only hardness response model which has a conforming response surface (Pavolo &

Chikobvu, 2022), and is shown in its expanded form in equation (14).

$$Hardness = 29.100 + 2.84T_i - 0.918R_i + 0.0321(T_i * R_i) - 0.061T_i^2 \quad (14)$$

Table 5 shows the cure time estimations obtained by simultaneous optimisation of the adhesion response models with the hardness response model $[T_i, R_i, (T_i * R_i), T_i^2]$. The shading in yellow shows where there is majority agreement of estimated cure times in each column of a given rubber thickness. As an example, for the column $i$ (mm) = 19mm rubber thickness the first seven adhesion response models agree on the cure time estimate of 29 minutes, with the last response model estimating 28 minutes.

*Table 5: Cure time estimates after simultaneous optimisation*

| $i$ (mm) | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $[(T_i * R_i), R_i^2]$ | 21 | 22 | 22 | 23 | 23 | 24 | 24 | 24 | 25 | 25 | 26 | 27 | 29 | 30 |
| $[T_i, (T_i * R_i), R_i^2]$ | 21 | 22 | 22 | 23 | 23 | 24 | 24 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| $[R_i, (T_i * R_i), R_i^2]$ | 21 | 22 | 22 | 23 | 23 | 24 | 24 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| $[(T_i * R_i), T_i^2, R_i^2]$ | 21 | 22 | 22 | 23 | 23 | 24 | 24 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| $[T_i, R_i, (T_i * R_i), R_i^2]$ | 21 | 22 | 22 | 23 | 23 | 24 | 24 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| $[T_i, (T_i * R_i), T_i^2, R_i^2]$ | 21 | 22 | 22 | 23 | 23 | 24 | 24 | 24 | 25 | 25 | 26 | 27 | 29 | 30 |
| $[R_i, (T_i * R_i), T_i^2, R_i^2]$ | 21 | 22 | 22 | 23 | 23 | 24 | 24 | 24 | 25 | 26 | 27 | 28 | 29 | 30 |
| $[T_i, R_i, (T_i * R_i), T_i^2, R_i^2]$ | 21 | 22 | 22 | 23 | 23 | 24 | 24 | 24 | 25 | 25 | 25 | 27 | 28 | 31 |

Frequency analysis of the estimated solutions in Table 5 shows three competing solutions coded $S_1$, $S_2$ and $S_3$ in Table 6 and occurring 1, 2, and 5 times respectively.

*Table 6: Frequency analysis of solutions*

| $i$ (mm) | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | Frequency |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $S_1$ | 21 | 22 | 22 | 23 | 23 | 24 | 24 | 24 | 25 | 25 | 25 | 27 | 28 | 31 | 1 |
| $S_2$ | 21 | 22 | 22 | 23 | 23 | 24 | 24 | 24 | 25 | 25 | 26 | 27 | 29 | 30 | 2 |
| $S_3$ | 21 | 22 | 22 | 23 | 23 | 24 | 24 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 5 |

The three solutions in Table 6 and their frequencies are graphed in Figure 5 for a visual presentation of their distribution.

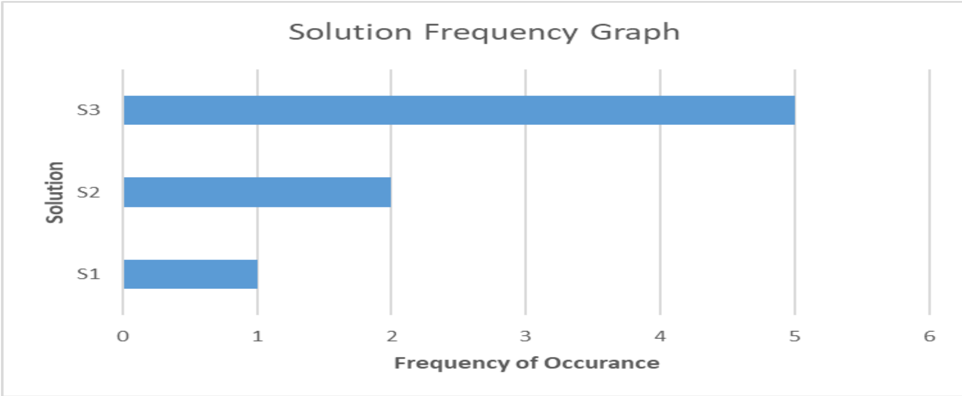


*Figure 5: The solution frequency graph*

Using ensembling, the frequencies 5, 2, and 1 (sum = 8) are converted to weights $\frac{5}{8} = 0.625, \frac{2}{8} = 0.250$ and $\frac{1}{8} = 0.125$ respectively, where the sum of the weights is 1. For any rubber thickness $R_i$, the weighted average cure time is computed as:

$$Wtd.Average = Wt_1 * S_1 + Wt_2 * S_2 + Wt_3 * S_3 \quad (15)$$

where $Wt_i$ is the weight for solution $S_i$, and $\sum_1^n Wt_i = 1$. For example, for a rubber thickness $R_{20}$ (of 20mm), the weighted average is computed as:

