

**Un modello di ottimizzazione per l'imputazione delle mancate risposte
statistiche nell'indagine sui trasporti marittimi dell'Istat**

Tiziana Pichiorri, Anna Maria Sgamba e Valerio Papale

Aprile 2004

Sommario

Il lavoro presenta l'utilizzo di metodi della ricerca operativa per il trattamento delle mancate risposte totali nei dati dell'indagine censuaria svolta dall'Istat sul trasporto marittimo.

Lo studio si pone come obiettivo la ricerca della configurazione ottima nell'attribuzione degli scostamenti nel flusso di natanti, tra porti di origine e destinazione, dovuti ad omissioni di compilazione dei modelli di rilevazione.

Il lavoro si sviluppa attraverso una fase di analisi teorica del problema finalizzata alla progettazione di due modelli matematici, il primo con funzione obiettivo lineare, il secondo con funzione obiettivo non lineare, atti a simulare il sistema in esame ed in una successiva fase di ricerca nella quale vengono presentati, e posti a confronto, i risultati ottenuti da una sperimentazione dei due modelli.

Viene inoltre presentata un'applicazione del modello lineare ai dati relativi alla rilevazione statistica dei Trasporti Marittimi con riferimento all'anno 2002.

La redazione dei capitoli e paragrafi è attribuibile come segue:

T. Pichiorri (DPTS/DCSC/SCO/E): Capitoli 1, 2 ,3, Appendice 1

A.M. Sgamba (DPTS/DCPC/PRE/C): Capitoli 1, 2 ,3, Appendice 1

V. Papale: Paragrafo 2.2.1, Appendice 2

INDICE

Introduzione

1. Analisi del problema

- 1.1 Il contesto di lavoro
- 1.2 L'indagine
- 1.3 I dati raccolti
- 1.4 Descrizione del problema
- 1.5 Le dimensioni del problema

2. I modelli proposti

- 2.1 Le assunzioni di base
- 2.2 Il modello matematico
 - 2.2.1 Formulazione mediante un modello non lineare
 - 2.2.2 Formulazione mediante un modello lineare

3. Sperimentazione dei modelli proposti e analisi di sensibilità

- 3.1 Applicazione e confronto dei modelli
 - 3.1.1 Sperimentazione del modello non lineare
 - 3.1.2 Sperimentazione del modello lineare
 - 3.1.3 Confronto dei risultati ottenuti
- 3.2 Applicazione ai dati del Trasporto Marittimo

Bibliografia

Appendice 1

Appendice 2

Introduzione

Il presente lavoro riguarda l'analisi di un reale problema decisionale aziendale e lo sviluppo di una sua possibile strategia risolutiva.

Il problema è quello relativo al trattamento delle mancate risposte, totali e parziali, nelle informazioni riportate in modelli di rilevazione statistica di dati sul trasporto marittimo, nell'ambito dell'indagine censuaria svolta dall'Istituto Nazionale di Statistica.

Come si vedrà più chiaramente in seguito, l'obiettivo è quello di muovere i primi passi verso la realizzazione di uno strumento di ricerca della configurazione ottima nell'attribuzione degli scostamenti nel flusso di natanti, merci e passeggeri tra porti di origine e destinazione, dovuti ad omissioni ed errori di compilazione dei modelli di rilevazione. In particolare, oggetto di analisi è il flusso di natanti tra ogni coppia di porti relativamente alla sola navigazione di cabotaggio.

Il lavoro di ricerca qui di seguito presentato è articolato in diverse sezioni logicamente distinte: una prima parte verrà dedicata allo studio teorico del problema cercando di descriverne caratteristiche, valenza gestionale e criticità di soluzione; la seconda parte propone invece un'analisi sperimentale, progettando ed implementando, anche attraverso delle esemplificazioni, due modelli matematici, il primo con f.o. lineare, il secondo con f.o. non lineare, atti a simulare il sistema in esame. Nell'ultima sezione vengono presentati, e posti a confronto, i risultati ottenuti da una sperimentazione dei due modelli precedentemente esposti. Infine, è presentata un'applicazione del modello lineare ai dati relativi alla rilevazione statistica dei Trasporti Marittimi con riferimento all'anno 2002.

