Experimental design sequential generation and overall D-efficiency criterion for electron beam grafting of corn starch

E Koleva\textsuperscript{1,2}, L Kolev\textsuperscript{2}, M Braşoveanu\textsuperscript{3} and M R Nemţanu\textsuperscript{3}

\textsuperscript{1}Institute of Electronics, Bulgarian Academy of Sciences, 72 Tzarigradsko shossee blvd., Sofia 1784, Bulgaria
\textsuperscript{2}University of Chemical Technology and Metallurgy, 8 Kl. Ohridski blvd., Sofia 1756, Bulgaria
\textsuperscript{3}National Institute for Lasers, Plasma and Radiation Physics, Electron Accelerators Laboratory, 409 Atomistilor St., P.O. Box MG-36, 077125 Bucharest-Magurele, Romania

e-mail: eligeorg@abv.bg

Abstract. An experimental investigation of the electron beam irradiation induced grafting of acrylamide onto corn starch was performed. In this way the synthesis of water-soluble copolymers having flocculation abilities is realized. The acrylamide monomer provides a polar graft side chain which turns it into a hydrophilic copolymer. The robust engineering approach was implemented for the estimation of models for the means and the variances of the investigated quality parameters in production conditions, aiming the assurance of low toxicity of the synthesised copolymer, process economic efficiency, copolymer efficiency in the flocculation process and its good solubility in water. In order to improve the accuracy of the estimated models, performing new experiments was planned, based on the sequentially generated D-optimal designs. An approach for the generation of such designs for several quality parameters is presented.

1. Introduction

The current environmental considerations impose rigorous protection rules and sustainable progress that minimize the impact of wastes on the environment. Accordingly, there is a strong demand to develop economically viable and eco-friendly replacements of conventional synthetic flocculants, based upon the renewable organic materials that are low cost and degrade naturally when are released in the environment [1]. Grafting is the most effective way of regulating the properties of natural polysaccharides according to the needs and produces high efficient graft copolymers [2] that can be used as flocculating agents for treatment of different wastewaters.

Electron beam (EB) irradiation can be successfully used to develop a wide variety of ion exchers, polymer-ligand exchers, chelating copolymers, hydrogels, affinity graft copolymers and polymer electrolytes, having various applications in water treatment, chemical industry, biotechnology, biomedicine, etc. [3, 4].

During the experimental investigations in front of the researchers always stays the question what experimental design to choose. This choice is connected with the properties of the designs, the type of the estimated dependencies, the type of the process characteristics, the desired number of experimental
runs, etc. According to their properties, the designs can be orthogonal or non-orthogonal, symmetrical or non-symmetrical, with fixed number of experiments or sequentially generated, rotatable, balanced, etc. [5-7].

The design efficiencies are measures of design goodness. Common measures of the efficiency are based on the extended design matrix \( F \), the information matrix \( G = F^T F \) and the covariance matrix \( C = (F^T F)^{-1} \) [5], where \( F \) is a matrix, defined by the design matrix and the estimated regression model. According to the efficiency criteria, there are different types of optimal experimental designs: A-optimal, D-optimal, E-optimal, G-optimal, M-optimal etc. A design with extended matrix \( F \) is called D-optimal, if it maximizes the determinant of the information matrix \( G \), consequently it minimizes the volume of the confidence ellipsoid of the estimates of the regression coefficients (\( \hat{\theta}_j \)). The property A-optimality is connected with minimizing the trace \( tr(C) \), representing the sum of the variances of the estimates of the regression coefficients \( \sigma^2(\hat{\theta}_j) \). The G-optimality is connected with minimizing the maximal variance of the predicted by the estimated regression models quality parameters \( \sigma^2(y_u) \).

There are many algorithms for generating information-efficient or optimal designs.

The advantages of the sequential generation of D-optimal designs are connected with the flexible augmentation of already performed experimental designs, with the generation of designs for not full second order polynomials, for quantitative and qualitative process parameters, when the qualitative process parameters have different number of levels, when there is irregular shape of the regions of interest of the factors [5]. The sequentially generated D-optimal designs can be performed in two ways – when initial design is already conducted or when there are no initial experiments.