$Wtd.Average = 0.125 * 31 + 0.25 * 30 + 0.625 * 30 = 30$

*Table 7: Weighted average estimation of cure times*

| $R_{t\ (mm)}$ | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | Wt. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $S_1$ | 21 | 22 | 22 | 23 | 23 | 24 | 24 | 24 | 25 | 25 | 25 | 27 | 28 | 31 | 0.125 |
| $S_2$ | 21 | 22 | 22 | 23 | 23 | 24 | 24 | 24 | 25 | 25 | 26 | 27 | 29 | 30 | 0.25 |
| $S_3$ | 21 | 22 | 22 | 23 | 23 | 24 | 24 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 0.625 |
| Wtd. Average | 21 | 22 | 22 | 23 | 23 | 24 | 24 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | |

The ensemble estimated weighted average solution in Table 7 is equivalent to the highest frequency solution **$S_3$** in Table 6 since this is the most dominant solution. The ensemble theoretical accuracy is shown in Table 8.

*Table 8: Showing the ensemble theoretical accuracy at simultaneous optimisation*

| | Adhesion | | | | Hardness | | | |
|---|---|---|---|---|---|---|---|---|
| Model | MSPE | Bias | Var. | Covar | MSPE | Bias | Var. | Covar. |
| $[(T_i * R_i), R_i^2]$ | 0.143 | 0.385 | 0.061 | | 0.184 | 0.402 | 0.023 | |
| $[T_i, (T_i * R_i), R_i^2]$ | 0.142 | 0.354 | 0.016 | | 0.096 | 0.285 | 0.014 | |
| $[R_i, (T_i * R_i), R_i^2]$ | 0.151 | 0.370 | 0.014 | | 0.096 | 0.285 | 0.014 | |
| $[(T_i * R_i), T_i^2, R_i^2]$ | 0.105 | 0.301 | 0.015 | | 0.096 | 0.285 | 0.014 | |
| $[T_i, R_i, (T_i * R_i), R_i^2]$ | 0.099 | 0.277 | 0.022 | | 0.096 | 0.285 | 0.014 | |
| $[T_i, (T_i * R_i), T_i^2, R_i^2]$ | 0.495 | 0.619 | 0.113 | | 0.076 | 0.250 | 0.012 | |
| $[R_i, (T_i * R_i), T_i^2, R_i^2]$ | 0.127 | 0.332 | 0.016 | | 0.096 | 0.285 | 0.014 | |
| $[T_i, R_i, (T_i * R_i), T_i^2, R_i^2]$ | 0.156 | 0.616 | 0.137 | | 0.065 | 0.220 | 0.012 | |
| Ensemble Accuracy | 0.177 | 0.407 | 0.049 | 0.021 | 0.101 | 0.287 | 0.015 | 0.023 |

The ensemble theoretical accuracy values at simultaneous optimisation for both adhesion and hardness are fairly accurate at MSPE = 0.177, bias = 0.407, variance = 0.049, covariance = 0.021 and MSPE = 0.101, bias = 0.287, variance = 0.015 and covariance =0.023, respectively.

An analysis of the model complexity with the highest number of solutions can indicate which complexities maybe under- or over-fit. Figure 6 shows the frequency distribution of each model complexity among the solutions.

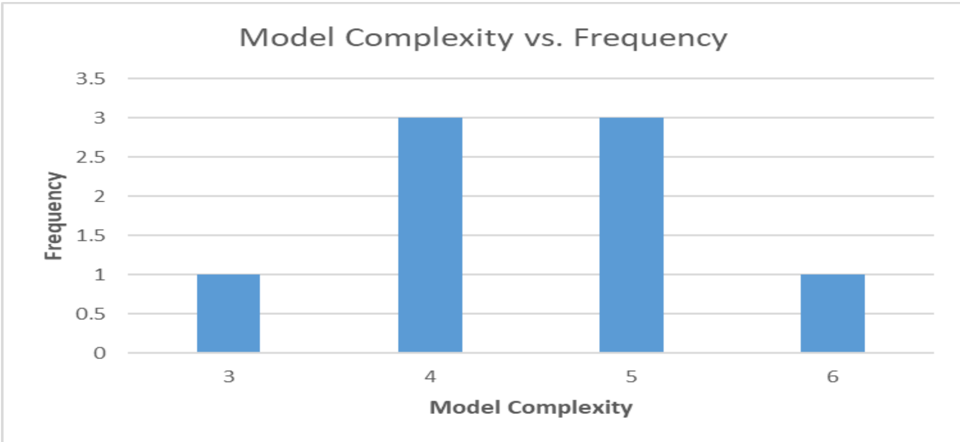


*Figure 6: Showing the model complexity frequency graph*

It is observed in Figure 6 that:

- There is only one adhesion response model with a model complexity of 3.
- That both the model complexities of 4 and 5 have three models each.
- There is only one model with a complexity of 6, the full model.