CAPITOLO 1

ANALISI DEL PROBLEMA

1.1 - Il contesto di lavoro

La crescente domanda di informazione statistica nel settore dei trasporti marittimi è strettamente legata ad una serie di interventi recentemente attuati e finalizzati ad una razionalizzazione e ad un miglioramento del servizio, che hanno amplificato la rilevanza economica del settore.

In un contesto come quello italiano, il settore dei trasporti marittimi ha sempre avuto un contenuto di forte strategicità dovuto principalmente alla posizione geo-economica del paese nel Mediterraneo ed alla sua vocazione storica. Recentemente inoltre, intense azioni di governo attuate a livello nazionale ed europeo hanno fortemente rinnovato l'attenzione verso questo sistema di movimentazione definito sempre più *capital intensive*.

Le politiche di regolamentazione e sviluppo del trasporto marittimo sono state rivolte soprattutto alla promozione della navigazione a corto raggio ed al passaggio dal trasporto su strada a quello via mare, motivate da vantaggi in termini di costi, tempi e impatto ambientale.

Nell'ambito del crescente processo di integrazione europea, la Direttiva Comunitaria n. 64 dell'8/12/1995 regola la raccolta dei dati del trasporto marittimo nei paesi comunitari, rispondendo all'esigenza di disporre di un'accurata base informativa armonizzata a livello europeo che si ponga come strumento di supporto conoscitivo per una corretta pianificazione delle politiche dei trasporti nazionali e comunitari.

Dopo un periodo di transizione, concesso agli Stati membri per consentire di adeguare i sistemi statistici nazionali, la rilevazione statistica del trasporto marittimo effettuata in Italia risponde ai requisiti della Direttiva dal 2000.

1.2 - L'indagine

L'informazione statistica sul trasporto marittimo del paese è prodotta dall'Istituto Nazionale di Statistica attraverso un'indagine censuaria che ha per oggetto la

navigazione marittima a scopo di commercio. L'indagine rientra tra le rilevazioni di interesse pubblico inserite nel Programma Statistico Nazionale.

Il campo di osservazione è costituito dai movimenti di merci e di passeggeri per mezzo di navi contro pagamento di nolo o prezzo di passaggio, su rotte seguite totalmente o parzialmente per mare. L'unità di rilevazione è rappresentata dal natante mercantile (ovvero qualunque imbarcazione atta al trasporto marittimo) battente qualsiasi bandiera che arriva o parte da porti del territorio nazionale indipendentemente dalle sue caratteristiche dimensionali, funzionali e così via. Non rientrano nel campo d'applicazione della Direttiva, e sono pertanto escluse dalla rilevazione, le navi da pesca, le navi da trivellazione e da esplorazione, le navi-officina per il trattamento del pesce, le navi da guerra e le imbarcazioni utilizzate unicamente a fini non commerciali.

L'insieme di tutti gli arrivi e le partenze presso i porti italiani costituisce l'universo di eventi d'interesse per la rilevazione.

Organi di rilevazione sono gli Uffici doganali e gli Uffici marittimi: per ogni arrivo o partenza un'agenzia delegata è preposta alla compilazione della dichiarazione statistica di arrivo o di partenza (modelli I.501 e I.502) ed alla consegna delle dichiarazioni stesse alle autorità doganali. Inoltre l'Autorità marittima è tenuta a fornire alle Autorità doganali gli elenchi dei natanti mercantili arrivati o partiti (modelli 61A e 61P) dai porti posti nella propria giurisdizione. Lo scopo di questi modelli è quello di mettere in grado l'Autorità doganale di verificare che tutti gli aventi obbligo abbiano consegnato la dichiarazione di arrivo o partenza, e che ci sia coincidenza tra quanto dichiarato nei modelli 61 e la relativa dichiarazione statistica tramite i modelli I.501 e I.502. L'autorità doganale è dunque tenuta a verificare la completezza e l'esattezza delle informazioni riportate sulle dichiarazioni ed a provvedere, conseguentemente, al loro invio all'Istituto Nazionale di Statistica con periodicità mensile. Ogni natante viene dunque rilevato in ciascun porto unitamente al carico sbarcato o imbarcato tante volte quanti sono i suoi arrivi e le sue partenze durante il mese di rilevazione.

