POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

Gestione ottima di un impianto di produzione di biogas: modelli di ottimizzazione con vincoli di incertezza

Laurea Magistrale in Automation and control Engineering - Ingegneria dell'automazione e del controllo

Author: Elvis Sheyler Jimenez Portillo

Advisor: Prof. Pietro Luigi Belotti

Co-advisor:

Academic year: 2021-2022

---

1. Introduzione

A fronte dei cambiamenti climatici e della ricerca di nuove forme di energia rinnovabili, il biogas si presenta come una nuova opportunità ed una soluzione ad entrambi i problemi. L'obiettivo di questo lavoro è lo studio e l'implementazione di un modello di impianto a digestione anaerobica per la produzione di biogas, con lo scopo di ottimizzare i costi relativi al funzionamento dell'impianto. Sono stati implementati modelli di ottimizzazione robusta per far fronte alla possibile presenza di parametri incerti presenti nell'impianto. Più modelli di incertezza sono stati usati con l'obiettivo di ottenere soluzioni che possano far fronte all'incertezza nel modo meno conservativo.

2. La struttura dell'impianto

Un impianto di biogas trasforma le biomasse in biogas, attraverso una serie di trasformazioni anaerobiche. In genere la struttura di un impianto di biogas è abbastanza standardizzata, e i suoi principali componenti sono: le aree di stoccaggio, il digestore, i gasometri, i sistemi di trattamento del biogas, la cogenerazione e le aree per lo stoccaggio del digestato. Il processo che porta alla formazione del biogas si chiama digestione anaerobica, che è, in estrema sintesi, un complesso processo biologico nel quale, in assenza di ossigeno, la sostanza organica viene trasformata in biogas. Dalla digestione anaerobica, oltre al biogas, si forma un altro prodotto, detto digestato, che è costituito dalla frazione organica non degradata ricca soprattutto di azoto. Il digestato è spesso usato come fertilizzante naturale. In questo lavoro è considerato un impianto di biogas da 1 MW. L'impianto è di dimensioni ridotte e lavora su scala locale. In questo caso le biomasse trattate per la produzione di biogas sono gli scarti della raccolta di mais e i liquami suini. Questi ultimi sono stati aggiunti come biomassa entrante nell'impianto per diverse ragioni. La prima è che apportano sostanze utili per il processo di produzione di biogas, mentre la seconda è che sono usati per portare la percentuale di sostanza secca sotto una certa soglia (in seguito vedremo più nel dettaglio cosa significa questo secondo punto). Le biomasse sono comprate e trasportate dalle rispettive fattorie/allevamenti, per poi essere preparate al processo all'interno dei digestori anaerobici. All'interno dei digestori, tramite il processo di digestione anaerobica, si formerà il biogas e il digestato che sarà venduto come concime.

3. Il modello

Innanzitutto è stato implementato un modello di impianto di biogas "ideale", ovvero senza prendere in considerazione nessun parametro incerto. In particolare è stato costruito un modello di ottimizzazione, perché come descritto in precedenza, l'obiettivo di questo lavoro è quello di ottenere una gestione ottimale dell'impianto, e quindi di massimizzarne il profitto. Un modello di ottimizzazione è un sistema composto da caratteristiche specifiche che sono gli obiettivi, le variabili, i parametri ed i vincoli. L'obiettivo di solito è definito da una funzione obiettivo. Le variabili sono controllate da noi e dipendono dalle relazioni definite all'interno del modello, mentre i parametri, che descrivono i dati del problema, sono predeterminate e non sono modificabili. Nel nostro modello le variabili considerate sono la quantità di mais e liquami acquistati, essenziali per il funzionamento dell'impianto, la quantità di digestato/concime venduto, ed il numero di viaggi necessari per il trasporto delle biomasse e del concime. Parlando dei vincoli invece, questi definiscono le condizioni necessarie che il modello deve rispettare per trovare le soluzioni ammissibili. In genere i vincoli sono rappresentati da equazioni o disequazioni. Nel nostro modello sono presenti diversi vincoli che possono essere divisi a seconda delle funzioni o caratteristiche che rappresentano. Per esempio si hanno i vincoli di produzione, che limitano la massima quantità di mais/liquami prodotti per una fattoria/allevamento; ci sono vincoli che mettono in relazione le biomasse e il biogas/concime, oppure i vincoli di relazione per il trasporto, che definiscono il legame tra le quantità di biomasse acquistate ed il numero di viaggi necessari al loro trasporto. A titolo dimostrativo rappresentiamo in forma matematica il vincolo che definisce il limite di sostanza secca presente nel digestore. Infatti va sottolineato che le biomasse non possono entrare nel digestore così come sono, ma vanno lavorate e trattate in precedenza in apposite aree di stoccaggio per prepararle al processo di digestione anaerobica. È importante che la percentuale di sostanza secca, che va ricordato essere la parte del campione del materiale che rimane dopo l'eliminazione d'acqua, non superi una certa soglia. Se la parte secca dovesse superare tale soglia il processo di digestione anaerobica non garantirebbe la produzione corretta di biogas e di digestato. La soglia totale di sostanza secca ammissibile nel digestore in questo caso non può essere più del 10%, mentre per il mais e per i liquami è rispettivamente al trenta e al cinque per cento (dati trovati in letteratura). Il vincolo che lo descrive è il seguente:

