

SOLUCIÓN DE UN PROBLEMA DE ASIMILACIÓN DE DATOS Y UN PROBLEMA DE LOCALIZACIÓN ÓPTIMA MEDIANTE MÉTODOS DE OPTIMIZACIÓN BINIVEL

Paula Monserratte Castro Castro[†]; Juan Carlos De los Reyes Bueno[‡]

[†]*Escuela Politécnica Nacional, Facultad de Ciencias, Centro de Modelización Matemática (MODEMAT), Quito, Ecuador*

e-mail: paula.castro@epn.edu.ec

[‡]*Escuela Politécnica Nacional, Centro de Modelización Matemática (MODEMAT), Quito, Ecuador*

e-mail: juan.delosreyes@epn.edu.ec

Resumen: Los problemas de asimilación de datos han sido ampliamente estudiados en la predicción numérica del tiempo y del clima, como una técnica para la reconstrucción del estado inicial de la atmósfera. Tomando este problema como motivación, la meta de este proyecto es encontrar la solución de un problema de localización óptima de ubicaciones en una ecuación parabólica, de tal manera que se reconstruya la condición inicial, es decir se resuelva un problema de asimilación de datos. Se consideró un modelo de optimización a dos niveles, donde el nivel inferior se encarga de la reconstrucción de la condición inicial del estado del sistema, mientras el nivel superior resuelve el problema de localización óptima de observaciones.

Para resolver el problema de asimilación de datos se utilizó el enfoque variacional, específicamente el $4D$ -VAR. Se demostró la existencia y unicidad del problema de asimilación de datos utilizando técnicas usuales en optimización, mientras que para derivar el sistema de optimalidad se utilizó el enfoque Lagrangiano. Debido a la estructura del funcional objetivo se obtuvo como ecuación adjunta una EDP con medidas de Borel en el lado derecho y para demostrar la existencia de solución de esta ecuación se utilizó el método de la transposición. Para resolver el problema de localización, reemplazamos su restricción, es decir, el problema de asimilación de datos, por todo el sistema de optimalidad de este y utilizamos nuevamente la técnica Lagrangiana para derivar el sistema de optimalidad del problema en el nivel superior.

La resolución numérica del problema fue desarrollada también en dos niveles. El problema de asimilación de datos se resolvió utilizando el algoritmo BFGS, mientras que para encontrar direcciones de descenso y resolver el problema de localización óptima se utilizó el método BFGS proyectado mediante la estimación de conjuntos ϵ -activos, conjuntamente con una regla de búsqueda lineal de Armijo modificada y considerando un criterio de parada que utilice pasos completos.

Palabras claves: Asimilación de datos, localización óptima, optimización binivel, métodos proyectados.

Abstract: Data assimilation problems have been widely studied in numerical weather prediction as a technique for the reconstruction of the atmosphere's initial condition. Taking this problem as motivation, our goal in this project is finding the optimal placements of the locations of a data assimilation problem, represented by a parabolic equation. We worked in a bilevel optimization problem where the inner level solves the data assimilation problem, and the upper level solves the optimal placement problem.

To solve the data assimilation problem, we used a variational approach $4D$ -VAR. The existence and uniqueness of the data assimilation problem were demonstrated using usual techniques in optimization. By using the Lagrangian approach, we derive the optimality system of the inner problem. As result, we got an adjoint equation with measures on the right-hand side. This was due to the structure of the objective functional. We proved the existence of a unique very weak solution of the adjoint equation by using the transposition method.

The numerical solution was also worked in two levels. The inner problem was solved by using the BFGS method while the upper level used the BFGS projected method through an estimation of ϵ -active sets jointly with a modified Armijo line search and a stop criterion which takes full steps.

Keywords: Data assimilation, optimal placement problem, bilevel optimization, projected methods.

1. INTRODUCCIÓN

El proceso de asimilación de datos se describe como el procedimiento mediante el cual se analiza toda la información disponible tal como observaciones o fenómenos físicos, para tratar de estimar de mejor manera el estado de un sistema que evoluciona en el tiempo. La asimilación de datos es ampliamente usada en la predicción numérica del tiempo y del clima, ya que si se conociera la condición inicial exacta del estado de la atmósfera, sería posible obtener un pronóstico meteorológico preciso para cualquier instante de tiempo [15]. Una de los enfoques utilizados para resolver el problema de asimilación de datos es el variacional, específicamente el 4D-VAR que considera observaciones distribuidas en un intervalo de tiempo $[t_0, t_n]$ [7].

Los datos necesarios para la asimilación pueden ser recopilados de diversas fuentes: radio sondeos, imágenes satelitales, estaciones de medición, entre otras. Sin embargo, debido al costo o a la dificultad en la instalación de los equipos que permiten obtener estos datos se busca poder localizarlos de manera óptima, de tal forma que la información obtenida de ellos sea útil para el proceso de asimilación y con esto a la reconstrucción de la condición inicial del estado de la atmósfera.

Tomando en cuenta la resolución de este problema común en meteorología como motivación, lo que se busca en este proyecto es resolver un problema de localización óptima de observaciones de tal manera que se logre la reconstrucción de la condición inicial de un problema parabólico. Para ello, se considerará un modelo de optimización a dos niveles donde el nivel inferior se encargará de la reconstrucción de la condición inicial del estado del sistema, es decir, de encontrar la solución a un problema de asimilación de datos, mientras que el nivel superior resolverá el problema de localización óptima de observaciones.

En este trabajo se propone la utilización de la técnica de conjuntos de entrenamiento, ampliamente utilizada en Machine Learning, la cual consiste en predecir información de un dato desconocido basándose en la información aprendida de una muestra de datos conocidos. En nuestro caso el conjunto de entrenamiento estará formado por simulaciones de la condición inicial y observaciones del estado del sistema parabólico estudiado respectivamente. Lo que se busca obtener es un vector de localizaciones que en promedio sea óptimo para todos los pares de entrenamiento considerados. La ventaja de abordar el problema de localización de esta manera es que no se realizan supuestos sobre la distribución de probabilidad que siguen los datos a ser asimilados, como es el caso de las estimaciones Bayesianas [2], [1].

Finalmente, la resolución numérica del problema también se desarrollará en dos niveles. Para el problema interior, es decir el de asimilación de datos, se utilizará el método BFGS. Se considera la utilización de un método de segundo orden debido a su velocidad de convergencia, ya que el sistema de optimalidad del problema del

nivel inferior será la restricción del problema del nivel superior, es decir del problema de localización óptima. Para la resolución numérica del problema del nivel superior se emplearán métodos Quasi Newton proyectados, específicamente el método BFGS proyectado mediante la estimación de conjuntos ϵ -activos.