1.3 - I dati raccolti

I modelli di rilevazione sono articolati in sezioni distinte concernenti informazioni sul natante mercantile, informazioni sul carico ed elenchi di codifiche.

La prima sezione è relativa ad informazioni di tipo strutturale e giuridico del natante: riporta dati relativi alla tipologia (ad es. nave da crociera, petroliera, porta container ecc.), alla classe di grandezza (tonnellate di stazza netta e lorda), alla portata ed alla nazionalità di registrazione della nave (bandiera). Sono presenti inoltre informazioni sul viaggio, quali il porto dichiarante, la direzione del movimento (entrata o uscita), il primo porto di destinazione (o di provenienza), la data dell'evento ed informazioni sull'agenzia che ha curato la compilazione del modello.

La seconda sezione contiene informazioni sul carico di merci e passeggeri. Riporta dati sul peso lordo delle merci in tonnellate, la descrizione delle merci, il tipo di carico (rinfusa liquida, rinfusa solida, container, roro ecc.), il numero di passeggeri e il porto ed il paese di imbarco o sbarco a seconda che si tratti rispettivamente di una dichiarazione di arrivo o partenza.

La terza sezione contiene gli elenchi di codifiche. Riporta i codici dei porti e delle relative zone costiere marittime di appartenenza, le tipologie di natanti e delle merci. Queste ultime due classificazioni sono riportate ad un livello di disaggregazione maggiore rispetto a quello richiesto, al fine di facilitare l'attribuzione da parte del compilatore alla classe opportuna.

La classificazione merceologica adottata dall'Istat è la NSTR – Nomenclatura statistica dei trasporti – che codifica le merci in 10 capitoli e 24 gruppi. Il livello di disaggregazione della classificazione impiegato per la diffusione dei dati è il capitolo.

1.4 - Descrizione del problema

Come accennato in sede di introduzione, il problema oggetto di studio costituisce un caso particolare nell'ambito del noto ed importante problema statistico delle mancate risposte.

Il fenomeno delle mancate risposte è molto frequente in tutte le indagini di tipo sociale ed economico ed il suo verificarsi può introdurre significative distorsioni nel processo di costruzione del dato.

In generale, vengono definiti come “*errori di mancata risposta*” quegli errori dovuti al rifiuto o all'impossibilità a rispondere da parte delle unità statistiche contattate. Sono suddivisi in totali, se l'unità non partecipa affatto all'indagine, e parziali, quando l'unità non fornisce solo alcune particolari risposte.

Nel caso dell'indagine Istat sui trasporti marittimi, ci troviamo in presenza di entrambe le tipologie di errore poiché la mancata risposta nasce sia da un'omissione di compilazione dell'intero modello di rilevazione di arrivo e/o partenza, sia da una non corretta trascrizione dell'informazione contenuta in esso. Ciò è dovuto a molteplici cause che, nel caso di errore parziale, affliggono sia il dato di testo che quello numerico e che vanno dal banale errore di annotazione a quello dell'uso di una grafia poco comprensibile o scorretta, dall'improprio arrotondamento di una quantità all'uso di un codice errato.

La conseguenza di ciò si traduce in assenza di conformità, sia di tipo numerico che di contenuto, tra le dichiarazioni statistiche effettuate nel porto di partenza ed in quello di arrivo per ogni natante.