$$\sum_{j=1}^{J} X_j \; 0.3 + \sum_{k=1}^{K} Y_k \; 0.05 \leq 0.10 \; (\sum_{j=1}^{J} X_j + \sum_{k=1}^{K} Y_k)$$

dove J e K rappresentano rispettivamente il numero di fattorie e allevamenti considerati, mentre X e Y sono le variabili che rappresentano la quantità di mais e liquami acquistati.
La funzione obiettivo, come descritto in precedenza ha lo scopo di ottimizzare il profitto dell'impianto di biogas, e cioè di massimizzare la differenza tra i ricavi ed i costi complessivi. Sono considerati ricavi tutti i profitti che otteniamo dalla vendita di biogas e concime che vengono prodotti dall'impianto. I costi invece si riferiscono all'acquisto delle biomasse e ai costi di trasporto (delle biomasse e del concime).

3.1. L'incertezza nel modello

Dopo aver implementato il modello in assenza di incertezza, si passa a considerare l'incertezza che può essere presente nell'impianto. Ciò serve a rendere il modello più realistico a causa della natura aleatoria dei raccolti, ma richiede l'uso di un altro tipo di ottimizzazione, detta robusta, che verrà descritta in seguito. Ora le quantità di mais prodotte dalle fattorie (S_j^M) non sono più considerate costanti ma soggette ad incertezza. É molto probabile pensare infatti che nel corso dell'anno, a causa di diversi fattori, la produzione di mais non sia costante, ma soggetta ad incertezza. Fattori climatici ed ambientali possono verificarsi nel corso dell'anno e influenzare in maniera significativa la produzione di mais. Ovviamente il modello sarà soggetto a modifiche, la più importante delle quali è l'introduzione di una nuova variabile, denominata r, che rappresenta la quantità totale di mais importata nell'impianto, che adesso è incerta, e dipendente dalle X_j, variabile che viene ridefinita per rappresentare la percentuale di capacità di produzione della fattoria j. La variabile X_j è

stata cambiata di significato perché se fosse rimasta come quantità di mais comprato, come era nel modello iniziale, sarebbe diventata una variabile incerta in quanto la quantità di mais prodotta è incerta in questo modello. Perciò è diventato necessario introdurre una nuova variabile, ovvero r, che definendo la somma totale di mais importato nell'impianto, non è dipendente in maniera diretta da S^M. In questo modo si potrà costruire un modello robusto non troppo conservativo, in quanto avremmo nel sistema solo un vincolo in cui sarà presente il parametro incerto, e che determina anche il valore di r (tale vincolo sarà mostrato in seguito). L'importanza della variabile r nasce proprio dal fatto che mette in relazione il parametro incerto S^M con il nostro modello di ottimizzazione.
I vincoli implementati sono per lo più quelli definiti nel modello senza incertezza. Di particolare interesse però è il vincolo che definisce la relazione tra quantità totale di mais prodotto in ogni fattoria e la quantità totale di mais importata nell'impianto, descritto di seguito:

$$\sum_{j=1}^{J} S_j^M \; X_j \geq r, \tag{1}$$

Questo vincolo è fondamentale perché ci dice che r è limitata dall'alto dalla quantità totale di mais prodotto, inteso come somma dei prodotti tra le capacità incerte S_j^M e le proporzioni X_j. Si potrebbe dire che questo vincolo definisce il livello di incertezza del modello, perché al suo interno sono presenti tutte le variabili ed i parametri affetti da incertezza. É proprio da questo vincolo che si partirà per l'implementazione del modello di ottimizzazione robusta.
Riguardo alla funzione obiettivo invece, è sempre la stessa, con lo scopo di massimizzare il profitto dell'impianto.
Iniziamo ora a parlare del modello di ottimizzazione che useremo per trattare l'incertezza.

4. L'ottimizzazione robusta

L'ottimizzazione robusta è un paradigma per modellare problemi di ottimizzazione sotto incertezza. In un problema di questo tipo oltre ai parametri noti e alle variabili ci sono anche i parametri incerti, non noti, di cui non abbiamo il controllo. Questi parametri possono variare in un ben definito intervallo di incertezza. Tale intervallo dipende dal tipo di incertezza che si sta trattando e può essere definito da disequazioni e equazioni sia lineari che non lineari. Per ogni valore dei parametri incerti, i vincoli che definiscono l'intervallo di incertezza devono essere soddisfatti. Inoltre è importante sottolineare il fatto che bisogna cercare di definire un insieme/intervallo di incertezza non troppo grande ma abbastanza ristretto, in modo da evitare di avere un modello troppo conservativo.
L'obiettivo dell'ottimizzazione robusta è trovare una soluzione che sia ammissibile per tutti i possibili valori dei parametri incerti dentro l'insieme di incertezza, indipendentemente dalla realizzazione di tali parametri. Per spiegarla in maniera più semplice possibile, si può pensare che i parametri incerti siano decisi e controllati da un "avversario" [1]. Dopo che risolviamo il nostro problema di ottimizzazione, "l'avversario" vede la nostra soluzione e sceglie i valori dei parametri incerti che più possono danneggiarci. In pratica l'avversario risolve a sua volta un problema di ottimizzazione in cui i parametri incerti sono le sue variabili, e con lo scopo di trovare una soluzione che più danneggia la nostra (per esempio violando i nostri vincoli). Di conseguenza dovremo cercare di anticipare le mosse del nostro avversario. Quindi dobbiamo costruire una "controparte robusta" al modello di ottimizzazione, cioè incorporare il problema di ottimizzazione del nostro avversario al modello. In pratica dovremo risolvere un'altro problema di ottimizzazione, detto appunto la "controparte robusta", che avrà come variabili il nostro parametro incerto e come obiettivo avrà quello di trovare la soluzione che è più sfavorevole nel nostro caso. Di seguito un esempio di quanto detto:

$$\max \; 2x + 3y$$
$$1 \leq x \leq 7$$
$$2 \leq y \leq 9$$
$$x + y \geq 6$$
$$\max(ux + y - v : u \geq 0, v \geq 0, u + v \leq 1) \leq 5$$

dove x e y che sono le variabili, mentre u e v sono i parametri incerti. L'ultimo vincolo, quello in cui sono presenti i parametri incerti, rappresenta la controparte robusta, ovvero il problema di ottimizzazione risolto dal nostro avversario, che in questo caso, tramite l'operatore di massimizzazione cerca di violare il vincolo di minore e

uguale. Risolvendo il modello di ottimizzazione in questo modo però, può capitare di ottenere nel modello un vincolo non lineare che aumenta in maniera significativa la complessità del problema, come accade nell'esempio sopra riportato. Una soluzione a questa eventualità può essere trovata applicando il concetto di dualità.
Il duale di un problema è un concetto fondamentale nella programmazione lineare. Infatti ad ogni problema di programmazione lineare può essere associato un altro problema, detto appunto duale, che fornisce informazioni importanti relative alla soluzione. Ricorrendo alla dualità potremmo trasformare il problema non lineare in lineare, senza ricorrere a sforzi computazionali troppo complessi. In questo caso si potrebbe considerare il modello implementato dall'avversario come il problema primale, su cui poi sarà costruito il duale applicando le regole di corrispondenza primale/duale che possono essere trovate in un qualsiasi manuale di programmazione lineare. Infine sostituendo il problema duale al vincolo non lineare in cui sono presenti i parametri incerti si otterrà il modello di ottimizzazione robusta.
Nella sezione seguente mostreremo come si è implementato il modello di ottimizzazione robusta utilizzando l'approccio appena descritto.