When solution complexity, e.g. S3(5) implying solution $S_3$ of complexity 5, of adhesion response models is sorted in increasing order of MSPE and Bias at simultaneous optimisation and plotted as a bar chart, as shown in Figure 7, important observations emerge.

| Solution | MSPE | Bias |
|---|---|---|
| S3(5) | 0.099 | 0.277 |
| S3(4) | 0.105 | 0.301 |
| S3(5) | 0.127 | 0.332 |
| S3(4) | 0.142 | 0.354 |
| S3(4) | 0.151 | 0.370 |
| S2(3) | 0.143 | 0.385 |
| S1(6) | 0.156 | 0.616 |
| S2(5) | 0.495 | 0.619 |

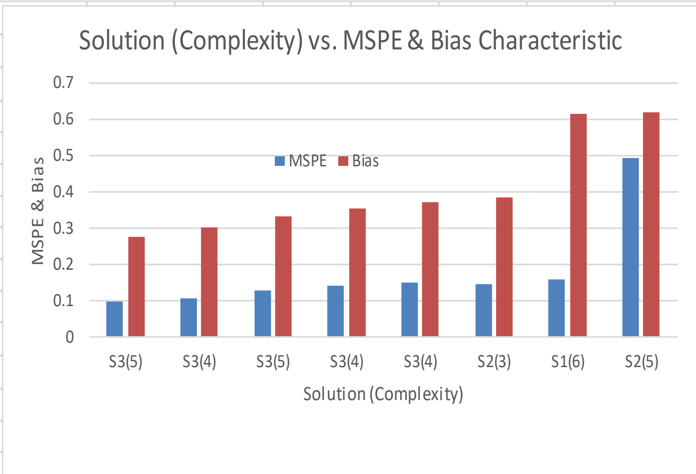


*Figure 7: Showing the trend of Solution (Complexity) plotted against MSPE and Bias at simultaneous optimisation.*

It is observed in Figure 7 that:

i) Solution $S_3$ has the least MSPE and Bias and hence is the best solution estimate to the problem. However, complexity 5 adhesion response models with solution $S_3$, have better average MSPE and Bias than complexity 4 models of the same solution.
ii) The solution of the “best” fit and parsimonious adhesion response model, $S_2$ with complexity 3, has a larger MSPE and bias than all the five adhesion response models with solution estimate $S_3$. This could be the result of the compromise that occurs at simultaneous optimisation with the selected hardness model manifesting as the error of optimism.
iii) Solution $S_1$ which is from the adhesion full model of complexity 6, and solution $S_2$ from the adhesion response model with model complexity 5 appear to be over-fit.

## 4.0 Discussion

Results vs. objective of reduction of number of simultaneous optimisations

The proposed methodology has shown that it is possible to reduce the number of simultaneous optimisations performed by the Multiple Simultaneous Optimisations Ensemble of Pavolo and Chikobvu (2022) whilst obtaining the same result and with similar accuracy. This reduces computational time and opens possibilities for generalisation of the methodology to other more complex simultaneous optimisation problems.

### 4.1 Results vs. objective of maintaining credibility.

The current MRSM conceptual framework selects a single best model for each response for simultaneous optimisation and accommodates model selection criterion uncertainty and hence model uncertainty. The proposed methodology accounts for model selection uncertainty by voting for the most parsimonious model with the best fit to the dataset using fifteen model selection criteria. Uncertainty is further neutralised by ensembling multiple results obtained from simultaneous optimisation of models nesting the selected best model. Ensembling takes advantage of a point raised by Zucchini (2000) that “small sample size datasets produce many competing candidate models with a measure of good fitness to the small sample size dataset without a clear best model because data will be insufficient to effectively approximate a single best”. This is because ensembling the competing good models by making them base models of the same ensemble minimises loss of information contained in them.

The fact that five of the eight estimated results were similar and also similar to the ensemble result further buttresses the credibility. At the same time, the fact that the ensemble estimated result is similar to the result obtained by the Multiple Simultaneous Optimisations Ensemble of Pavolo and Chikobvu (2022) shows the credibility of the proposed methodology.

### 4.3The model complexity perspective and error of optimism

Of interest is how the majority voted best fit and parsimonious model failed to give the credible result. The failure to agree with the majority predicted and ensemble result has been explained in machine learning literature as the error of optimism. The inefficiency of a response model with best fit to a dataset to predict beyond the dataset is termed the error of optimism and gets worse the smaller the sample size.