Più dettagliatamente: ogni natante viene rilevato in ciascun porto di partenza, unitamente al carico imbarcato, attraverso la dichiarazione statistica contenuta nel modello I.502. In questo caso, tale dichiarazione conterrà, tra le altre informazioni, quella relativa al porto di partenza e quella relativa al primo porto di destinazione. Questo ultimo sarà a sua volta il porto dichiarante che rileverà, per mezzo del modello I.501, ogni natante in arrivo unitamente al suo carico sbarcato. In questo caso, tale dichiarazione conterrà, tra le altre informazioni, quella relativa al porto di arrivo e quella relativa all'ultimo porto di provenienza. E' ovvio osservare che, per ogni viaggio del natante tra due porti, il numero delle dichiarazioni statistiche effettuate debba essere pari a due e che, il contenuto informativo (dati anagrafici, carico imbarcato, etc.) della dichiarazione di partenza, debba essere corrispondente con il contenuto informativo (dati anagrafici, carico sbarcato, etc.) della dichiarazione di arrivo. Accade invece, sovente, di riscontrare difformità tra le due registrazioni.

Se poi, come accade realmente nell'indagine, prolunghiamo nel tempo l'osservazione dei viaggi di un natante tra due porti ed estendiamo l'analisi all'insieme di tutti i porti nazionali che ospitano, in arrivo ed in partenza, qualsiasi natante che si sposta per fini commerciali è facilmente intuibile come l'errore possa assumere una dimensione importante.

Il problema che si pone ai responsabili dell'indagine è di tipo decisionale e richiede, nel caso di errore totale, di scegliere a quali porti attribuire le omissioni di dichiarazioni (viaggi) e, nel caso di errore parziale, di scegliere a quali porti (dichiarazioni statistiche di arrivo o partenza) attribuire i differenziali di flusso (merci e passeggeri) rilevati.

Attualmente, il monitoraggio della completezza e conformità dei dati ha luogo in vari stadi dell'indagine.

Nella fase iniziale di arrivo dei questionari in Istat una attenta revisione viene espletata da personale esperto, al quale è affidato il compito di verificare la correttezza delle informazioni riportate nei modelli di rilevazione, di riscontrare per quanto è possibile la presenza per ogni viaggio di entrambe le dichiarazioni di arrivo e partenza e di effettuare confronti con i dati presenti nei modelli 61A e 61P forniti dalle Capitanerie di Porto. Le situazioni più critiche in termini di errori di compilazione o di significative lacune nel numero di dichiarazioni pervenute vengono risolte tramite solleciti e verifiche telefoniche e via fax.

Successivamente alla registrazione su supporto informatico, i dati subiscono un processo di correzione e controllo mediante software specifici, attraverso i quali vengono risolti tutti gli errori dovuti alla digitazione, eventuali valori mancanti sfuggiti ai controlli precedenti e sono verificate tutte le relazioni logiche presenti nei dati tramite l'impostazione di regole formali e controlli visivi nelle situazioni più complesse. In ultimo, viene effettuata un'operazione di linkage dei natanti registrati con quelli presenti nel Registro Lloyds, sulla base del nome del natante, delle sue caratteristiche strutturali (dimensioni e tipologia) e della nazionalità di registrazione. A causa delle imprecisioni presenti nelle variabili coinvolte nel processo, l'abbinamento viene considerato valido se supera alcuni criteri soglia di accettazione.

Inoltre sono effettuati confronti rispetto ai dati dell'anno precedente e monitorato l'andamento mensile del flusso di viaggi e merci.

A valle del processo di correzione e controllo vengono calcolate statistiche "a specchio" per ogni singola coppia di porti, agendo con interventi mirati nelle situazioni di maggiore sbilanciamento. L'enorme numero di confronti infatti rende inattuabile in pratica la soluzione di ogni singolo caso (per i dati riferiti all'anno 2002 i collegamenti tra tutte le possibili coppie di porti, per ogni trimestre, per ogni classe dimensionale del natante, per ogni tipologia, per ogni nazionalità di registrazione

hanno dato origine a circa 16.400 confronti per i viaggi e ad oltre 50.000 confronti per le merci, relativamente alla sola navigazione di cabotaggio).