2. PROBLEMA DE ASIMILACIÓN DE DATOS

2.1 Planteamiento del problema

Para resolver el problema de asimilación de datos se utilizará el enfoque variacional mediante la resolución del problema 4D-VAR que considera diferentes observaciones distribuidas en un intervalo de tiempo $[t_0, t_l]$ [7]. Matemáticamente resolver el problema 4D-VAR significa resolver el siguiente problema de optimización

$$\begin{aligned} \min_u J(y, u) &= \frac{1}{2} \sum_{i=1}^l [H(y(t_i) - z_o(t_i))]^T R_i^{-1} [H(y(t_i) - z_o(t_i))] \\ &\quad + \frac{1}{2} [u - u^b]^T B^{-1} [u - u^b] \\ \text{sueto a:} & \\ y(t_i) &= M(y(t_0)) \quad (\text{Sistema de EDP's}) \\ y(t_0) &= u \quad (\text{Condición inicial}). \end{aligned} \quad (1)$$

En la formulación anterior z_o representa el estado observado, u^b la información previa o *background*. H es un operador de observación, que transforma las variables del modelo en variables observables. Para todo $i = 1, \dots, l$, los R_i representan matrices de covarianza respecto a los errores de las observaciones y B a la matriz de covarianza de los errores del background. Dependiendo si el análisis se realiza en dimensión finita o infinita se trabajará con matrices u operadores de covarianza respectivamente, para el caso de dimensión infinita los productos matriciales de la formulación (1) serán remplazados por productos internos definidos en espacios funcionales convenientes.

Para pasar de la formulación general de un problema 4D-VAR al problema específico de asimilación de datos con el cual se trabajará, se fijan los operadores de covarianza del background y de los errores iguales al operador identidad, se toma además $u^b \in L^2(\Omega)$ igual a cero. El operador de observaciones H viene dado por el vector en \mathbb{R}^{μ_E} de localizaciones $w = (w_k)$, $k = 1, \dots, \mu_E$ de entradas binarias, donde μ_E representa el número total de puntos en el dominio espacial en el que se este trabajando. Los w_k toman el valor de uno si la posición x_k es seleccionada y cero caso contrario. El operador de observaciones H permite además la evaluación de las variables en entradas puntuales en el espacio. El sistema de EDP's que representa la restricción del problema corresponde a un problema parabólico lineal y continuo, específicamente la ecuación del calor con condiciones de Dirichlet homogéneas. Así, el problema de asimilación de datos en concreto que se va a resolver es el siguiente:

$$\min_u J(y, u) = \frac{1}{2} \sum_k \sum_i w_k (y(x_k, t_i) - z_o(x_k, t_i))^2 + \frac{\alpha}{2} \int_{\Omega} |u|^2 \quad (2)$$

$$\text{s.a } \begin{cases} \frac{\partial y}{\partial t} - \Delta y = 0 & \text{en } Q = \Omega \times]0, T[\\ y = 0 & \text{en } \Sigma = \Gamma \times]0, T[\\ y(0) = u & \text{en } \Omega. \end{cases} \quad (3)$$

2.2 Existencia de solución del problema de optimización

Sean $\Omega \subset \mathbb{R}^n$ un dominio acotado tipo Lipschitz, con frontera Γ . Definimos $Q = \Omega \times (0, T)$ y $\Sigma = \Gamma \times (0, T)$ con un número real $T > 0$ fijo. Notemos que (3) corresponde a la ecuación del calor con condiciones de Dirichlet homogéneas, que es una ecuación lineal correctamente definida, es decir, para cada $u = u(x) \in L^2(\Omega)$, existe una única solución $y(x, t) \in W(0, T)$. Donde además se verifica la desigualdad siguiente

$$\|y\|_{W(0, T)} \leq c_w \|u\|_{L^2(\Omega)}, \text{ con } c_w > 0.$$

(Ver [14], pp.149-151). Consecuentemente, la función solución que asigna $u \mapsto y$ define un operador lineal y continuo, notado G_0 que va de $L^2(\Omega)$ en $W(0, T)$, en particular en $C([0, T]; L^2(\Omega))$. Consideremos el operador que representa la inyección canónica

$$E : W(0, T) \rightarrow L^2(0, T; L^2(\Omega))$$

también lineal y continuo que asigna a cada $y \in W(0, T)$ la misma función en $L^2(Q)$, con $L^2(0, T; L^2(\Omega)) \cong L^2(Q)$ y definiendo el operador S como la composición de G_0 con E , tenemos

$$S : L^2(\Omega) \rightarrow L^2(Q) \\ u \mapsto Su = y$$

El operador S recibirá el nombre de operador control-estado. Notemos además que dicho operador es a su vez lineal y continuo por ser la composición de operadores lineales y continuos. Reescribiendo el funcional objetivo (2) tomando en cuenta el operador control-estado S obtenemos el funcional reducido

$$\min_u f(u) = \frac{1}{2} \sum_k \sum_i w_k (Su(x_k, t_i) - z_o(x_k, t_i))^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega)}^2 \quad (4)$$

TEOREMA 1. Si $\alpha > 0$. Entonces:

- (i) $f : L^2(\Omega) \rightarrow \mathbb{R}$ definida como en (4) es débilmente semicontinua inferior y radialmente no acotada.
- (ii) El problema de minimización (4) tiene solución óptima única.

Demostración.

(i) Notemos que

$$f(u) = \frac{1}{2} \sum_k \sum_i w_k (Su(x_k, t_i) - z_o(x_k, t_i))^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega)}^2 \\ \geq \frac{\alpha}{2} \|u\|_{L^2(\Omega)}^2.$$

De donde se sigue inmediatamente que f es radialmente no acotada, pues $f(u) \rightarrow +\infty$ cuando $\|u\| \rightarrow +\infty$.

Por otro lado, debido a que f es continua y convexa es también semicontinua inferior (ver [14], pp.47).