In this paper a number of 20 experiments for electron beam grafting of acrylamide onto starch in order to obtain water-soluble copolymers having flocculation abilities were considered. In order to improve the accuracy of the estimated models (regression coefficients) for the means and variances of the quality parameters of copolymers (residual monomer concentration, monomer conversion coefficient, intrinsic viscosity and Huggins’ constant) additional sequentially generated experimental runs are proposed. In this way the performed experiment will be augmented by new experimental points by the implementation of D-efficiency criterion, the robust engineering approach and taking into account several quality parameters.

2. Methodologies

2.1. Robust engineering design approach

Robust (not sensitive to noises and errors) engineering approach can be implemented to analyse experiments where the variance is non-homogeneous over the factor (process parameters’) space [5, 8] and the noise factors cannot be identified or experimentally studied. The observations in this case are called heteroscedastic (variance varies with the factor levels). Minimization of the variances of the investigated quality parameters will practically lead to industrial production with better repeatability and quality.

The estimated mean value \( \hat{y}_u \) and variance \( s^2_u \) can be considered as two responses at the design points, and ordinary least squares method can be used to fit the regression models of the mean value and the variance for the quality characteristic [5]:

\[
\hat{y}(\hat{x}) = \sum_{i=1}^{k_x} \hat{\theta}_{yi} f_{yi}(\hat{x})
\]

\[
\ln(\hat{s}^2(\hat{x})) = \sum_{i=1}^{k_x} \hat{\theta}_{si} f_{si}(\hat{x})
\]
where \( \hat{\theta}_{yi} \) and \( \hat{\theta}_{ci} \) are estimates of the regression coefficients, and \( f_{yi} \) and \( f_{ci} \) are known functions of the process parameters \( x_i \).

The models are estimated for coded in the region \([-1; 1]\) values of the process parameters, using the following equation:

\[
x_i = (2z_i - z_{i,\text{max}} - z_{i,\text{min}})/(z_{i,\text{max}} - z_{i,\text{min}})
\]

where \( x_i \) and \( z_i \) are the coded and the natural values of the process parameter.

2.2. \( D \)-efficiency criterion

A design with probability measure \( \xi \), which satisfies the condition [5]:

\[
|M(\xi)| = \max_{\xi} |M(\xi)|
\]

is called continuous D-optimal design. The continuous designs have properties that do not depend on the number of observations [5]. The normalized information matrix in (4) is given by:

\[
M = \frac{1}{N} F^T F
\]

with elements:

\[
m_{ij} = \frac{1}{N} \sum_{l=1}^{N} f_{il} f_{uj}
\]

where \( N \) – the number of observations. If the \( N \) observations are allocated at \( h \) support points and \( r_l \) observations are allocated at each of them, then:

\[
m_{ij} = \sum_{l=1}^{h} \frac{r_l}{N} f_{il} f_{ji} = \sum_{l=1}^{h} \xi_l f_{il} f_{ji}
\]

where \( \xi_l \) is the proportion of the observations at the \( l \)-th design support point. The probability measure function \( \xi(\vec{x}) \) has non-zero values only in the design support points and the values are integer for the continuous D-optimal designs.

The D-efficiency of a design can be defined as follows:

\[
D_{\text{eff}} = \left( \frac{|M(\xi)|}{|M(\xi_0)|} \right)^{1/k}
\]

where \( k \) is the number of the coefficients in the regression model. The determinant of the normalized information matrix of any investigated design is compared to that of the continuous D-optimal design.

The \( D_{\text{eff}} \) is zero for singular information matrices \( G = F^T F \) and are equal to 1 for continuous D-optimal designs.

3. Experimental conditions

Experiments for grafting of starch with acrylamide by using electron beam irradiation were performed in order to synthesize water-soluble copolymers having flocculation abilities. The synthesis of graft copolymers was performed by two steps: (1) preparation of solutions containing starch and monomer and (2) irradiation of prepared solutions by electron beam.

1. Starch aqueous solutions were prepared by dissolving corn starch in distilled water. Then, acrylamide was added to these solutions with further stirring, resulting in various acrylamide/starch (AMD/St weight ratios) homogenous aqueous solutions.