1.5 - Le dimensioni del problema

Il numero di porti italiani attualmente interessati alla rilevazione è di 126. Considerando che in ogni porto vengono compilate sia dichiarazioni di arrivo che di partenza, cioè ogni porto può essere considerato sia come porto di arrivo che come porto di partenza, avremmo una matrice di base di 126x126.

Ogni natante viene rilevato in ciascun porto tante volte quanti sono i suoi arrivi e le sue partenze durante il mese di rilevazione, dando origine ad un flusso medio di circa 13.000 modelli al mese, circa 160.000 ogni anno.

Le informazioni riportate su ciascun modello alimentano, a livello informatico, 8 file dati differenti, uno relativo alle caratteristiche del natante e del viaggio, corrispondente alla testata del modello di rilevazione, ed altri sette, uno per ogni sezione di cui è composto il modello, sono riservati alle informazioni sulle merci e sui passeggeri. Ogni modello quindi contribuisce con 1 record alla dimensione del file relativo alla testata e con un numero che va da 0 ad n record per ciascuno degli altri sette file.

La Direttiva Comunitaria cui è sottoposta l'indagine impone la trasmissione periodica dei dati ad Eurostat con cadenza trimestrale e con un livello di disaggregazione molto elevato.

Limitando l'analisi ai soli viaggi in navigazione di cabotaggio, tralasciando cioè tutto ciò che concerne il traffico internazionale ed il flusso di merce e passeggeri, e considerando che il livello di dettaglio temporale richiesto è il trimestre, la classificazione del tipo di nave prevede 10 categorie, la classificazione della grandezza delle navi è di 22 gruppi, le nazionalità di registrazione delle navi contemplate sono 248, avremo che la matrice di tutti i confronti possibili ha dimensioni $126 \times 126 \times 4 \times 10 \times 22 \times 248$ pari a 3.464.778.240 combinazioni.

Naturalmente un porto non ha collegamenti con tutti gli altri, né in ogni trimestre transitano navi appartenenti ad ogni categoria. Come già evidenziato in precedenza per i dati relativi al 2002 hanno dato origine a circa 16.400 confronti.

CAPITOLO 2

IL MODELLO PROPOSTO

2.1 - Le assunzioni di base

Come affermato in sede di introduzione, l'analisi e la modellizzazione del nostro sistema è finalizzata all'individuazione delle scelte che consentano di realizzare l'obiettivo di una gestione ottimale del problema delle mancate risposte, ai fini di un conseguente miglioramento della qualità dell'informazione statistica finale.

L'assunzione alla base della formulazione del problema è di avere dati già sottoposti ad un processo di correzione e controllo a livello di singolo record, perciò si esclude la presenza di errori nelle informazioni a disposizione. Limitiamo l'analisi alla sola navigazione di cabotaggio, per la quale soltanto infatti sono disponibili dichiarazioni statistiche di arrivo e partenza.

A partire da dati "puliti" si fisserà l'attenzione esclusivamente sul problema delle *mancate risposte*, errori non campionari che conducono a sottostime nella misurazione del fenomeno oggetto della rilevazione.

Nel caso in analisi, definiamo mancate risposte totali tutti quei viaggi, avvenuti nella realtà, che sono sfuggiti alla rilevazione, per i quali, quindi, non risulta compilata nessuna dichiarazione statistica di arrivo/partenza. Si è in presenza di mancate risposte parziali, se invece la dichiarazione statistica non è stata compilata completamente in tutte le sue parti. In entrambi i casi, avendo ipotizzato di disporre di dati già sottoposti ad un processo di correzione, la sottostima non può evidenziarsi da un'analisi, seppur attenta, della singola dichiarazione, ma solo mediante il confronto congiunto dei flussi in ingresso ed in uscita da un medesimo porto e tra coppie di porti.