(ii) Tomemos $\{u_n\}_{n \geq 1} \subset L^2(\Omega)$ una sucesión minimizante, es decir

$$\inf_{u \in L^2(\Omega)} f(u) = \lim_{n \rightarrow +\infty} f(u_n).$$

Notemos que la sucesión $\{u_n\}$ es acotada. Razonando por contradicción, si suponemos que no lo es, se contradice que f sea radialmente no acotada, lo cual fue demostrado en (i). Por otro lado, como $L^2(\Omega)$ es un espacio reflexivo sabemos que existe una sub-sucesión $\{u_{n_k}\} \subset \{u_n\}$ tal que $u_{n_k} \rightharpoonup \bar{u}$ cuando $k \rightarrow +\infty$ (ver [14], pp.46). Luego, gracias a la semicontinuidad inferior de f se sigue que

$$f(\bar{u}) \leq \lim_{k \rightarrow \infty} \inf f(u_{n_k}) = \inf_{u \in L^2(\Omega)} f(u).$$

De donde $\bar{u} \in L^2(\Omega)$ es una solución óptima de (4).

La unicidad de esta solución se sigue de la convexidad estricta de la función f , lo cual se obtiene al exigir $\alpha > 0$. □

Notemos además que al estar minimizando sobre todos los $u \in L^2(\Omega)$, necesitamos que el parámetro de regularización de Tikhonov sea estrictamente positivo para poder garantizar la existencia de la solución del problema de optimización, por lo cual no consideraremos el caso $\alpha = 0$.

Gracias al Teorema 1, se demuestra la existencia y unicidad de $\bar{u} \in L^2(\Omega)$ control óptimo del problema, con $\bar{y} = y(\bar{u}) \in W(0, T)$ estado óptimo asociado.

2.3 Sistema de optimalidad

Definamos $X := L^2(0, T; W_0^{1,r})$ el espacio de las funciones de $[0, T]$ a $W_0^{1,r}(\Omega)$ medibles con $r \in [1, \frac{n}{n-1}]$ tal que

$\int_0^T \|g(t)\|_{W_0^{1,r}(\Omega)}^2 dt < \infty$. Notando como es usual a la ecuación de estado como $e(y, u) = 0$. Donde $e(y, u) = 0$ verifica que

$$\langle e(y, u), p \rangle_{X^*, X} = \int_{\Omega} (y(T)p(T) - up(0)) dx \\ - \int_{\Omega} \int_0^T (p_t y + \nabla y \nabla p) dx dt - \int_{\Gamma} \int_0^T p \partial_\nu y dx dt$$

para cualquier $p \in X$ y donde $\partial_\nu y$ representa la derivada de y en el sentido de la normal. Para obtener el sistema de optimalidad del problema de asimilación de datos se utilizará el enfoque Lagrangiano. Sea p el multiplicador

de Lagrange asociado a la restricción $e(y, u) = 0$, el operador Lagrangiano queda expresado por

$$\mathcal{L} : W(0, T) \times L^2(\Omega) \times X \rightarrow \mathbb{R}$$

$$(y, u, p) \mapsto \mathcal{L}(y, u, p) = J(y, u) + \langle e(y, u), p \rangle_{X^*, X}$$

es decir

$$\mathcal{L}(y, u, p) = \sum_k \sum_i w_k [y(x_k, t_i) - z_o(x_k, t_i)]^2 + \frac{\alpha}{2} \|u\|_{L^2(\Omega)}^2$$

$$+ \int_{\Omega} \int_0^T (-p_t - \Delta p) y dx dt - \int_{\Gamma} \int_0^T p \partial_\nu y dx dt$$

$$+ \int_{\Omega} (y(T)p(T) - up(0)) dx.$$

El sistema de optimalidad estará correctamente definido si $\mathcal{L}_y(\bar{y}, \bar{u}, p) = 0$, que corresponde a la ecuación adjunta del sistema, tiene solución.

El estado adjunto se obtiene al derivar $\mathcal{L}(y, u, p)$ respecto a y en alguna dirección v e igualar el resultado a cero. Así, notando $\mathcal{L}_y = \mathcal{L}_y(y, u, p)$, tenemos

$$\mathcal{L}_y(v) = \sum_k \sum_i w_k (y(x_k, t_i) - z_o(x_k, t_i)) v(x_k, t_i)$$

$$+ \int_{\Omega} \int_0^T (p_t - \Delta p) v dx dt + \int_{\Omega} v(T)p(T) dx$$

$$- \int_{\Gamma} \int_0^T p \partial_\nu v dx dt = 0,$$

en donde la doble sumatoria puede escribirse formalmente como

$$\int_{\Omega} \int_0^T \sum_{k,i} w_k (y(x, t) - z_o(x, y)) v(x, t) \delta(x - x_k) \delta(t - t_i) dx dt,$$

donde $\delta(x - x_k)$ y $\delta(t - t_i)$ representan la masa de Dirac concentrada en los puntos x_k y t_i respectivamente ([9], pp.79). Estas distribuciones devuelven el valor de un funcional cuando es evaluado en un punto, es decir

$$\int_{\Omega} g(x) \delta(x - x_k) dx = g(x_k).$$

De donde

$$\mathcal{L}_y(v) = \int_{\Omega} \int_0^T \left(\sum_{k,i} w_k [y - z_o] \delta(x - x_k) \delta(t - t_i) - p_t - \Delta p \right) v$$

$$- \int_{\Gamma} \int_0^T p \partial_\nu v dx dt + \int_{\Omega} p(T)v(T) dx = 0.$$

Así, la formulación fuerte de la ecuación adjunta está dada por

$$-p_t - \Delta p = \sum_{k,i} w_k [z_o - y] \otimes \delta(x - x_k) \otimes \delta(t - t_i) \quad \text{en } Q$$

$$p = 0 \quad \text{en } \Sigma$$

$$p(x, T) = 0 \quad \text{en } \Omega,$$

la formulación matemática del problema es la siguiente:

$$(5) \quad \min_{w \in \{0,1\}} J(y, w) = \sum_{j=1}^N \left(\|y_j^{Train} - y_j\|^2 + \beta \|u_j^{Train} - u_j\|^2 \right) + \gamma \|w\|_{l_0}$$

$$g(x) \otimes \delta(x - x_k) = \int_{\Omega} g(x) \delta(x - x_k) dx = g(x_k).$$

De manera similar, la ecuación del gradiente se obtiene al derivar el operador Lagrangiano con respecto a u en alguna dirección h e igualar a cero el resultado, es decir

$$\mathcal{L}_u(h) = \alpha \langle u, h \rangle_{L^2(\Omega)} - \int_{\Omega} hp(0) dx$$

$$= \int_{\Omega} (\alpha u - p(0)) h dx = 0, \quad \forall h \in L^2(\Omega)$$

de donde la ecuación del gradiente del problema (2) queda expresada por

$$\alpha u - p(0) = 0 \quad \text{en } L^2(\Omega). \quad (6)$$

Finalmente para que el sistema de optimalidad quede bien definido resta demostrar que (5) tiene solución en un espacio funcional apropiado.