2. Homogenous aqueous solutions resulted in step 1 were exposed to electron beam irradiation. The irradiations were carried out at ambient temperature and pressure by using linear electron accelerators of mean energy of 5.5 MeV with different irradiation doses and dose rates.

The synthesized graft copolymers were characterized by the following performance quality parameters: \( y_1 \) [%] - residual monomer concentration, \( y_2 \) [%] - monomer conversion coefficient, \( y_3 \)
[dL/g] - intrinsic viscosity and $y_4$ - Huggins’ constant. The variation regions $[z_{\text{min}}-z_{\text{max}}]$ of the process parameters are presented in table 1.

**Table 1. Process parameters for corn starch modified by electron beam irradiation of 5.5 MeV.**

| Parameter                        | $z_{i,\text{min}}$ | $z_{i,\text{max}}$ |
|----------------------------------|--------------------|--------------------|
| EB irradiation dose [kGy]        | 0.64               | 1.44               |
| EB irradiation dose rate         | 0.45               | 1.40               |
| AMD/St weight ratio              | 5.00               | 10.02              |

The concentration of St for these experiments varied from 2.00% to 6.15% and the concentration of AMD varied from 10.00% to 33.67% [8].

The conducted experimental design with 20 experimental process parameter sets for coded values in the region $[-1,1]$ by using equation (3) is presented in table 2, and the obtained regression models are presented in table 3 and table 4, together with the values of the corresponding multiple correlation coefficients $R^2$. These coefficients are tested for significance and their values are measures of the accuracy of the estimated models. The closer to 1 the value of $R^2$ is, the better the model describes the variations of the quality characteristics as a function of the process parameters.

**Table 2. Experimental design.**

| $N_0$ | $x_1$   | $x_2$   | $x_3$   |
|-------|---------|---------|---------|
| 1     | -0.550  | -0.68421| -1.00000|
| 2     | 0.275   | -0.32632| -0.99602|
| 3     | 0.175   | 1.00000 | -1.00000|
| 4     | -0.600  | -0.72632| 0.97211 |
| 5     | -0.150  | -0.38947| 1.00000 |
| 6     | 0.275   | -0.32632| 1.00000 |
| 7     | -0.175  | 0.43158 | -0.99602|
| 8     | 1.000   | 0.07368 | -0.99602|
| 9     | 0.325   | -0.30526| -0.99602|
| 10    | -1.000  | -0.15789| -0.99602|
| 11    | -1.000  | -0.60000| -0.99602|
| 12    | -0.125  | -0.55789| -0.99602|
| 13    | -0.900  | -0.51579| -0.99602|
| 14    | -0.100  | -1.00000| -0.99602|
| 15    | -0.175  | 0.43158 | 0.98406 |
| 16    | -1.000  | -0.15789| 0.99203 |
| 17    | -1.000  | -0.60000| 0.98406 |
| 18    | -0.125  | -0.55789| 0.98406 |
| 19    | -0.900  | -0.51579| 0.98805 |
| 20    | -0.100  | -1.00000| 0.98805 |

The estimated models in table 3 can be used for optimisation at formulation of specific product and quality requirements. Taking into account the models in table 4, the optimal solutions and the corresponding process parameter choice will guarantee robustness toward production errors and higher product and process quality. The set production requirements were the following:

- residual monomer concentration: $\hat{y}_1(\mathbf{x}) < 5\% \rightarrow$ assurance of low toxicity,
- monomer conversion coefficient: $\hat{y}_2(\mathbf{x}) > 90\% \rightarrow$ economic efficiency,
- intrinsic viscosity: $\hat{y}_3(\mathbf{x}) > 6 \text{ dL/g} \rightarrow$ copolymer efficiency in flocculation process,
Huggins’ constant: $0.3 \leq \tilde{y}_4(\tilde{x}) \leq 1$ (or $-1.20397 \leq \ln(\tilde{y}_4) \leq 0$) → good solubility in water.