Nel caso delle mancate risposte totali si rende necessario ricostruire tutte le informazioni riguardanti il viaggio, mentre per le mancate risposte parziali occorre ricostruire soltanto la parte di carico mancante. In questa sede verranno affrontati esclusivamente i problemi legati ai dati riguardanti i viaggi, cioè la stima delle mancate risposte totali, rinviando a sviluppi futuri la stima delle mancate risposte parziali.

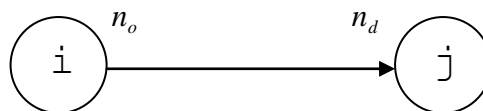
2.2 - Il modello matematico

La definizione di un modello matematico consente di rappresentare il problema dato per mezzo di relazioni tra variabili che esprimono i criteri secondo i quali confrontare diverse soluzioni ammissibili del problema allo scopo di scegliere quella migliore.

Dato l'insieme P dei porti coinvolti nell'indagine, ciascun porto p_l , ($l=1, \dots, n$) può essere analizzato sia dal punto di vista degli eventi "partenza" che degli eventi "arrivo" che si verificano nel periodo di rilevazione.

Denotando con O l'insieme dei porti di *origine* e con D l'insieme dei porti di *destinazione*, nel seguito indicheremo con o_i ($i=1, \dots, n$) o con d_j ($j=1, \dots, n$) i porti a seconda che siano luoghi origine o destinazione del viaggio.

La presenza di mancate risposte totali appare evidente confrontando il numero n_o delle dichiarazioni relative alle navi partite da un dato porto i verso un porto j , con il numero n_d delle dichiarazioni relative alle navi che, nello stesso periodo, sono arrivate in j provenendo da i . Nel caso migliore $n_o = n_d$.



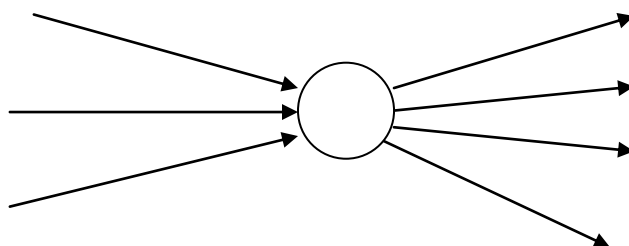
La disuguaglianza tra i valori n_o ed n_d indica esplicitamente un'omissione, da parte del porto con numerosità inferiore, nella trasmissione delle informazioni. In realtà, non è possibile escludere che anche i dati relativi al porto a cui corrisponde il valore più alto siano affetti da mancate risposte totali. Ad alcune delle n_o partenze, dichiarate da i , può non corrispondere nessuno degli n_d arrivi dichiarati da j e viceversa. Nel caso peggiore, non esiste nessuna coppia arrivo-partenza relativa allo stesso viaggio, il valore "vero" è dato da $n_o + n_d$.

Riportiamo di seguito una tabella di esemplificazione dei casi possibili.

Porto di origine (<i>i</i>)		Porto di destinazione (<i>j</i>)	
Nome nave	N. partenze	Nome nave	N. arrivi
Nina	10	Nina	14
Pinta	0	Pinta	5
Santa Maria	18	Santa Maria	13
Totale	28		32

In realtà, l'individuazione esatta della coppia arrivo-partenza relativa ad uno specifico viaggio effettuato da un natante è impossibile a causa dell'attuale configurazione dei modelli di rilevazione che prevedono l'informazione relativa alla data dell'evento, ma non quella relativa all'orario o altro codice che ne consenta l'abbinamento univoco. In più, per i natanti che effettuano più volte nel corso del giorno la stessa linea, è prevista la compilazione di modelli riepilogativi mensili, rendendo di fatto inattuabile l'individuazione delle singole coppie.