Complexity analysis makes the best adhesion response model appear as an under-fit. This is because MSPE, in this case, estimates prediction capability within the context of simultaneous optimisation, not just good fitness to the small sample size MRSM dataset used to learn the model, hence exposing the error of optimism. The prediction performance is, though, affected by simultaneous optimisation compromise and the hardness response model it simultaneously optimises with.

### 4.4 Methodology generalisability

The critical question is for the generalisability of the proposed methodology to more complex small sample size MRSM simultaneous optimisation problems. Generalisability to (i) two response problems with multiple good models for each response then (ii) multiple response problems in general.

The current problem had one response with multiple good models and a second response with a single good model. However, most two-response small sample size MRSM simultaneous optimisation problems have multiple good models for each response. The methodology is applicable to each response before simultaneous optimisations and ensembling of results. Coding the methodology in R can make the process of obtaining the solution much faster.

Multiple response problems with two or more responses are obviously a challenge. Research needs to be done on the best way to obtain simultaneously optimised estimated results for ensembling. This may require research into how other methodologies like Derringer's (1994) desirability functions can be used to simultaneously obtain optimised results for ensembling. Artificial neural networks could be another possible methodology.

## 5.0 Conclusion

The method of pruning base models for the ensemble by (i) selecting the best fitted and parsimonious response model using multiple model selection criteria voting, then (ii) adding all the other response models with the "best model" nested in them reduces the number of simultaneous optimisations performed to the multiple simultaneous optimisations ensemble of Pavolo and Chikobvu (2022).
It is also noted that,

(a) Response models can be classified using model complexity as under- or over-fit with a simple complexity analysis.

(b) The credible solution does not necessarily come from the "best" fit and parsimonious model as chosen by majority vote of classical MS criteria. This is a manifestation of the error of optimism in which the best model to fit the small sample size MRSM dataset is inefficient in generalisation.

This method of selecting base models for an ensemble for simultaneous optimisation is recommended for similar small sample size MRSM problems.

## 6.0 Implications of the Study

### 6.1 To the Conveyor Belting Manufacturing Industry

The study proposes a simple methodology of estimating credible conveyor belting curing times for productivity and quality for different total conveyor belting rubber thicknesses. Every time product quality requirements change which then causes the change of conveyor belting rubber compound specifications, cure times have to be re-estimated. The methodology becomes handy in all such incidences.

### 6.2 To generality of practitioners

To the generality of practitioners, the proposed approach is a novel and credible way of estimating process parameters from a small sample size MRSM dataset and it accounts for small sample size inefficiencies and problems that affect the accuracy of simultaneous optimisation results.

### 6.3 To academia in generality

The effect of the error of optimism on estimated results is demonstrated in the context of response models learned from an MRSM small sample size dataset through a simple model complexity analysis. This phenomenon is indeed a problem and should be accounted for in every MRSM small sample size dataset problem.

## 7.0 Limitations and Future Research Directions

### 7.1 Limitations

The proposed methodology was compared with the Multiple Simultaneous Optimisations Ensemble of Pavolo and Chikobvu (2022). It would gain more acclaim if it is compared with other methodologies as well as using other datasets. At the same time, the methodology needs to be tested with more complex examples.

A simulation would also need to be done to buttress credibility of methodology in such small sample size problems.

### 7.2 Future Research Directions

There is need to research on how the methodology fares with more complex problems, that is, (i) problems with more responses than two and (ii) problems with more than one good response models for each response.

Further research is also needed to look at how simulation methodology can be used to justify credibility in small sample size MRSM problems.

There is also need to look at how ensembles of other available simultaneous optimisation methodologies can fare with small sample size MRSM datasets.

## References

Ahangi, A., Langroudi, A. F., Yazdanpanah, F., & Mirroshandel, S. A. (2019). A novel fusion mixture of active experts algorithm for traffic signs recognition. Multimedia Tools and Applications, 78, 20217-20237. https://doi.org/10.1007/s11042-019-7391-0

Bakhtiarifar, M. H., Bashiri, M., & Amiri, A. (1999). Optimization of multi-response problems with continuous functional responses by considering dispersion effects.

