

introducción

La optimización de recursos materiales y económicos juega un papel fundamental en la experimentación, estando estrechamente ligada a las condiciones bajo las que deben realizarse los experimentos. En muchas disciplinas existe especial interés en analizar cómo varía una o varias características de un producto final dependiendo de su composición. El problema de diseño debe dirigirse a identificar las proporciones de la mezcla que describen mejor cierta característica de interés optimizando los recursos disponibles.

Los elementos de composición de este tipo de problemas son las variables controlables por el experimentador, en este caso, las proporciones de los ingredientes que intervienen en la mixtura, y la región de diseño donde estas cantidades sean factibles. El problema de diseño de mezclas se caracteriza por estar definido sobre una región restringida, el simplex, como resultado de la relación de dependencia existente entre variables, esto es, si una de las proporciones varía, al menos una de las restantes debe modificarse necesariamente por esta situación.

Para resolver este tipo de problemas, es necesario definir un modelo adecuado que describa la relación composición-respuesta. En la literatura se han desarrollado diferentes modelos para describir este tipo de comportamientos, siendo una reparametrización de los polinomios ordinarios, los de mayor acogida debido a su flexibilidad en multitud de situaciones prácticas. Éstos son los conocidos polinomios canónicos de Scheffé (1958)^[1].

La mayor parte de los trabajos dedicados al diseño de mezclas han recibido un enfoque clásico-determinista del problema. En este sentido, existe una serie de diseños estándar que por sistema han sido utilizados en la literatura. Algunos de los más populares son los simplex-lattice, simplex-centroid y

axiales (ver [1], [2] y [3] respectivamente).

Aplicar la teoría de diseño óptimo para experimentos con mezclas (DOEM) es un planteamiento relativamente nuevo y poco estudiado. Una guía actualizada de los trabajos existentes puede encontrarse en la monografía *Optimal Mixture Experiment* (Sinha et al. 2014, [4]).

Box y Draper en el trabajo *A basis for the selection of a response surface design* (1959) [5] advierten que asumir un modelo inexacto puede poner en riesgo las conclusiones del estudio. Las discrepancias existentes entre la característica observada y la estimada radican principalmente en dos fuentes de error: error de varianza, producido por el muestreo, y error de sesgo, debido a la no adecuación del modelo. En este mismo trabajo se muestra que cualquier desviación leve del modelo asumido disminuye la eficiencia de un diseño calculado cuando únicamente se analiza un funcional de la matriz de varianzas-covarianzas. Bajo este marco, el objetivo de este trabajo consiste en la búsqueda de estrategias de construcción de los diseños óptimos tomando como medida de calidad un funcional apropiado de la matriz del error cuadrático medio de los estimadores sobre un vecindario del modelo considerado.

Un algoritmo genético para generar diseños D-óptimos-robustos

Nuestro interés reside en calcular diseños D-óptimos-robustos para mezclas de tres ingredientes. El criterio de D-optimización es uno de los más conocidos en DOE y su finalidad es minimizar el volumen del elipsoide de confianza de los parámetros del modelo. En este trabajo se consideró un modelo de Scheffé cuadrático con término constante para aproximar la respuesta. La estrategia utilizada para la construcción de los diseños fue una estrategia minimax que

consiste en calcular el diseño que minimiza la variabilidad (el mejor de los diseños) cuando la diferencia entre el modelo considerado y la verdadera respuesta es máxima (en el peor de los casos).

En la mayoría de los casos no es posible encontrar una solución analítica para este tipo de problemas, siendo necesarios métodos numéricos para su resolución. Incluso utilizando este tipo de técnicas, los problemas de mezclas requieren un alto coste computacional en su resolución, y la complejidad que presentan las formas funcionales llevan a abortar el proceso de optimización en la mayoría de los casos. Una alternativa viable y de gran acogida en la actualidad, que proporciona buenas soluciones cuando las técnicas de optimización habituales basadas en el cálculo diferencial fallan, son los algoritmos genéticos (AGs). El proceso de optimización se basa en los principios de ensayo y error y probabilidad y, a pesar de ser un método heurístico (la convergencia no está garantizada), la práctica demuestra su buen funcionamiento y capacidad de búsqueda en tiempos razonables. Esta clase de algoritmos fue establecida por Holland en 1975 [6] y popularizada por Goldberg en otros campos del conocimiento (1989, [7]).