Inoltre, in generale è possibile affermare che, in un dato intervallo di tempo, il numero di navi che arrivano in un porto, indipendentemente dalla loro provenienza, dovrebbe essere sufficientemente "vicino" al numero delle navi che partono dal medesimo porto nello stesso periodo, qualunque sia la loro destinazione. Infatti, se il numero degli arrivi superasse di gran lunga quello delle partenze, si potrebbe ipotizzare una situazione in cui moltissime navi rimangono per periodi prolungati attraccate alla banchina del molo, condizione assai improbabile trattandosi di natanti utilizzati per operazioni commerciali. Nel caso opposto invece, si avrebbe il paradosso di navi che partono da un porto, pur non essendovi mai arrivate.



2.2.1 - Formulazione mediante un modello non lineare

In questa sezione verrà esposta una formulazione “intuitiva” del problema in esame rinviando invece, al paragrafo successivo, la possibilità di lettura dello stesso all’interno di una classe di esempi classici di applicazione della ricerca operativa quale quella dei problemi di flusso a costo minimo.

In base alle considerazioni precedentemente esposte, il caso di studio può essere formalizzato mediante un modello non lineare di ottimizzazione vincolata.

Indichiamo con:

$N_o(i, j)$ = Numero di partenze dichiarate dal porto di origine i verso il porto di destinazione j .

$N_d(i, j)$ = Numero di arrivi dal porto di origine i , dichiarati dal porto di destinazione j .

$M_o(i, j)$ = Numero di partenze omesse dal porto di origine i verso il porto di destinazione j .

$M_d(i, j)$ = Numero di arrivi dal porto di origine i , omessi dal porto di destinazione j .

$ct(i, j)$ = Peso associato alla tratta (i, j)

$cp(l)$ = Peso associato al porto l .

L’insieme delle soluzioni ammissibili sarà rappresentato da tutte le coppie, M_o ed M_d , di matrici quadrate di ordine $n \times n$ i cui elementi generici $M_o(i, j)$ e $M_d(i, j)$ rappresentano rispettivamente il numero di mancate dichiarazioni di partenza e arrivo sulla tratta (i, j) , che realizzano il pareggiamento dei flussi di natanti su ogni singola tratta (vincolo 1) soddisfacendo, contemporaneamente, i vincoli di capacità (vincoli 2 e 3). Questi ultimi tengono conto delle considerazioni sopra esposte, relativamente al caso migliore e peggiore nel numero di corrispondenze partenza-arrivo presenti tra due porti. La soluzione ottima assicura la minimizzazione dello scostamento tra il flusso in ingresso ed il flusso in uscita di natanti da ogni singolo porto, utilizzando il numero minimo di mancate dichiarazioni di partenza e di arrivo su ogni singola tratta.

$$\begin{aligned} \text{Min} \quad & \sum_l \left(cp(l) \times \left(\sum_i (N_d(i,l) + M_d(i,l)) - \sum_j (N_o(l,j) + M_o(l,j)) \right) \right)^2 + \\ & + \sum_i \sum_j (ct(i,j) \times M_o(i,j))^2 + \sum_i \sum_j (ct(i,j) \times M_d(i,j))^2 \end{aligned}$$

$$1) N_o(i,j) + M_o(i,j) = N_d(i,j) + M_d(i,j) \quad \forall i,j$$

$$2) 0 \leq M_o(i,j) \leq N_d(i,j) \quad \forall i,j \quad \text{intere}$$

$$3) 0 \leq M_d(i,j) \leq N_o(i,j) \quad \forall i,j \quad \text{intere}$$

con:

$$cp(l) = \begin{cases} 2 \cdot \max \left\{ \sum_i N_d(i,l), \sum_j N_o(l,j) \right\} & \text{se } \min \left\{ \sum_i N_d(i,l), \sum_j N_o(l,j) \right\} = 0 \\ 2 \cdot \frac{\max \left\{ \sum_i N_d(i,l), \sum_j N_o(l,j) \right\}}{\min \left\{ \sum_i N_d(i,l), \sum_j N_o(l,j) \right\}} & \text{altrimenti} \end{cases} \quad \forall l$$

$$e \quad ct(i,j) = \begin{cases} \frac{\min \{N_o(i,j), N_d(i,j)\}}{\max \{N_o(i,j), N_d(i,j)\}} & \text{se } \frac{\min \{N_o(i,j), N_d(i,j)\}}{\max \{N_o(i,j), N_d(i,j)\}} \geq 0,01 \\ 0,01 & \text{altrimenti} \end{cases} \quad \forall i,j$$