Utilizando el método de la transposición, se puede demostrar que existe $p \in L^2(Q)$ única solución muy débil de (5) (ver [9], pp.183). Además, como el lado derecho de este sistema pertenece a las medidas regulares de Borel en Q , $\mathcal{M}(Q)$, se puede concluir que $p \in L^2(0, T; W_0^{1,r})$ con $r \in [1, \frac{n}{n-1}]$ ([3], pp 31). Es decir, podemos concluir que $p \in X$.

3. PROBLEMA DE LOCALIZACIÓN ÓPTIMA DE OBSERVACIONES

Dependiendo de la calidad de la información, así como la cantidad de observaciones de las que se disponga, el proceso de asimilación de datos será más o menos útil para la reconstrucción de una condición inicial deseada. Los datos necesarios para la asimilación pueden ser recopilados, por ejemplo, de estaciones de medición. Sin embargo, debido al costo o a la dificultad en la instalación de los equipos que permiten obtener estos datos se busca poder localizarlos de manera óptima, de tal forma que la información obtenida sea útil para el proceso de asimilación. Para ello, se considerará un modelo de optimización a dos niveles donde el nivel inferior se encargará de la reconstrucción de la condición inicial del estado del sistema, es decir, de encontrar la solución a un problema de asimilación de datos, mientras que el nivel superior resolverá el problema de localización óptima de observaciones. Usando la notación

$$y = (y(w), p(w), u(w)),$$

$$\begin{cases} \min_u J(y, u) = \frac{1}{2} \sum_{k,i} w_k (y_j(x_k, t_i) - y_j^{Train}(x_k, t_i))^2 + \frac{\alpha}{2} \int_{\Omega} \|\nabla y\|^2 \\ \text{s.a.} \begin{cases} \frac{\partial y_j}{\partial t} - \Delta y_j = 0 \text{ en } Q \\ y_j = 0 \text{ en } \Sigma \\ y_j(0) = u_j \text{ en } \Omega. \end{cases} \end{cases} \quad (9)$$

$\forall j = 1, \dots, N$. Debido a la dificultad que implica trabajar con la "norma" l_0 o de conteo, se realiza una primera aproximación del problema utilizando $\|\cdot\|_{l_1}$ en lugar de $\|\cdot\|_{l_0}$, de donde el nuevo funcional objetivo a resolver sería no diferenciable. Por lo tanto no se pueden usar las técnicas usuales de optimización que asumen diferenciable de la función objetivo. Se considera entonces una relajación del problema, donde las entradas de w toman valores entre 0 y 1. Con este cambio se consigue la diferenciable del funcional objetivo, pero el problema se convierte en uno con restricciones tipo caja.

$$\min_{0 \leq w \leq 1} J(y, w) = \sum_{j=1}^N (\|y_j^{Train} - y_j\|^2 + \beta \|u_j^{Train} - u_j\|^2) + \gamma \sum_k w_k \quad (7)$$

El índice j recorre el conjunto de entrenamiento o training set y está dado por las duplas $(u_j^{Train}, y_j^{Train})$, que representan simulaciones de la condición inicial y el estado observado. La idea de trabajar con conjuntos de entrenamiento proviene del Machine Learning y consiste en aprender del conjunto dado y aplicar la información aprendida para predecir alguna característica en un dato desconocido. En el contexto en el que estamos trabajando lo que nos interesa aprender del conjunto de entrenamiento es precisamente el vector de localizaciones w , de tal manera que las ubicaciones dadas por él sean en promedio óptimas para todos los pares considerados.

3.1 Sistema de Optimalidad

Para obtener el sistema de optimalidad, consideraremos un caso particular del mismo tomado $j = 1$, es decir, se trabajará solamente con un elemento del conjunto de entrenamiento, ya que los resultados obtenidos pueden extenderse directamente al caso general, es decir, cuando $j > 1$. Con esta consideración, (7) puede ser reescrito como

$$\min_{0 \leq w \leq 1} J(y, w) = \|y^{Train} - y\|^2 + \beta \|u^{Train} - u\|^2 + \gamma \sum_k w_k, \quad (8)$$

donde (y, w) es la solución de:

$$\begin{cases} \frac{\partial y}{\partial t} - \Delta y = 0, \text{ en } Q \\ y = 0, \text{ en } \Sigma \\ y(0) = u, \text{ en } \Omega \\ -\frac{\partial p}{\partial t} - \Delta p = \sum_{k,i} w_k [y^{Train} - y] \otimes \delta(x - x_k) \otimes \delta(t - t_i) \\ p = 0, \text{ en } \Sigma \\ p(T) = 0, \text{ en } \Omega \\ \alpha u - p(0) = 0, \text{ en } \Omega. \end{cases} \quad (9)$$

Sea μ_E el número total de observaciones en el espacio, notemos que para cada $w \in \mathbb{R}^{\mu_E}$, la ecuación de estado

(9) tiene solución única. Este resultado se formaliza en el siguiente teorema.

TEOREMA 2. Si $n \leq 3$ y $\alpha > 0$ con α el parámetro de regularización de Tikhonov del problema de asimilación de datos, entonces para cada $w \in \mathbb{R}^{\mu_E}$ la ecuación de estado (9) tiene solución única determinada por w ,

$$(y, p, u) \in W(0, T) \times X \times L^2(\Omega).$$

Demostración. En efecto, el sistema (9) corresponde también al sistema de optimalidad del problema de asimilación de datos dados por los sistemas (3), (5) y (6), donde se estableció para cada $u \in L^2(\Omega)$ la existencia y unicidad de $y \in W(0, T)$ solución de (3) y de $p \in X$ solución muy débil de (5) tal que verifican la ecuación del gradiente (6), con lo cual se concluye el resultado. \square