Scientia Iranica, 25(4), 2267-2281. https://doi.org/10.24200/sci.2017.4458
Banfield, R. E., Hall, L. O., Bowyer, K. W., & Kegelmeyer, W. P. (2006). A comparison of decision tree ensemble creation techniques. IEEE transactions on pattern analysis and machine intelligence, 29(1), 173-180. https://doi.org/10.1109/TPAMI.2007.250609
Bauer, E. and Kohavi, R (1999) An empirical comparison of voting classification algorithms: Bagging, Boosting and variants. Machine Learning, 36(1/2): 105-139
Bernard, S., Heutte, L., & Adam, S. (2010). A study of strength and correlation in random forests. In Advanced Intelligent Computing Theories and Applications: 6th International Conference on Intelligent Computing, ICIC 2010, Changsha, China, August 18-21, 2010. Proceedings 6 (pp. 186-191). Springer. https://doi.org/10.1007/978-3-642-14831-6_25
Bortnowski, P., Kawalec, W., Król, R., & Ozdoba, M. (2022). Types and causes of damage to the conveyor belt–Review, classification and mutual relations. Engineering Failure Analysis, 140, 106520. https://doi.org/10.1016/j.engfailanal.2022.106520
Breiman, L. (2001). Random Forests. Machine learning, 45, 5-32. https://doi.org/10.1023/A:1010933404324
Burnham, K. P., & Anderson, D. R. (2002). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer New York, NY. https://doi.org/10.1007/b97636
Claeskens, G., & Hjort, N. L. (2008). Model Selection and Model Averaging (Vol. 330). Cambridge University Press. https://doi.org/10.1017/CBO9780511790485
Costa, N. R., Lourenço, J., & Pereira, Z. L. (2011). Desirability function approach: a review and performance evaluation in adverse conditions. Chemometrics and intelligent laboratory systems, 107(2), 234-244. https://doi.org/10.1016/j.chemolab.2011.04.004
Costa, N. R., & Lourenço, J. A. (2014). Optimization criteria ability to depict Pareto frontiers. In Proceedings of the World Congress on Engineering (Vol. 2). https://www.iaeng.org/publication/WCE2014/WCE2014_pp958-961.pdf
Dächert, K., Gorski, J., & Klamroth, K. (2012). An augmented weighted Tchebycheff method with adaptively chosen parameters for discrete bicriteria optimization problems. Computers & Operations Research, 39(12), 2929-2943. https://doi.org/10.1016/j.cor.2012.02.021
Derringer, G. C. (1994). A balancing act-optimizing a products properties. Quality progress, 27(6), 51-58. https://cdnm.statease.com/pubs/derringer.pdf
Dietterich, T. G. (2000). Ensemble methods in machine learning. In International workshop on multiple classifier systems (pp. 1-15). Springer. https://doi.org/10.1007/3-540-45014-9_1
Emmert-Streib, F., & Dehmer, M. (2019). Evaluation of regression models: Model assessment, model selection and generalization error. Machine learning and knowledge extraction, 1(1), 521-551. https://doi.org/10.3390/make1010032
Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. Neural computation, 4(1), 1-58. https://doi.org/10.1162/neco.1992.4.1.1
Govindaluri, S. M., & Cho, B. R. (2007). Robust design modeling with correlated quality characteristics using a multicriteria decision framework. The International Journal of Advanced Manufacturing Technology, 32, 423-433. https://doi.org/10.1007/s00170-005-0349-6
Hansen, L. K., & Salamon, P. (1990). Neural network ensembles. IEEE transactions on pattern analysis and machine intelligence, 12(10), 993-1001. https://doi.org/10.1109/34.58871
Hu, R., Zhou, S., Liu, Y., & Tang, Z. (2019). Margin-based Pareto ensemble pruning: an ensemble pruning algorithm that learns to search optimized ensembles. Computational intelligence and neuroscience, 2019, 7560872.