Es habitual encontrar terminología específica para denominar conceptos del DOE por sus "homólogos" en Genética. Al conjunto de diseños iniciales se denota por P y se conoce como población. Los cromosomas corresponden a las soluciones potenciales del problema, es decir, a los diseños y a sus puntos se les llama genes. El criterio de optimización elegido se conoce como función fitness.

El algoritmo comienza generando una población de M diseños aleatorios. Llamaremos generación a la renovación completa de todos los diseños de la población que se llevará a cabo mediante

la actuación de los operadores genéticos. Éstos se clasifican en tres grandes grupos: operadores de selección, cruce y mutación. Se elegirán adecuadamente dos diseños padre y, a partir de ellos, se creará un nuevo diseño hijo que pasará a la siguiente generación. Los operadores de selección actuarán siempre hasta haber completado una generación; mientras que cruce y mutación actuarán o no dependiendo de un test de probabilidad. El proceso de elección de los padres se realiza aleatoriamente considerando que los diseños de la población con más calidad, tienen mayor probabilidad de salir elegidos. La misión del operador de cruce es generar nuevos diseños similares a sus padres, mientras que la mutación se encarga de la "biodiversidad" de las soluciones con el objetivo de evitar caer en óptimos locales. Una regla de parada que resulta adecuada está ligada al hecho de no encontrar mejoras sustanciales en el fitness del mejor diseño de cada generación. En este trabajo se consideró que el diseño D-óptimo-robusto había sido alcanzado cuando el mejor de los diseños de la población permaneciera invariante durante 1000 iteraciones consecutivas. Denotando por ν a la importancia relativa que el experimentador da al sesgo frente a la varianza, la figura 1 muestra los diseños D-óptimos-robustos calculados para diferentes valores de ν .

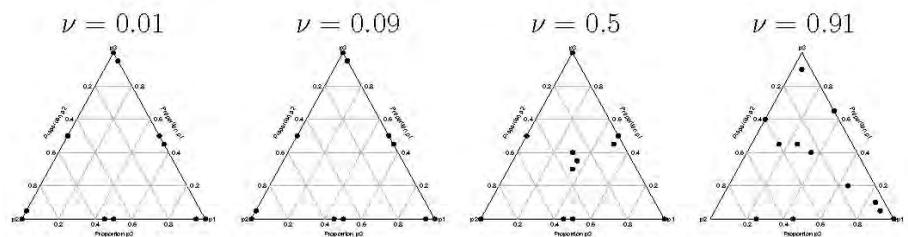


Figura 1: Diseños D-óptimos-robustos obtenidos con AG. El simplex para tres ingredientes se representa por un triángulo donde los vértices corresponden a los componentes puros, los lados a las mezclas binarias y el interior a las mixturas en las que intervienen todos los componentes. La cercanía/lejanía de un punto de diseño a cada vértice representa la proporción presencia/ausencia de ese componente en la mixtura.