Notemos como \mathbf{Z} y \mathbf{Z}^* al siguiente espacio y su dual respectivamente

$$\mathbf{Z} = X \times W(0, T) \times L^2(\Omega)$$

$$\mathbf{Z}^* = X^* \times W(0, T)^* \times L^2(\Omega).$$

Bajo el supuesto de que existe un control óptimo único $\bar{w} \in \mathbb{R}^{\mu_E}$ del problema (8) con $\bar{y} := (y(\bar{w}), p(\bar{w}), u(\bar{w})) \in \mathbf{Y} := W(0, T) \times X \times L^2(\Omega)$ su estado óptimo asociado. Notando $\varepsilon(y, w) = 0$, la restricción del problema de localización, donde

$$\varepsilon : \mathbf{Y} \times \mathbb{R}^{\mu_E} \rightarrow \mathbf{Z}^*$$

$$(y, w) \mapsto \varepsilon(y, w) = (\varepsilon_1(y, w), \varepsilon_2(y, w), \varepsilon_3(y, w))$$

y donde $\varepsilon_1, \varepsilon_2, \varepsilon_3$ verifican lo siguiente:

$$\langle \varepsilon_1(y, w), \eta \rangle_{X^*, X} = \int_{\Omega} (y(T)\eta(T) - u\eta(0)) - \int_{\Omega} \int_0^T (\eta_t y + \nabla \eta \nabla y) - \int_{\Gamma} \int_0^T \eta \partial_\nu y, \forall \eta \in X$$

$$\begin{aligned} \langle \varepsilon_2(y, w), \sigma \rangle_{W^*, W} &= \int_{\Omega} (p(0)\sigma(0)) + \int_{\Omega} \int_0^T (\sigma_t p + \nabla \sigma \nabla p) \\ &\quad - \int_{\Omega} \int_0^T \sigma \left(\sum_{k,i} w_k [y^{Train} - y] \delta(x - x_k) \delta(t - t_i) \right) \\ &\quad - \int_{\Gamma} \int_0^T \sigma \partial_\nu p, \forall \sigma \in W := W(0, T) \end{aligned}$$

$$\langle \varepsilon_3(y, w), \theta \rangle_{L^2(\Omega), L^2(\Omega)} = \int_{\Omega} \theta (\alpha u - p(0)), \forall \theta \in L^2(\Omega).$$

Para derivar el sistema de optimalidad del problema de localización óptima se utilizará nuevamente el enfoque Lagrangiano. Sea \mathbf{p} el multiplicador de Lagrange, con $\mathbf{p} := (\eta, \sigma, \theta)$ asociado a la restricción $\varepsilon(y, w) = 0$, el operador Lagrangiano queda expresado por

$$\mathcal{L} : \mathbf{Y} \times \mathbb{R}^{\mu_E} \times \mathbf{Z} \rightarrow \mathbb{R}$$

$$(\mathbf{y}, w, \mathbf{p}) \mapsto \mathcal{L}(\mathbf{y}, w, \mathbf{p}) = J(\mathbf{y}, w) + \langle \varepsilon(\mathbf{y}, w), \mathbf{p} \rangle_{Z^*, Z}$$

Es decir,

$$\begin{aligned} \mathcal{L}(\mathbf{y}, w, \mathbf{p}) &= J(\mathbf{y}, w) + \int_{\Omega} (y(T)\eta(T) - u\eta(0)) - \int_{\Omega} \int_0^T (\eta_t + \Delta\eta) \\ &\quad - \int_{\Gamma} \int_0^T \eta \partial_\nu y + \int_{\Omega} (p(0)\sigma(0)) dx + \int_{\Omega} \int_0^T (\sigma_t - \Delta\sigma) p \\ &\quad + \int_{\Omega} \int_0^T \sigma \left(- \sum_{k,i} w_k [y^{Train} - y] \delta(x - x_k) \delta(t - t_i) \right) \\ &\quad - \int_{\Gamma} \int_0^T \sigma \partial_\nu p + \int_{\Omega} \theta (\alpha u - p(0)) \end{aligned} \quad (10)$$

La ecuación adjunta resulta de fijar en cero la derivada de $\mathcal{L}(\mathbf{y}, w)$ respecto a $\mathbf{y} = (y, p, u)$ en alguna dirección $v = (v_1, v_2, v_3)$, es decir,

$$\mathcal{L}_y(v) = \mathcal{L}_y(v_1) + \mathcal{L}_p(v_2) + \mathcal{L}_u(v_3) = 0. \quad (11)$$

Para hacerlo, vamos a considerar los sistemas de manera separada. Así,

$$\begin{aligned} \mathcal{L}_y(v_1) &= \int_{\Omega} \int_0^T \left(-\frac{\partial\eta}{\partial t} - \Delta\eta - 2(y^{Train} - y) \right) v_1 \\ &\quad + \int_{\Omega} \int_0^T \left(\sum_{k,i} w_k \sigma(x, y) \delta(x - x_k) \delta(t - t_i) \right) v_1 \\ &\quad - \int_{\Gamma} \int_0^T \eta \partial_\nu v_1 + \int_{\Omega} (\eta(T)v_1(T)) = 0. \end{aligned}$$

La formulación fuerte de este primer sistema está dada por

$$\begin{aligned} -\frac{\partial\eta}{\partial t} - \Delta\eta &= 2(y^{Train} - y) - \sum_{k,i} w_k \sigma \otimes \delta(x - x_k) \otimes \delta(t - t_i) \\ \eta &= 0 \text{ en } \Sigma \\ \eta(T) &= 0 \text{ en } \Omega \end{aligned} \quad (12)$$

Ahora,

$$\begin{aligned} \mathcal{L}_p(v_2) &= \int_{\Omega} \int_0^T \left(\frac{\partial\sigma}{\partial t} - \Delta\sigma \right) v_2 + \int_{\Gamma} \int_0^T \sigma \partial_\nu v_2 \\ &\quad + \int_{\Omega} (v_2(0)[\sigma(0) - \theta]) = 0 \end{aligned}$$

Y su formulación fuerte es

$$\begin{aligned} \frac{\partial\sigma}{\partial t} - \Delta\sigma &= 0 \text{ en } Q \\ \sigma &= 0 \text{ en } \Sigma \\ \sigma(0) &= \theta \text{ en } \Omega. \end{aligned} \quad (13)$$

Finalmente,

$$\begin{aligned} \mathcal{L}_u(v_3) &= -2\beta \int_{\Omega} (u^{Train} - u) v_3 - \int_{\Omega} \eta(0)v_3 + \int_{\Omega} \theta \alpha v_3 \\ &= \int_{\Omega} \left(-2\beta(u^{Train} - u) - \eta(0) + \theta \alpha \right) v_3 = 0 \end{aligned}$$

de donde

$$\theta = \frac{1}{\alpha} [2\beta(u^{Train} - u) + \eta(0)]. \quad (14)$$

La solución de (11) se da al resolver (12), (13) y (14) simultáneamente, así la ecuación adjunta del problema de localización está dada por

$$\begin{aligned} -\frac{\partial\eta}{\partial t} - \Delta\eta &= 2(y^{Train} - y) - \sum_{k,i} w_k \sigma(x, t) \otimes \delta(x - x_k) \otimes \delta(t - t_i) \\ \eta &= 0 \text{ en } \Sigma \\ \eta(T) &= 0 \text{ en } \Omega \\ \frac{\partial\sigma}{\partial t} - \Delta\sigma &= 0 \text{ en } Q \\ \sigma &= 0 \text{ en } \Sigma \\ \sigma(0) &= \frac{1}{\alpha} [2\beta(u^{Train} - u) + \eta(0)] \text{ en } \Omega. \end{aligned} \quad (15)$$