5. Modello di ottimizzazione robusta

5.1. Incertezza poliedrale

In questo caso l'intervallo di incertezza in cui è definito il parametro incerto S_J^M è il seguente:

$$I = \{S_j^M, j \in J : S_j^M \in [SmLB, SmUB],$$
$$\textstyle\sum_{j=1}^J S_j^M \geq \sigma_{LB},$$
$$\textstyle\sum_{j=1}^J S_j^M \leq \sigma_{UB},$$
$$S_{j_1}^M \geq 0.9\ S_{j_2}^M,\ S_{j_2}^M \geq 0.9\ S_{j_1}^M\ \forall (j_1, j_2) \in E\}$$

L'intervallo di incertezza stabilisce che la produzione di mais per ogni fattoria è compresa in un certo intervallo, definito da $SmLB$ e $SmUB$, che sono rispettivamente i limiti inferiori e superiori di produzione di mais per ogni fattoria. σ_{LB} e σ_{UB} sono due parametri che rappresentano il limite inferiore e superiore di produzione totale di mais. Infine ci sono anche delle relazioni di vicinanza tra fattorie vicine, ovvero in una coppia di fattorie (j_1,j_2) comprese nell'insieme E, che definisce l'insieme delle coppie di fattorie vicine, la fattoria j_1 non può produrre più di un certa quantità rispetto alla fattoria j_2 e viceversa. L'insieme E sarà formato dalle coppie di fattorie che saranno distanti non più di 3 km. Le equazioni/disequazioni che rappresentano questo intervallo definiscono un poliedro di possibili valori di S_j^M.
Ora per costruire il modello robusto applichiamo il seguente ragionamento. Il parametro incerto S_j^M è presente solo nel vincolo rappresentato dall'equazione (1). Per fare in modo di violare questo vincolo l'opponente potrebbe decidere di costruire un modello di ottimizzazione dove S^M è la variabile che controlla, e la funzione obiettivo ha lo scopo di minimizzare la parte a sinistra del vincolo in questione, in questo modo:

$$\min(\sum_{j=1}^{J} S_j^M\ X_j) \geq\ r \tag{2}$$

Ovviamente i vincoli saranno quelli definiti nell'intervallo di incertezza. La controparte robusta sarà data sostituendo il modello appena descritta al vincolo (1). Così facendo però il modello da lineare diventa non lineare con tutte le problematiche che ne derivano, tra cui per esempio una maggiore difficoltà nella sua risoluzione. Ricordando però quanto detto in precedenza, ricorrendo alla dualità, potremo implementare un modello di ottimizzazione lineare ed ovviare a questo problema. Quindi dal problema primale, che sarebbe il modello implementato dall'opponente, si costruisce tramite le regole di corrispondenza primale/duale il problema duale. Infine sostituendo tale problema al vincolo (1) del modello originale si otterrà il modello di ottimizzazione robusta.

5.2. Budget uncertainty

Ora verrà testato un'altra classe di insiemi di incertezza per implementare il modello robusto, che prende il nome di "Budget uncertainty" o "Gamma uncertainty". Sono stati Bertsimas e Sim a proporre questo nuovo approccio nel 2004 [3], con lo scopo di ridurre il livello di conservatorismo dei modelli robusti implementati fino a quel momento. In breve l'incertezza di budget è una costante che controlla quanti parametri incerti possono deviare dai loro valori nominali. Per esempio consideriamo un insieme di parametri a_j appartenenti ad un insieme J, e di questi alcuni sono soggetti ad incertezza. I parametri incerti si discostano dal loro valore nominale in un determinato modo. Statisticamente è improbabile che tutti i parametri $a_j, j \in J$ siano soggetti ad incertezza, quindi si può introdurre un parametro θ che rappresenta il numero di parametri che possono scostarsi dal valore nominale. θ regola il numero di parametri incerti che possono essere considerati, e di conseguenza θ controlla il "trade-off" tra il livello di incertezza e i suoi effetti sulla funzione obiettivo del problema nominale. Questo trade-off viene definito come "price of robustness". In pratica regolando θ, si può controllare la robustezza del modello rispetto al livello di conservatorismo della soluzione.