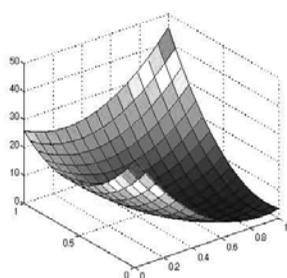


Figura 2: Superficie-respuesta obtenida del ECM

- [1] H. Scheffé, "Experiments with mixtures", J. Roy. Statist. Soc., vol. 20, no. 2, series B, pp.344-360, 1958.
- [2] H. Scheffé, "The simplex-centroid design for experiments with mixtures", J. Roy. Statist. Soc., vol. 25, no. 2, series B, pp.235-263, 1963.
- [3] J. A. Cornell, "Experiments with Mixtures", Wiley, New York, 2002.
- [4] B. K. Sinha, N.K. Mandal, Manisha Pal, y P. Das, "Optimal mixture experiments", Springer, Series: Lecture Notes in Statistics, vol. 1028, 2014.
- [5] G.E.P. Box, and N.R. Draper, "A basis for the Selection of a Response Surface Design", Journal of the American Statistical Association, vol. 54, pp.622-654, 1959.
- [6] J. H. Holland, "Adaptation in Natural and Artificial Systems", MIT Press: Cambridge, 1975.
- [7] D. E. Goldberg, "Genetic Algorithm in Search, Optimization, and Machine Learning", Addison-Wesley: New York, 1989.
- [8] J. J. Hernández, J. Sanz-Argent, J. Benajes y S.Molina, "Selection of a diesel fuel surrogate for the prediction of auto-ignition under HCCI engine conditions", Fuel, vol. 87, pp.655-665, 2008.
- [9] M. Barrero, "El hormigón del futuro", maet, vol. 1, pp.117, 2013.

Aplicación en problemas reales

La idea de emplear la metodología anterior para resolver un caso real surge del trabajo de J. J. Hernández et al. (2008,^[8]). El interés de esta investigación es encontrar la composición de sucedáneo que mejor simule al diésel en el autoencendido bajo las condiciones de un motor HCCI (*Homogeneous Charge Compression Ignition*). El desarrollo de estos motores de combustión interna alternativos es objeto de estudio en numerosos centros de investigación y fabricantes de automóviles. Su rendimiento a carga media es significativamente mayor que el de los motores tradicionales, y sus bajas emisiones de NOx cumplen con las normativas europeas sobre la emisión de gases contaminantes.

El principal reto que debe superarse para su puesta en el mercado es lograr el control sobre el inicio del autoencendido. Este proceso es controlado en este tipo de motores fundamentalmente por la cinética química del combustible que es extremadamente sensible a la composición de la carga. Los combustibles de los motores comerciales consisten en mezclas de cientos de hidrocarburos, y en los estudios cinéticos sólo un número pequeño de representantes de cada familia es utilizado para simular las propiedades de los combustibles originales. En J. J. Hernández et al. ^[8] consideraron mezclas de n-heptano, tolueno y ciclohexano para simular el diésel en representación de los grupos: parafinas, aromáticos y naftenos respectivamente.

La robustez en este problema cobra especialmente sentido debido a que cambios en las condiciones de operación de motor (presión, temperatura, revoluciones, etc) perturban fuertemente el fenómeno estudiado. La respuesta a modelizar fue el error cuadrático medio ECM de ángulos de autoencendido que mide las discrepancias

entre el sucedáneo y el diésel original en relación a éste. La metodología empleada fue la descrita en la sección 2 y los diseños propuestos fueron los mostrados en la figura 1.

La segunda etapa del problema consistirá, por tanto, en determinar qué composición da un valor mínimo de la respuesta. Para ilustrar, considerando los diseños obtenidos para $v=0.01, 0.09$ y tras el desarrollo de la fase experimental, se obtuvo la superficie-respuesta de mejor ajuste por mínimos-cuadrados que se muestra en figura 2. Utilizando una rutina de minimización sobre el conjunto de soluciones factible se obtuvo que la composición óptima, la que mejor simula el autoencendido del diésel, es: 24% n- heptano, 45% tolueno y 31% ciclohexano.

La generalidad y versatilidad de la metodología desarrollada hacen esta teoría fácilmente extensible a cualquier problema de diseño de mezclas. Otra línea de investigación consiste en el desarrollo de nuevos hormigones que persiguen optimizar alguna de sus propiedades. Por ejemplo, minimizar en hormigones de alta resistencia a compresión y flexión, las fisuraciones debidas a la retracción [9]. Este problema puede plantearse desde el diseño de mezclas puesto que éste se identifica como una mezcla de cemento, áridos y agua y, en definitiva, el fin es optimizar cierta característica deseable. Se abre así un nuevo campo de investigación donde nuevas posibilidades están siendo exploradas.