**Table 3.** Models for the means of the product quality characteristics.

| Models | $R^2$ | Sign. |
|--------|-------|-------|
| $\hat{y}_1(\tilde{x})$ | 1.8416958 - 1.4711672x_1 + 0.39465799x_2 + 0.76523861x_1^2 - 1.4618933x_1x_2 | 0.82 | Model 1 |
| $\hat{y}_2(\tilde{x})$ | 89.612983 + 6.1859389x_1 + 2.921934x_1 + 5.2145372x_2 - 2.1297402x_1x_3 | 0.75 | Model 2 |
| $\hat{y}_3(\tilde{x})$ | 4.2786392 - 0.95051438x_2 + 2.2118238x_3 + 1.2412826x_2^2 - 7.0991376x_1^2x_2 + 0.7214555x_3 - 1.2935142x_1x_3 + 12.089739x_1x_2^2 | 0.79 | Model 3 |
| $\ln(\tilde{y}_4)$ | 0.86268086 + 1.5188592x_2 + 1.6109201x_3 + 0.87526587x_1x_3 - 1.2935142x_1x_3 | 0.62 | Model 4 |

**Table 4.** Models for the variances of the product quality characteristics.

| Models | $R^2$ | Sign. |
|--------|-------|-------|
| $\ln(\xi_1^2(\tilde{x}))$ | -3.0237265 - 0.60469566x_2 - 1.0612263x_3 - 4.3482641x_1^2 - 4.2402799x_2^2 + 5.604919x_1x_2 - 2.2657548x_2x_3 | 0.72 | Model 5 |
| $\ln(\xi_2^2(\tilde{x}))$ | 0.09316988 - 1.0210602x_2 - 1.7785371x_3 - 4.6387008x_1^2 - 4.9999832x_2^2 + 6.047145x_1x_2 - 2.242544x_2x_3 | 0.76 | Model 6 |
| $\ln(\xi_3^2(\tilde{x}))$ | -2.5452688 - 0.6782522x_2 - 2.7034767x_1^2 + 2.1048188x_2^2 + 3.2229606x_1x_2 + 0.88934159x_2^2x_3 | 0.50 | Model 7 |

The result of the graphical optimization is shown in figure 1 [8]. The figure presents the contour plot with the contour lines of the constraints for $\tilde{y}_1(\tilde{x})$ - the residual monomer concentration, $\tilde{y}_2(\tilde{x})$ - monomer conversion coefficient, $\tilde{y}_3(\tilde{x})$ - intrinsic viscosity and $\tilde{y}_4(\tilde{x})$ - Huggins’ constant, as a function of the EB irradiation dose ($z_1$) and EB irradiation dose rate ($z_2$) for AMD/St weight ratio $z_3 = 7.51$. The colored area corresponds to possible process parameter choices, where all the constraints are fulfilled. The final choice of regime parameters can be done after minimizing the variances of these quality parameters by the models in table 4.

![Figure 1. Graphical optimization. Contour plot of the constraints for $\tilde{y}_1(\tilde{x})$ - the residual monomer concentration, $\tilde{y}_2(\tilde{x})$ - monomer conversion coefficient, $\tilde{y}_3(\tilde{x})$ - intrinsic viscosity and $\tilde{y}_4(\tilde{x})$ - Huggins’ constant, as a function of the EB irradiation dose ($z_1$) and EB irradiation dose rate ($z_2$) for AMD/St weight ratio $z_3 = 7.51$.](image)

The accuracy of the predicted quality characteristics can be defined by their variances. As an example, the variance of the predicted variable $\tilde{y}_2(\tilde{x})$, estimated on the base of the initial experimental
design and the model in table 3, is presented in figure 2 as a function of the process parameters: a) \(x_1\) and \(x_2\) at \(x_3 = 0\), b) \(x_1\) and \(x_3\) at \(x_2 = 0\), c) \(x_2\) and \(x_3\) at \(x_1 = 0\). The maximum value of \(\sigma^2(\hat{y}_2) = 1.4489\) was obtained for the values of the process parameters: EB irradiation dose \(z_1 = 1.44\) kGy and EB irradiation dose rate \(z_2 = 1.40\) kGy/min and AMD/St weight ratio \(z_3 = 10.02\).