Ritornando al nostro problema, costruiamo un nuovo insieme di incertezza che definisce il parametro incerto S^M, richiamando l'approccio di budget uncertainty definito in precedenza. L'insieme di incertezza è il seguente:

$$
\begin{aligned}
&I = \{S_j^M, j \in J : S_j^M \in [SmLB_j, SmUB_j],\\
&S_j^M \geq SmNOM_j\ (1 - \lambda b_j)\ \forall j \in J,\\
&\textstyle\sum_{j=1}^J b_j \leq \theta,\\
&b_j \geq 0\ \forall j \in J,\\
&b_j \leq 1\ \forall j \in J\}.
\end{aligned}
$$

L'insieme di incertezza definisce che la produzione di mais per ogni fattoria è compresa nell'intervallo definito da $SmLB_j$ e $SmUB_j$. Inoltre è introdotta una nuova variabile chiamata b, che determina il numero di parametri incerti presenti. La variabile b è definita nell'intervallo tra zero e uno. Se b_j è diverso da zero infatti significa che la produzione di mais S_j^M è diversa dal valore nominale. Il parametro θ regola il "price of robustness", ovvero controlla il trade-off tra il livello di incertezza del modello e i suoi effetti sulla funzione obiettivo. Se θ è uguale a zero la somma delle variabili b_j deve essere uguale a zero e quindi non ci sono fattorie con produzioni di mais incerte. Aumentando il valore di θ invece aumenta il limite superiore sulla somma dei valori di b_j, e quindi aumenta il livello di incertezza. Sono introdotte dei nuovi parametri, $SmNOM$ e λ, che definiscono rispettivamente la produzione nominale di mais per ogni fattoria e il decremento del valore di produzione nominale del mais, quando tale produzione è incerta. Per costruire il modello di ottimizzazione robusta, anche in questo caso, definiamo la controparte robusta del nostro problema di ottimizzazione. Il ragionamento è sempre lo stesso: dobbiamo pensare di avere un avversario che voglia in qualche modo danneggiarci. L'avversario controlla i parametri incerti, quindi cercherà di danneggiarci violando l'unico vincolo in cui è presente il parametro incerto S^M, ovvero la disuguaglianza (1), cercando di minimizzare la quantità a sinistra del segno. Come fatto in precedenza il problema di ottimizzazione che svolge il nostro avversario è definito come problema primale, con la differenza che i vincoli imposti in questo caso sono quelli definiti nel nuovo intervallo d'incertezza. Poi sempre con lo stesso procedimento è stato costruito il problema duale, ed infine sostituendolo all'equazione (1) del modello originale, si è ricavato il nuovo modello di ottimizzazione robusta.

5.3. Incertezza basata sugli scenari

Per ultimo trattiamo un tipo di incertezza basato sugli scenari. Gli scenari sono un modo efficiente di costruire problemi di ottimizzazione robusta usando dati storici di incertezze [2]. Supponiamo di avere un vettore di parametri incerti, ma di non conoscere un intervallo di incertezza che vada bene per il nostro modello. Se comunque avessimo a disposizione un database di registri contenenti il valore di ogni parametro incerto per un dato periodo di tempo passato, si potrebbe costruire un modello per cui uno o più vincoli siano soddisfatti per ogni registro del vettore d'incertezza. Quindi pur non avendo a disposizione un intervallo di incertezza ben definito, si potrebbe costruire un modello robusto considerando l'insieme di scenari come una sorta di insieme di incertezza. In questo caso verranno considerati i valori storici di produzione di mais in un certo range temporale(i dati non fanno riferimento ad un impianto reale ma saranno ricavati casualmente utilizzando delle librerie di Python). In pratica ora il parametro incerto S^M non è un vettore riga, ma una matrice di valori. Infatti S^M deve rappresentare i valori di produzione di mais per tutte le fattorie considerate nel corso degli anni. Di conseguenza ora l'equazione

(1) viene modificata in questo modo:

$$\sum_{j=1}^{J} S_{n,j}^{M}\ X_j \geq r,\ \forall n \in 1..N$$

con N che definisce il numero di anni passati presi in considerazione. È da notare che aumentando il numero di scenari, e quindi di anni passati presi in considerazione, si aumenta il livello di incertezza del problema e quindi si ottiene un modello più conservativo.