La ecuación del gradiente del problema de localización es el resultado de fijar en cero la derivada de $\mathcal{L}(\mathbf{y}, w)$ respecto a w en alguna dirección $h = (h_1, \dots, h_{\mu_E})$. Así,

$$\mathcal{L}_{w_k}(h_k) = \left(\gamma - \sum_i \sigma(x_k, t_i) (y^{Train}(x_k, t_i) - y(x_k, t_i)) \right) \cdot h_k,$$

de donde, la k -ésima componente de la ecuación del gradiente del problema a dos niveles está dada por:

$$\nabla f(w)_k = \gamma - \sum_i \sigma(x_k, t_i) (y^{Train}(x_k, t_i) - y(x_k, t_i)), \quad \forall k \quad (16)$$

donde (16) satisface la desigualdad variacional

$$\nabla f(w)(v - w) \geq 0, \quad \forall v \in U_{ad} \quad (17)$$

Si definimos,

$$\begin{aligned} \lambda_a &= \max\{0, \nabla f(w)\} \\ \lambda_b &= |\min\{0, \nabla f(w)\}|, \end{aligned}$$

la desigualdad variacional (17) es equivalente al siguiente sistema

$$\begin{cases} (\lambda_a)_k \geq 0, (\lambda_b)_k \geq 0 & , \forall k = 1, \dots, \mu_E \\ (\lambda_a)_k(0 - w_k) = (\lambda_b)_k(w_k - 1) = 0 & , \forall k = 1, \dots, \mu_E \\ 0 \leq w_k \leq 1 & , \forall k = 1, \dots, \mu_E \end{cases} \quad (18)$$

Aquí, asumiremos la existencia de solución de la ecuación adjunta (15), con lo cual el sistema de optimalidad del problema de optimización a dos niveles quedaría correctamente definido.

3.2 Métodos Quasi-Newton proyectados

De manera general los métodos de proyección usan la dirección de descenso de un problema sin restricciones y proyectan la nueva iteración dentro del conjunto admisible ([5], pp.75). Así, las actualizaciones de w_k estarán dadas por

$$w_{k+1} = P_{U_{ad}}(w_k + \alpha_k d_k),$$

donde U_{ad} es el conjunto de soluciones admisibles definido por a_i y b_i y donde P representa el operador proyección. d_k representa alguna dirección de descenso y $\alpha_k \in (0, 1)$ un parámetro de búsqueda lineal, que no es calculado de la misma manera que en los métodos no

proyectados. Así por ejemplo, una regla de Armijo modificada puede ser usada para hallar el valor más grande de α tal que

$$f(P_{U_{ad}}(w_k + \alpha_k d_k)) - f(w_k) \leq -\frac{\hat{\gamma}}{\alpha_k} \|P_{U_{ad}}(w_k + \alpha_k d_k) - w_k\|^2 \quad (19)$$

El criterio de parada para los métodos de proyección también difiere en relación con los métodos no proyectados, podemos escoger por ejemplo el siguiente criterio de parada que considera pasos completos

$$\|w_k - P_{U_{ad}}(w_k - \nabla f(w_k))\| < \hat{\epsilon}$$

para algún $0 < \hat{\epsilon} \ll 1$ ([5], pp.77).

En los métodos de segundo orden proyectados, en lugar de utilizar la información de la Hessiana, se utiliza la información proporcionada por la Hessiana reducida, basada en la estimación de conjuntos ϵ -activos, los cuales están dados por

$$A^\epsilon(w) = \{i : a_i \leq w_i \leq a_i + \epsilon \text{ o } b_i \geq w_i \geq b_i - \epsilon\}.$$

Si se considera el caso cuando $\epsilon = 0$, el conjunto recibe el nombre de conjunto activo. Por otro lado, $I^\epsilon(w)$ e $I(w)$ representan el conjunto ϵ -inactivo e inactivo, y son los complementos de los conjuntos activos, respectivamente.

De manera general, si S representa un conjunto de índices cualquiera, entonces R_S denotará la matriz

$$R_S = (\delta_{ij}), \text{ si } i \in S \text{ o } j \in S$$

donde δ_{ij} representa la delta de Kronecker. Así, la matriz Hessiana reducida para la iteración w_k se define como

$$\begin{aligned} \tilde{R}(w_k, \epsilon_k, H_k) &= R_{A^{\epsilon_k}(w_k)} + R_{I^{\epsilon_k}(w_k)} H_k R_{I^{\epsilon_k}(w_k)} \\ &= \begin{cases} \delta_{ij}, & \text{si } i \in A^{\epsilon_k}(w_k) \text{ o } j \in A^{\epsilon_k}(w_k); \\ (H_k)_{ij}, & \text{caso contrario.} \end{cases} \end{aligned}$$

y las iteraciones del método estarán dadas por

$$w_{k+1} = P_{U_{ad}}(w_k - \alpha_k \tilde{R}(w_k, \epsilon_k, H_k)^{-1} \nabla f(w_k)). \quad (20)$$

Si la matriz H_k simboliza a la matriz Hessiana, (20) representa las iteraciones del método de Newton proyectado. Si en cambio H_k representa la aproximación de la matriz Hessiana dada por las actualizaciones del método BFGS, el método implementado será el BFGS proyectado.