https://doi.org/10.1155/2019/7560872
Hurvich, C., Tsai, C. (1989) Regression and Time Series Model Selection in Small Samples, Biometrika, 76, 297-307
Jenkins, D. G., & Quintana-Ascencio, P. F. (2020). A solution to minimum sample size for regressions. PloS one, 15(2), e0229345. https://doi.org/10.1371/journal.pone.0229345
Kazienko, P., Lughofer, E., & Trawiński, B. (2013). Hybrid and Ensemble Methods in Machine Learning. Journal of Universal Computer Science, 19(4), 457-461. https://doi.org/10.3217/jucs-019-04
Kleinberg, E. M. (1990). Stochastic discrimination. Annals of Mathematics and Artificial intelligence, 1, 207-239. http://machine-learning.martinsewell.com/ensembles/Kleinberg1990.pdf
Kovach, J., Cho, B. R., & Antony, J. (2008). Development of an experiment-based robust design paradigm for multiple quality characteristics using physical programming. The International Journal of Advanced Manufacturing Technology, 35, 1100-1112. https://doi.org/10.1007/s00170-006-0792-z
Lee, D.-H., Kim, K.-J., & Köksalan, M. (2012). An interactive method to multiresponse surface optimization based on pairwise comparisons. IIE transactions, 44(1), 13-26. https://doi.org/10.1080/0740817X.2011.564604
Liberman, M. (2010) Fred Jelinek. Comput. Linguist, 36(4):595-599
Menahem, E., Shabtai, A., Rokach, L., & Elovici, Y. (2009). Improving malware detection by applying multi-inducer ensemble. Computational Statistics & Data Analysis, 53(4), 1483-1494. https://doi.org/10.1016/j.csda.2008.10.015
Merkwirth, C., Mauser, H., Schulz-Gasch, T., Roche, O., Stahl, M., & Lengauer, T. (2004). Ensemble methods for classification in cheminformatics. Journal of chemical information and computer sciences, 44(6), 1971-1978. https://doi.org/10.1021/ci049850e
Murphy T. Tsai, K., Allen, J. (2005). A Review of Robust Design Methods for Multiple Responses. Research in Engineering Design, 15(4), pp 201-215
Noorossana, R., Davanloo Tajbakhsh, S., & Saghaei, A. (2009). An artificial neural network approach to multiple-response optimization. The International Journal of Advanced Manufacturing Technology, 40, 1227-1238. https://doi.org/10.1007/s00170-008-1423-7
Ortiz, M., Sarabia, L., Herrero, A., & Sánchez, M. (2006). Vectorial optimization as a methodological alternative to desirability function. Chemometrics and intelligent laboratory systems, 83(2), 157-168. https://doi.org/10.1016/j.chemolab.2005.11.005
Pavolo, D., & Chikobvu, D. (2022). Estimating Rubber Covered Conveyor Belting Cure Times Using Multiple Simultaneous Optimizations Ensemble. Operational Research in Engineering Sciences: Theory and Applications, 5(1), 90-106. http://dx.doi.org/10.31181/oresta180222016p
Peterson, J. J., Miro-Quesada, G., & del Castillo, E. (2009). A Bayesian reliability approach to multiple response optimization with seemingly unrelated regression models. Quality Technology & Quantitative Management, 6(4), 353-369. https://doi.org/10.1080/16843703.2009.11673204
Polikar, R., Topalis, A., Parikh, D., Green, D., Frymiare, J., Kounios, J., & Clark, C. M. (2008). An ensemble based data fusion approach for early diagnosis of Alzheimer's disease. Information Fusion, 9(1), 83-95. https://doi.org/10.1016/j.inffus.2006.09.003
Rawlings, J. O., Pantula, S. G., & Dickey, D. A. (1998). Applied regression analysis: a research tool. Springer. https://doi.org/10.1007/b98890

Schapire R.E., Freund, Y., Bartlett, P., Lee, W. (1998). Boosting the margin: A new explanation for the effectiveness of voting methods. The Annals of Statistics, 26(5), 1651 - 1686

Schomaker, M., & Heumann, C. (2020). When and when not to use optimal model averaging. Statistical papers, 61(5), 2221-2240. https://doi.org/10.1007/s00362-018-1048-3

Sivertsen, E., Bjerke, F., Almøy, T., Segtnan, V., & Næs, T. (2007). Multivariate optimization by visual inspection. Chemometrics and intelligent laboratory systems, 85(1), 110-118. https://doi.org/10.1016/j.chemolab.2006.05.005

Sohn, S. Y., & Shin, H. (2007). Experimental study for the comparison of classifier combination methods. Pattern Recognition, 40(1), 33-40. https://doi.org/10.1016/j.patcog.2006.06.027

Ueda, N., & Nakano, R. (1996). Generalization error of ensemble estimators. In Proceedings of International Conference on Neural Networks (ICNN'96) (Vol. 1, pp. 90-95). IEEE. https://doi.org/10.1109/ICNN.1996.548872

Wang, W. (2008). Some fundamental issues in ensemble methods. In 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence) (pp. 2243-2250). IEEE. https://doi.org/10.1109/IJCNN.2008.4634108

Xu, K., Lin, D. K., Tang, L.-C., & Xie, M. (2004). Multiresponse systems optimization using a goal attainment approach. IIE transactions, 36(5), 433-445. https://doi.org/10.1080/07408170490426143

Xu, Y., Goodacre, R. (2018) On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning. J Anal Test. 2018; 2(3):249-262. DOI: 10.1007/s41664-018-0068-2. Epub 2018 Oct 29. PMID: 30842888; PMCID: PMC6373628.