5.4. Soluzione

Nella tabella seguente sono mostrati i valori di biogas prodotti, in metri cubi, per i diversi modelli di ottimizzazione implementati. Mostriamo questi valori in quanto il biogas è la principale fonte di guadagno, e quindi il profitto del nostro impianto dipende in modo diretto da questi risultati. La quantità di biogas prodotta, senza incertezza, è ottenuta risolvendo il modello non robusto per $SmNOM$ dato. In questo caso il valore corrisponde a 6220k. Nella tabella seguente invece sono riportati i valori ottenuti per i modelli di ottimizzazione robusta implementati.

Δ	Poliedr.	θ	Budget	N	Scenari
0.99	6155k	1	6088k	5	6161k
0.95	5907k	2	5955k	10	5953k
0.90	5597k	3	5822k	20	5892k
0.75	5443k	5	5573k	40	5852k
0.5	5443k	10	5443k	80	5757k

Tabella 1: Quantità di biogas prodotto per i modelli implementati

I valori presenti in tabella sono stati ricavati partendo sempre da un vettore di $SmNOM$ dato e considerando un solo insieme di fattorie/allevamenti con le loro coordinate. Δ è un parametro che regola i valori di σ_{LB} e σ_{UB} nell'incertezza poliedrale. Più basso è questo valore, più l'intervallo definito da σ_{LB} e σ_{UB} si allarga e di conseguenza aumenta il livello di incertezza nel modello. Allo stesso modo aumentando il valore di θ e N per la budget uncertainty e per l'incertezza a scenari, aumentiamo il livello di incertezza nei rispettivi modelli. Ovviamente ad un aumento di incertezza corrispondono dei modelli più conservativi e di conseguenza delle soluzioni più stringenti in termini di valori, come è possibile osservare anche dalla tabella.

6. Conclusione

Dalla tabella possiamo trarre qualche conclusione sull'utilità del tipo di insiemi di incertezza scelti. La tabella ci mostra che aumentando il livello di incertezza in ogni modello assistiamo ad una diminuzione della quantità di biogas prodotta. Di fatto con l'aumento dell'incertezza i modelli diventano più conservativi e di conseguenza le soluzioni riportano dei valori più bassi. Questa diminuzione dei valori però è più o meno rapida a seconda del tipo di incertezza considerato.
In generale se si implementa la controparte robusta per due insiemi di incertezza diversi, si ottengono dei valori diversi. Se però un insieme di incertezza è un sottoinsieme di un altro, allora la sua controparte robusta otterrà dei valori maggiori della funzione obiettivo indipendentemente dal tipo di classi degli insiemi di incertezza(nel nostro caso poliedrale, di budget o a scenari). In definitiva si può dire che a seconda dei casi che possono capitare nella realtà, un insieme di incertezza può risultare più o meno utile. Per esempio si può pensare che non tutti gli impianti abbiano a disposizione dati storici, e quindi in quel caso non si potrebbe usare il modello di incertezza per scenari. Nel caso invece in cui si vorrebbe avere un controllo più stretto sul trade-off tra il livello di incertezza del modello e il valore della funzione obiettivo, si potrebbe usare la budget uncertainty, e così via.

7. Bibliografia

Riferimenti bibliografici

[1] Pietro Belotti, Zsolt Csizmadia, Susanne Heipcke, and Sebastien Lannez. Robust Optimization with FICO™Xpress. 2014.

[2] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization*, volume 28. Princeton university press, 2009.

[3] Dimitris Bertsimas and Melvyn Sim. The price of robustness. *Operations research*, 52(1):35–53, 2004.