3.2.1 BFGS Proyectado

Sea H_k la aproximación de la matriz Hessiana dada por el método del BFGS. Por simplicidad en la notación de aquí en adelante, notaremos de igual manera a su matriz reducida, es decir

$$H_{k+1} = R_{I^{\epsilon_k}(w_k)} H_k R_{I^{\epsilon_k}(w_k)} - R_{I^{\epsilon_k}(w_k)} \frac{H_k s_k s_k^T H_k}{s_k^T H_k s_k} R_{I^{\epsilon_k}(w_k)} + \frac{y_k^\# (y_k^\#)^T}{s_k^T y_k^\#},$$

que representa la matriz de iteraciones del método BFGS proyectado ([8], pp.102), con

$$y_k^\# = R_{I^{\epsilon_k}(w_k)} (\nabla f(w_{k+1}) - \nabla f(w_k)). \quad (21)$$

De manera similar al caso no proyectado existe una actualización de la matriz inversa para el método del BFGS proyectado. Sea $B_k = H_k^{-1}$, si los conjuntos activos en las iteraciones w_k y w_{k+1} no cambian, es decir, $A(w_k) = A(w_{k+1})$. Entonces se puede actualizar a la matriz B_{k+1} como sigue

$$B_{k+1} = \left(I - \frac{s_k (y_k^\#)^T}{s_k^T y_k^\#} \right) B_k \left(I - \frac{y_k^\# s_k^T}{s_k^T y_k^\#} \right) + \frac{s_k s_k^T}{s_k^T y_k^\#}, \quad (22)$$

con

$$s_k^\# = R_{I^{\epsilon_k}(w_k)} (w_{k+1} - w_k).$$

Por otro lado, si los conjuntos activos cambian de iteración en iteración, (22) ya no se verifica. En su lugar, se considera una aproximación basada en la forma recursiva de las actualizaciones de la matriz BFGS ([8], pp.103).

$$B_{k+1} = \left(I - \frac{s_k^\# (y_k^\#)^T}{(y_k^\#)^T s_k^\#} \right) R_I B_k R_I \left(I - \frac{y_k^\# (s_k^\#)^T}{(y_k^\#)^T s_k^\#} \right) + \frac{s_k^\# (s_k^\#)^T}{(y_k^\#)^T s_k^\#}, \quad (23)$$

donde $R_I = R_{I^{\epsilon_k}(w_k)}$.

Para encontrar direcciones de descenso para resolver el problema de localización óptima de observaciones se utilizará las actualizaciones de la matriz inversa del BFGS proyectado conjuntamente con una regla de búsqueda lineal. El conjunto sobre el cual se buscará el parámetro α_k mediante la regla de Armijo modificada es el siguiente

$$\left\{ \frac{1}{2^i \|\nabla f(w_0)\|} \text{ con } i = \{0, 1, 2, \dots\} \right\},$$

donde $\nabla f(w_0)$ representa el gradiente del problema evaluado en un vector de localizaciones iniciales w_0 dado. Notemos que este es un conjunto factible para la búsqueda del parámetro si se toma w_0 tal que

$$0 < \frac{1}{2^k \|\nabla f(w_0)\|} < 1, \quad \forall k > 1.$$

4. IMPLEMENTACIÓN NUMÉRICA

Para la implementación numérica del problema de localización en primer lugar se debe resolver la ecuación de estado, lo que significa resolver todo el sistema de optimalidad del problema de asimilación de datos. Gracias a la velocidad de convergencia superlineal del método BFGS, se utilizará este para calcular la solución del problema de asimilación. Después, con esta información se calculará el sistema adjunto, que corresponde a resolver un sistema acoplado. Luego, con los valores de los sistemas estado y adjunto se calculará la ecuación del gradiente del problema de localización y finalmente con el método del BFGS proyectado se obtendrá una dirección de descenso, necesaria para calcular el vector de localizaciones óptimas.

Experimento

El objetivo principal de este experimento es observar cómo varía la estructura del vector de localizaciones óptimas al trabajar con distintos valores de los parámetros γ y β , presentes en el funcional objetivo del problema en el nivel superior. Donde γ representa un parámetro que controla en cierta manera el grado de dispersión del vector de localizaciones y β el parámetro de regularización de Tikhonov del problema de localización óptima de ubicaciones.

En el Cuadro 1 se presentan los resultados obtenidos con algunos valores de γ y fijando $\beta = 10^{-4}$. Se usará la siguiente notación: J_0 y J_{end} representarán los valores inicial y final del funcional objetivo, mientras que $0's$, I_1 , I_2 , y I_3 , $1's$ representaran los intervalos indicados anteriormente.

A continuación se presentan los resultados gráficos obtenidos, la Figura (1) muestran el decrecimiento de la función objetivo para diferentes valores de γ . La Figura (2) muestra la reconstrucción de la condición inicial deseada cuando se varía el valor de γ . Finalmente, la Figura (3) muestra la estructura del vector de localizaciones óptimas obtenido para diferentes valores de γ .

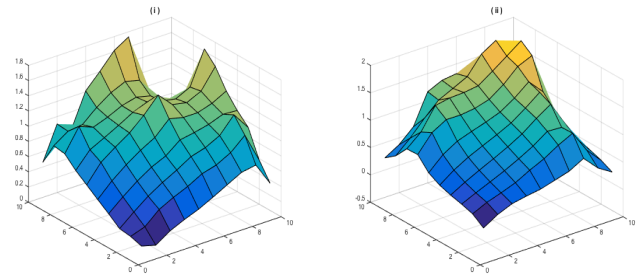


Figura 2. Control óptimo. (i) $\gamma = 0,1$, (ii) $\gamma = 10$.

Como se puede observar que la reconstrucción de la condición inicial deseada es más exacta cuando el vector de localizaciones w tiene menos entradas nulas.

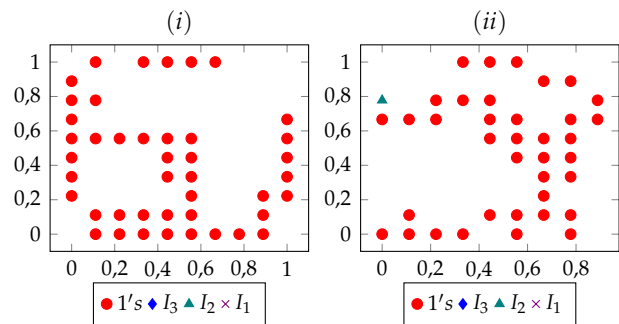


Figura 3. Estructura del vector de localizaciones. (i) $\gamma = 0,1$, (ii) $\gamma = 10$.

γ	$0's$	I_1	I_2	I_3	$1's$	J_0	J_{end}	iter
10^{-3}	68	0	0	1	31	0,063	0,033	6
10^{-2}	67	0	0	0	33	0,62	0,33	21
0,1	57	0	0	0	43	6,19	4,30	21
1	95	0	0	0	5	61,79	5,01	9
10	62	0	1	0	37	617,98	375,37	6
100	75	25	0	0	0	$6,17 \times 10^3$	0,021	17

Tabla 1. Experimento 1 - Diferentes valores de γ y β .