![Figure 2](image)

**Figure 2.** Variance of the predicted variable \(\hat{y}_2(x)\), estimated on the base of the initial experimental design, as a function of the process parameters: a) \(x_1\) and \(x_2\) at \(x_3 = 0\), b) \(x_1\) and \(x_3\) at \(x_2 = 0\), c) \(x_2\) and \(x_3\) at \(x_1 = 0\).

4. Results and discussion

The sequential generation of D-optimal designs, considered here, involves the use of the initially performed experimental design with 20 experiments (table 2) and the use of new candidate experimental points. They represent all possible combinations of the chosen levels of the process parameters (factors - \(x_1\), \(x_2\) and \(x_3\)) without repeated experimental runs. For example, if for \(x_1\) are chosen levels with a step 0.25 within the region \([-1, 1]\) the following values: -1, -0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75, 1 are possible.

The sequentially generated in table 5 additional experiments (5, 10 and 15 experiments) consider the maximization of the criterion:

\[
\hat{x}_{N+1} = \arg \max_x f_{N+1}^x (F_N^x F_N)^{-1} f_{N+1}
\]

(8)

by scanning all candidate experimental points, with different levels of the process parameters (\(x_1\), \(x_2\), and \(x_3\)) within the coded experimental regions \([-1, 1]\). The different levels of the values of the factors
are defined by different variation steps 0.25, 0.5 and 1. The results for the individual D-efficiencies are obtained by investigation of each of the models independently from the others. That is why the generated additional experimental points are different in each single case.

**Table 5. Individual D-efficiencies (additional experiments/ coded step size).**

| Model  | $D_{\text{eff}}$ (initial) | $D_{\text{eff}}$ (5 / 0.5) | $D_{\text{eff}}$ (10 / 1) | $D_{\text{eff}}$ (10 / 0.5) | $D_{\text{eff}}$ (10 / 0.25) | $D_{\text{eff}}$ (15 / 0.5) | $D_{\text{eff}}$ (overall) |
|--------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| Model 1| 0.2905                      | 0.4967                      | 0.5860                      | 0.5372                      | 0.5134                      | 0.5658                      | 0.5757                      |
| Model 2| 0.6252                      | 0.8156                      | 0.9062                      | 0.9030                      | 0.8167                      | 0.8950                      | 0.8154                      |
| Model 3| 0.3151                      | 0.6226                      | 0.8497                      | 0.6736                      | 0.6157                      | 0.6887                      | 0.7193                      |
| Model 4| 0.5833                      | 0.8736                      | 0.9648                      | 0.9732                      | 0.9344                      | 0.9727                      | 0.8017                      |
| Model 5| 0.3734                      | 0.6439                      | 0.8137                      | 0.6948                      | 0.6461                      | 0.7095                      | 0.6872                      |
| Model 6| 0.3734                      | 0.6439                      | 0.8137                      | 0.6948                      | 0.6461                      | 0.7095                      | 0.6872                      |
| Model 7| 0.4902                      | 0.7689                      | 0.9462                      | 0.8082                      | 0.7328                      | 0.7976                      | 0.9395                      |

* 9 experiments

The maximization of the criterion (8) means that by adding the new N+1 experimental point to the already existing experimental design with N points will be in the place of the factor space, where the variance of the estimated regression coefficients $(F_0^T F_N)^{-1}$ is maximum.

In order to consider all estimated models in table 3 and table 4 an overall sequential generation criterion is necessary. The proposed here criterion is the simultaneous maximization of (8) for all models, when generating each new experimental point.

**Table 6. Experimental design – augmentation.**

| N  | $x_1$ | $x_2$ | $x_3$ | $z_1$ [kGy] | $z_2$ [kGy/min] | $z_3$ |
|----|-------|-------|-------|-------------|-----------------|-------|
| 21 | -1.000| 1.000 | -1.000| 0.6400      | 1.4000          | 5.0000 |
| 22 | -1.000| 1.000 | 1.000 | 0.6400      | 1.4000          | 10.0200|
| 23 | 1.000 | -1.000| -1.000| 1.4400      | 0.4500          | 5.0000 |
| 24 | -1.000| -1.000| -1.000| 0.6400      | 0.4500          | 5.0000 |
| 25 | -1.000| 1.000 | 0.000 | 0.6400      | 1.4000          | 7.5100 |
| 26 | -1.000| -1.000| 0.000 | 0.6400      | 0.4500          | 7.5100 |
| 27 | -1.000| -1.000| 1.000 | 0.6400      | 0.4500          | 10.0200|
| 28 | 1.000 | 0.000 | -1.000| 1.4400      | 0.9250          | 5.0000 |
| 29 | 1.000 | 0.000 | 1.000 | 1.4400      | 0.9250          | 10.0200|
| 30 | -1.000| 0.000 | -1.000| 0.6400      | 0.9250          | 5.0000 |