Yang, P., Hwa Yang, Y., B Zhou, B., & Y Zomaya, A. (2010). A review of ensemble methods in bioinformatics. Current Bioinformatics, 5(4), 296-308. https://www.ingentaconnect.com/content/ben/cbio/2010/00000005/00000004/art00006

Zimroz, R., & Król, R. (2009). Failure analysis of belt conveyor systems for condition monitoring purposes. Mining Science, 128(36), 255–270. http://www.miningscience.pwr.edu.pl/Failure-analysis-of-belt-conveyor-systems-for-condition-monitoring-purposes,59825,0,2.html

Zucchini, W. (2000). An Introduction to Model Selection. Journal of mathematical psychology, 44(1), 41-61. https://doi.org/10.1006/jmps.1999.1276

## Annexures

**Annexure 1:** Small sample size MRSM dataset

The MRSM dataset generation for the rubber covered conveyor belting problem is explained in detail in Pavolo and Chikobvu (2020a).

**The two-factor CCD experiment MRSM dataset**

| Run | T (min.) | $R_t$ (mm) | Ave. Hardness ($^0$shore A) | Ave. Adhesion(N/mm) |
|---|---|---|---|---|
| 1 | 16 | 7.2 | 60 | 10.60 |
| 2 | 30 | 7.2 | 63 | 13.34 |
| 3 | 16 | 22.8 | 53 | 6.20 |
| 4 | 30 | 22.8 | 61 | 12.10 |
| 5 | 23 | 15 | 58 | 11.80 |
| 6 | 23 | 15 | 58 | 12.10 |
| 7 | 13 | 15 | 44 | 6.50 |
| 8 | 33 | 15 | 63 | 13.30 |
| 9 | 23 | 4 | 63 | 13.30 |
| 10 | 23 | 26 | 56 | 3.50 |
| 11 | 23 | 15 | 58 | 12.20 |
| 12 | 23 | 15 | 57 | 12.30 |
| 13 | 23 | 15 | 58 | 12.10 |

**Annexure 2:** *The fifteen classical model selection criteria used for best model voting*

| CRITERION | FORMULAR | DESCRIPTION |
|---|---|---|
| AIC | $n * \ln\left(\frac{SS_{Res}}{n}\right) + 2k$ | Akaike (1973) |
| BIC | $n * \ln\left(\frac{SS_{Res}}{n}\right) + k * \ln(n)$ | Schwarz's Bayesian Criterion (1978) |
| HQ | $n * \ln\left(\frac{SS_{Res}}{n}\right) + 2p * \ln(\ln(n))$ | Hannan and Quinn (1979) |
| KIC | $n * \ln\left(\frac{SS_{Res}}{n}\right) + 3(k + 1)$ | Cavanaugh J.E. (1999) |
| TIC | $n * \ln\left(\frac{SS_{Res}}{n}\right) + 2(k + 1)$ | Takeuchi (1978) |
| SBC | $n * \ln\left(\frac{SS_{Res}}{n}\right) + \frac{2(k+2)n\sigma^2}{SS_{Res}} - \frac{2(n\sigma^2)^2}{SS_{Res}{}^2}$ | Sawa (1978) |
| AICc | $n * \ln\left(\frac{SS_{Res}}{n}\right) + 2k\,\frac{2k(k+1)}{(n-k-1)}$ | Corrected AIC, Sugiura (1978) |
| HQc | $n * 2\ln\left(\frac{SS_{Res}}{n}\right) + 2nk.\frac{\ln(\ln(n))}{(n-k-1)}$ | McQuarrie and Tsai (1998) |
| KICc | $n * \ln\left(\frac{SS_{Res}}{n}\right) + \frac{(k+1)(3n-k-2)}{(n-k-2)} + \frac{k}{(n-k)}$ | Bekara M (2004) |
| MKIC | $\frac{2(n-K-2)p\sigma^2}{\sigma^2 p} - 2p - 2n + 4$ | Cavanaugh J.E. (2004) |
| PRESS | $\sum_{i=1}^{n} e_i{}^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$ | Prediction Sum of Squares (Allen, 1971a) |
| $R^2$-prediction | $\left[1 - \frac{PRESS}{SS_{Total}}\right] * 100$ | Allen (1971b) |
| Adequate Precision | $\frac{[\max(\hat{Y}) - \min(\hat{Y})]}{\sqrt{(\tilde{V}_{\hat{y}})}} > 4$ | This is a signal-to-noise ratio |
| Cp - k | $\frac{SS_{Reg(p)}}{\hat{\sigma}^2} - n + k$ | Mallow's Cp (1973) |
| $APC_k$ | $\frac{(n+k)SS_{Reg(k)}}{n * (n-k)}$ | Amemiya's Prediction Criterion (1976) |

**Annexure 3:** The thirty-one all possible OLS adhesion response models

The table below shows the adhesion OLS response models respectively. For example, the first adhesion model in the table in its expanded form is:

$$Adhesion = 12.26 + 0T_i + 0R_i - 0.0039(T_i * R_i) + 0{T_i}^2 + 0{R_i}^2 \quad \text{(A1)}$$