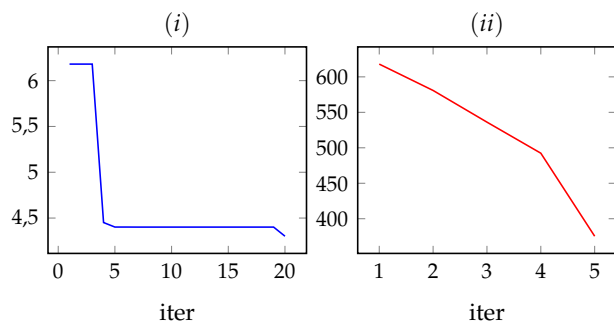


Figura 1. Decrecimiento de la función objetivo (i) $\gamma = 0,1$, (ii) $\gamma = 10$.

5. CONCLUSIONES

En el presente trabajo se ha tratado el problema de localización óptima de observaciones, de tal manera que se logre la reconstrucción de la condición inicial de un problema parabólico, como un problema de optimización continua a dos niveles.

Para resolver el problema de la reconstrucción de la condición inicial del estado del sistema, es decir, para encontrar la solución del problema de asimilación de datos, que corresponde al problema interior del problema binivel, se utilizó el enfoque variacional, específicamente la técnica del 4D-VAR que considera observaciones distribuidas en un intervalo de tiempo determinado. Mientras que la solución obtenida para el problema de localización, que corresponde al nivel superior del problema binivel, es un vector de localizaciones óptimas de tal manera que con las ubicaciones obtenidas se consiga una mejor reconstrucción de la condición inicial deseada y del estado del sistema observado, respectivamente.

Considerando que el número de observaciones que se pueden obtener es limitado, se incluyó en el planteamiento del problema al parámetro de penalización $\gamma > 0$, que actúa sobre w de la siguiente forma: a medida que más observaciones son tomadas en cuenta el producto $\gamma \sum_k |w_k|$, presente en el funcional objetivo, será mayor.

Al tratarse de un problema de minimización se favorecerá los casos en los cuales el vector w tenga más entradas

nulas. Este parámetro será denominado como parámetro de penalización del vector w . Sin embargo, debido a la no diferenciabilidad de este funcional se resolvió una relajación de este problema, haciéndolo diferenciable pero aumentando una restricción tipo caja al sistema.

La resolución numérica del problema también se la realizó a dos niveles. Debido a que todo el sistema de optimalidad del problema de asimilación es la restricción del problema en el nivel superior, se utilizó un método con una rápida velocidad de convergencia para la resolución de este primer sistema, el BFGS. Para encontrar direcciones de descenso y resolver el problema de localización óptima se utilizaron las actualizaciones de la matriz inversa del BFGS proyectado conjuntamente con una regla de búsqueda lineal de Armijo modificada. Es importante notar que el criterio de parada que se utilizó en la implementación numérica considera pasos completos.

Se desarrolló un algoritmo que permitió obtener un vector de localizaciones, control y estado óptimos que aproximen a la condición inicial simulada y al estado del sistema observado. Como una primera aproximación se realizó la parte experimental considerando solo una condición inicial simulada y un estado observado, es decir, se trabajó únicamente con la dupla $(u_1^{Train}, y_1^{Train})$.

Del experimento se observó que una potencial limitación del algoritmo implementado es la dependencia de los parámetros con los cuales se trabaja. Así también, el algoritmo no proporciona un control directo sobre el número de observaciones que se obtienen. En la práctica resolver el problema de localización óptima requerirá determinar de manera experimental los valores adecuados para los parámetros a utilizarse.

6. AGRADECIMIENTOS

Este trabajo fue parcialmente financiado por: el Instituto Nacional de Meteorología e Hidrología dentro del proyecto "Sistema de Pronóstico del Clima y el Tiempo para todo el territorio ecuatoriano: Modelización numérica y estadística - fase dos", la Escuela Politécnica Nacional dentro del proyecto "Flujos de materiales viscoplásticos en la industria alimenticia: modelización matemática, simulación numérica y optimización" y al programa de cooperación científico-tecnológica MATH-AmSud.

6. REFERENCIAS

- [1] A. Alexanderian, N. Petra, G. Stadler, and O. Ghattas. A-optimal design of experiments for infinite-dimensional bayesian linear inverse problems with regularized l_0 -sparsification. *SIAM Journal on Scientific Computing*, 36(5):A2122–A2148, 2014.
- [2] A. Alexanderian, N. Petra, G. Stadler, and O. Ghattas. A fast and scalable method for a-optimal design of experiments for infinite-dimensional bayesian nonlinear inverse problems. *SIAM Journal on Scientific Computing*, 38(1):A243–A272, 2016.
- [3] Eduardo Casas, Christian Clason, and Karl Kunisch. Parabolic control problems in measure spaces with sparse solutions. *SIAM Journal on Control and Optimization*, 51(1):28–63, 2013.
- [4] Eduardo Casas and Karl Kunisch. Parabolic control problems in space-time measure spaces. *ESAIM: Control, Optimisation and Calculus of Variations*, 22(2):355–370, 2016.
- [5] J. De los Reyes. *Numerical PDE-constrained optimization*. Springer, 2015.
- [6] LC Evans. *Partial differential equations (graduate studies in mathematics 19 american mathematical society)*. 1998.
- [7] Eugenia Kalnay. *Atmospheric modeling, data assimilation and predictability*. Cambridge university press, 2003.
- [8] C. Kelley. *Iterative methods for optimization*, volume 18. Siam, 1999.
- [9] J. Lions. *Optimal control of systems governed by partial differential equations*, volume 170. Springer Verlag, 1971.
- [10] Jacques-Louis Lions. Pointwise control for distributed systems. *Control and Estimation in Distributed Parameter Systems, Frontiers in Applied Mathematics*, 11:1–39, 1992.
- [11] Ehrhard Raschke and Daniela Jacob. *Energy and water cycles in the climate system*, volume 5. Springer Science & Business Media, 2013.
- [12] JP Raymond. Nonlinear boundary control of semilinear parabolic problems with pointwise state constraints. *Discrete and Continuous Dynamical Systems*, 3:341–370, 1997.
- [13] Tomáš Roubíček. *Nonlinear partial differential equations with applications*, volume 153. Springer Science & Business Media, 2013.
- [14] F. Tröltzsch. Optimal control of partial differential equations. *Graduate studies in mathematics*, 112, 2010.
- [15] Thomas Tomkins Warner. *Numerical weather and climate prediction*. Cambridge University Press, 2010.