

## Intisari

Response surface methodology (RSM) menjadi sebuah standar utama di dalam proses optimisasi sebuah proses. Terdapat tiga tahapan dalam RSM, yaitu desain eksperimen (DoE), pemodelan, dan optimisasi, ketiganya tersusun untuk menemukan area optimisasi dan titik optimalnya secara akurat. RSM mensyaratkan kendali dan randomisasi eksperimen secara penuh untuk menangkap hubungan antara faktor dan respon. Bagaimanapun, sebagai bagian dari RSM, penerapan DoE pada proses manufaktur yang kontinu memunculkan kendala yang tidak dapat dihindari, yaitu menghentikan proses yang sedang berjalan untuk melaksanakan eksperimen DoE, hal ini tentu saja akan menimbulkan kerugian waktu operasional proses. Pada sisi lain, proses manufaktur yang sudah menerapkan otomatisasi sistem pencatatan akan menyimpan data-data observasional berisi perubahan setting parameter proses termasuk karakteristik produknya, dan data-data ini berpotensi menyimpan banyak informasi untuk menyelidiki hubungan kausalitas antara faktor dan respon proses. Sehingga, jika dibandingkan dengan menghentikan proses yang sedang berjalan, mengekstrak informasi dari data observasional menjadi alternatif yang dapat dipertimbangkan.

Untuk menciptakan kondisi seperti eksperimen DoE dari data observasional, penelitian ini mengajukan pendekatan baru melalui penerapan *multiobjective optimal instance selection* untuk memenuhi ortogonalitas dan *space-filling* seakan-akan sebuah DoE. Algoritma genetika bekerja secara iteratif memilih subset observasi yang memenuhi kriteria spesifik berdasarkan sebuah DoE yang ideal, termasuk mengakomodasi adanya kovariat. Terlebih, pada kondisi ketika subset belum memenuhi kriteria DoE, algoritma serupa dapat membangkitkan rekomendasi eksperimen riil untuk melengkapi subset tersebut sehingga keberterimaan DoE-nya lebih tinggi. Subset terpilih bekerja sebagai pengganti eksperimen DoE karena telah memenuhi sifat DoE. Berikutnya, pemodelan RSM akan berlanjut melalui pemodelan model linear orde kedua untuk menangkap hubungan antara faktor dan respon serta memastikan adanya titik optimum global. Pada akhirnya, tahapan optimisasi dapat berjalan mengacu pada model RSM tersebut dengan menerapkan algoritma optimisasi standar. Implementasi dari kerangka RSM berbasis data observasional ini diterapkan pada beberapa data riil, berupa big-data dari proses flotasi dan proses CnC, dilengkapi dengan simulasi yang mengacu pada beberapa fungsi uji optimisasi.

Hasil penerapan menunjukkan bahwa subset yang terpilih dari data observasional telah memenuhi kemiripan dengan DoE, terutama ketika melalui penambahan rekomendasi titik eksperimen riil. Studi simulasi juga menunjukkan bahwa metode ini berhasil menemukan titik optimal pada beberapa jenis fungsi uji optimisasi, meskipun terdapat selisih antara hasil optimal teoritis dengan RSM berbasis observasional namun tidak berbeda secara signifikan. Secara praktis, penggunaan data observasi dalam optimisasi berbasis RSM ini akan mengurangi *downtime* dari proses yang sedang berjalan.

**Kata kunci:** response surface methodology, data observasional, design of experiment, instance subset selection, multi-objective genetic algorithm.

## Abstract

Response surface methodology (RSM) becomes a gold standard method in process optimization purposes. Three stages of RSM i.e. design of experiment (DoE), modeling, and optimization, ensure that the desired optimization region and optimal solution are accurately found. The RSM requires full control and randomizes the DoE to capture the pure relationship between experiment factors and response. However, the application of the DoE within RSM for continuous manufacturing raises an unavoidable constraint given that stopping a running process to carry out actual experiments raises operating time losses. Meanwhile, a settled manufacturing system provides observational data records regarding changes in the process parameters and product characteristics and thus contains useful information for investigative causation purposes. As such, rather than disrupting ongoing process, retrieving important causality information from big observed data provides a comparable contribution to the designed experiment.

To create a DoE-like condition from observational data, this research proposes a new approach using a multiobjective optimal instance selection of observational data subsets to satisfy orthogonality and space-filling DoE principles. A modified genetic algorithm works by iteratively selecting subset observations that fulfill specific criteria based on ideal DoE properties while considering the adoption of observational data characteristics, including the existence of covariates. Moreover, for certain unsatisfied criteria, a similar algorithm generates a new actual experiment to complete the existing dataset and increase the acceptance of the selected subset. The founded subset works as a substitute for a full DoE experiment and satisfies an orthogonal and balanced design. Thus, the RSM modeling stages partake in the factors-response relationship based on the subset using a standard second-order linear model to accommodate a single global optimal point. Eventually, the optimization stages employ standard available algorithms to search for the optimal solution based on the fitted RSM model. The implementation of the proposed frameworks involves some real-data case studies from a large observational dataset from a flotation and CnC process. Completing the analysis, a simulation study based on optimization function tests is performed as well.

The results show a successful implementation with acceptable performance in selecting the optimal subset based on the DoE model and criteria; additionally, some new actual experiment points have been generated and have increased the subset acceptance. The simulation study shows that the proposed framework works well in some types of multimodal optimization tests and gives an optimal solution with insignificant difference with theoretical calculation. Practically, the proposed method becomes an alternative in optimizing continuous process to reduce production downtime.

**Keywords:** response surface methodology, observational data, design of experiment, instance subset selection, multi-objective genetic algorithm.