Which simplifies to

$$Adhesion = 12.2600 - 0.0039(T_i * R_i) \quad \text{(A2)}$$

The second response model is

$$Adhesion = 7.95 + 0.3244T_i - 0.3127R_i + 0(T_i * R_i) + 0{T_i}^2 + 0{R_i}^2 \quad \text{(A3)}$$

Which simplifies to

$$Adhesion = 7.9500 + 0.3244T_i - 0.3127R_i \quad \text{(A4)}$$

The first two adhesion response models are represented in summary format as $[T_i * R_i]$ and $[T_i, R_i]$ respectively.

| MODEL | $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\beta}_{12}$ | $\hat{\beta}_{11}$ | $\hat{\beta}_{22}$ |
|---|---|---|---|---|---|---|
| $[T_i]$ | 3.26000 | 0.32440 | | | | |
| $[R_i]$ | 15.41000 | | -0.31270 | | | |
| $[T_i * R_i]$ | 12.26000 | | | -0.00389 | | |
| $[{T_i}^2]$ | 6.58000 | | | | 0.00668 | |
| $[{R_i}^2]$ | 13.64000 | | | | | -0.01114 |
| $[T_i, R_i]$ | 7.95000 | 0.32440 | -0.31270 | | | |
| $[T_i, T_i * R_i]$ | 3.26000 | 0.51000 | | -0.01235 | | |
| $[T_i, {T_i}^2]$ | -2.30000 | 0.83500 | | | -0.01110 | |
| $[T_i, {R_i}^2]$ | 6.18000 | 0.32440 | | | | -0.01114 |
| $[R_i, T_i * R_i]$ | 15.41000 | | -0.79100 | 0.02078 | | |
| $[R_i, {T_i}^2]$ | 11.67000 | | -0.31270 | | 0.00668 | |
| $[R_i, {R_i}^2]$ | 11.08000 | | 0.38000 | | | -0.02309 |
| $[T_i * R_i, {T_i}^2]$ | 8.96000 | | | -0.01189 | 0.01048 | |
| $[T_i * R_i, {R_i}^2]$ | 10.49700 | | | 0.02025 | | -0.02579 |
| $[{T_i}^2, {R_i}^2]$ | 9.91000 | | | | 0.00664 | -0.01109 |
| $[T_i, R_i, T_i * R_i]$ | 12.94000 | 0.10700 | -0.64600 | 0.01450 | | |
| $[T_i, R_i, {T_i}^2]$ | 2.41000 | 0.83500 | -0.31270 | | -0.01110 | |
| $[T_i, R_i, {R_i}^2]$ | 3.61000 | 0.32440 | 0.38000 | | | -0.02309 |
| $[T_i, T_i * R_i, {T_i}^2]$ | -2.28000 | 1.02000 | | -0.01235 | -0.01110 | |
| $[T_i, T_i * R_i, {R_i}^2]$ | 9.14000 | 0.09100 | | 0.01559 | | -0.02242 |
| $[T_i, {T_i}^2, {R_i}^2]$ | -0.25000 | 0.91900 | | | -0.01290 | -0.01122 |
| $[R_i, T_i * R_i, {T_i}^2]$ | 15.24000 | | -0.77100 | 0.01990 | 0.00031 | |
| $[R_i, T_i * R_i, {R_i}^2]$ | 11.08000 | | -0.09800 | 0.02078 | | -0.02309 |
| $[R_i, {T_i}^2, {R_i}^2]$ | 7.52000 | | 0.35800 | | 0.00661 | -0.02240 |
| $[T_i * R_i, {T_i}^2, {R_i}^2]$ | 10.39000 | | | 0.01890 | 0.00054 | -0.02485 |
| $[T_i, R_i, T_i * R_i, {T_i}^2]$ | 7.40000 | 0.61800 | -0.64600 | 0.01450 | -0.01110 | |
| $[T_i, R_i, T_i * R_i, {R_i}^2]$ | 8.61000 | 0.10700 | 0.04700 | 0.01450 | | -0.02309 |
| $[T_i, T_i * R_i, {T_i}^2, {R_i}^2]$ | -4.25000 | 1.02100 | 0.43000 | | -0.01510 | -0.02476 |
| $[T_i, R_t, {T_i}^2, {R_i}^2]$ | 1.95000 | 0.75900 | | 0.01676 | -0.01491 | -0.02336 |
| $[R_i, T_i * R_i, {T_i}^2, {R_i}^2]$ | 11.21000 | | -0.11300 | 0.02150 | -0.00026 | -0.02317 |
| $[T_i, R_i, T_i * R_i, {T_i}^2, {R_i}^2]$ | 0.74000 | 0.80400 | 0.09700 | 0.01450 | -0.01510 | -0.02476 |

**Annexure 4:** The Multiple Simultaneous Optimisation Ensemble of Pavolo and Chikobvu (2021)