In figure 3 is presented the variance of the predicted variable $\hat{y}_2(\bar{x})$, estimated on the base of the suggestion for the augmentation of the initial experimental design and the model in table 3, as a function of the process parameters: a) $x_1$ and $x_2$ at $x_3 = 0$, b) $x_1$ and $x_3$ at $x_2 = 0$, c) $x_2$ and $x_3$ at $x_1 = 0$. The maximum value of $\sigma^2(\hat{y}_2) = 0.7496$ is obtained for the values of the process parameters: EB irradiation dose $z_1 = 1.44$ kGy and EB irradiation dose rate $z_2 = 1.40$ kGy/min and AMD/St weight ratio $z_3 = 10.02$. It is seen that the variance of the predicted monomer conversion coefficient $\hat{y}_2(\bar{x})$ was reduced around two times.
Figure 3. Variance of the predicted variable \( \hat{y}_2(\hat{x}) \), estimated on the base of the augmented experimental design, as a function of the process parameters: a) \( x_1 \) and \( x_2 \) at \( x_3 = 0 \), b) \( x_1 \) and \( x_3 \) at \( x_2 = 0 \), c) \( x_2 \) and \( x_3 \) at \( x_1 = 0 \).

5. Conclusions
In this paper 20 experiments for the grafting of starch with acrylamide by using electron beam irradiation in order to synthesize water-soluble copolymers having flocculation abilities were considered. In order to improve the accuracy of the estimated models (regression coefficients) for the means and variances of the copolymer quality parameters (residual monomer concentration, monomer conversion coefficient, intrinsic viscosity and Huggins’ constant) additional sequentially generated experimental runs were proposed.

The implemented overall D-optimal design sequential generation criterion represents the maximization of the determinants of the normalized information matrices \( |M_r(\hat{x})| \) when adding each next experimental point. The generated points consider the maximization of the criterion (8) by scanning among the candidate experimental points for all estimated models simultaneously.

The initial experiment was conducted in the space of the process parameters – the electron beam irradiation dose, the dose rate and the AMD/St weight ratio – were not equally distributed, due to the prior experience for not obtaining good quality parameter results. At choosing the candidate points for new experiments, this information can be taken into consideration by excluding such points or making the experiment for improving the prediction and accuracy characteristics of the estimated models.
Acknowledgements
This research was conducted under a bilateral joint project between the Institute of Electronics at Bulgarian Academy of Sciences and the National Institute for Lasers, Plasma and Radiation Physics at the Romanian Academy.

This work was also partially supported by a project NUCLEU from the Romanian Ministry of Research and Innovation, by a grant of the Romanian National Authority for Scientific Research, CNDI–UEFISCDI (project number 64/2012) and by the Bulgarian National Science Fund under contract DN17/9.

References
[1] Sharma B 2006 J. Polym. Environ. 14 195-202
[2] Lee C 2014 Process Safety Environ. Protect. 92 489-508
[3] Lapin S 2015 UV+EB Technol. 1 44-49
[4] Nasef M 2012 Prog. Polym. Sci. 37 1597-1656
[5] Vuchkov I and Boyadjieva L 2001 Quality improvement with design of experiments (Dordrecht: Kluwer Academic Publishers)
[6] Toubia O and Hauser J 2007 Market. Sci. 26(6) 851–858
[7] Ranade S and Thiagarajan P 2016 Int. J. Pharm. Technol. 8(3) 16277-16287
[8] Koleva E, Koleva L, Tzotchev V, Nemtanu M, Brashoveanu M and Vutova K 2016 Sci. Eng. & Education 1 89-96