La struttura della funzione obiettivo è frutto di scelte dettate dalle seguenti considerazioni:

- tra tutte le soluzioni ammissibili che minimizzano lo scostamento tra il flusso in ingresso ed in uscita di natanti da ogni porto, richiediamo di ricercare, adottando un comportamento di prudenza, quella che contestualmente assicuri il valore minimo delle mancate risposte.

- Una tratta (i,j) per la quale i dati registrati forniscono valori prossimi di $N_o(i,j)$ ed $N_d(i,j)$ necessiterà di “poche unità” per il soddisfacimento della condizione di pareggiamento del flusso sulla tratta, quindi i valori $M_o(i,j)$ ed $M_d(i,j)$ che ne derivano sono entrambi molto piccoli e contribuiscono in maniera estremamente ridotta al “peso” complessivo della funzione obiettivo, ma sono anche una possibile indicazione di un buon livello della qualità del dato rilevato. La minimizzazione del flusso in ingresso ed in uscita di natanti da ogni porto comporta un’ulteriore allocazione di viaggi tra le singole tratte. Poiché i singoli termini della funzione obiettivo sono elevati al quadrato, aumentando valori molto piccoli di $M_o(i,j)$ ed $M_d(i,j)$ si hanno incrementi della f.o. di modesta entità, rispetto all’incremento, di pari unità, di valori $M_o(i,j)$ ed $M_d(i,j)$ più elevati e questo induce a “preferirli” nell’attribuzione dei viaggi mancanti. Per ristabilire un equilibrio, si associano dei coefficienti di peso $ct(i,j)$, inversamente proporzionali al differenziale sulla tratta, che assumono valore pari a 1 nel caso in cui $N_d(i,j) = N_o(i,j)$ e decrescono progressivamente all’aumentare della differenza tra le due quantità, fino ad una soglia minima fissata a 0,01.
- I pesi $cp(l)$ associati ai porti, esaltano scostamenti maggiori nei vincoli di conservazione del flusso, variando da un minimo di 2 nel caso di perfetto bilanciamento tra i flussi in ingresso ed in uscita ad un massimo fortemente dipendente dai dati di input.

Da un punto di vista puramente matematico, si tratta di determinare il minimo di una funzione obiettivo quadratica in presenza di vincoli lineari di eguaglianza e disuguaglianza. E’ possibile immediatamente notare che la $f(x)$ assume solamente valori non negativi, in quanto somma di quantità elevate al quadrato, e pertanto risulta limitata inferiormente dallo zero.

La presenza di espressioni non lineari in un modello di programmazione matematica è piuttosto frequente, tuttavia i modelli non lineari sono di solito molto più difficili da risolvere e quindi spesso si cerca di approssimarli con modelli lineari. I vantaggi che ne derivano sono quelli di generalità e flessibilità, di semplicità e di efficienza degli algoritmi risolutivi ed in ultimo di possibilità di analisi qualitative.

2.2.2 - Formulazione mediante un modello lineare

Un gran numero di problemi reali, negli ambiti più diversi, si modellizzano efficacemente come problemi di flusso su rete. La pianificazione della produzione e della distribuzione, la localizzazione di impianti di produzione e stoccaggio, lo studio di reti di comunicazione e di trasporto, sono solo alcuni esempi di problemi applicativi caratterizzati da una struttura di grafo che si manifesta in modo del tutto naturale.

Tra i diversi problemi di flusso su rete, il problema del flusso a costo minimo (MCF *da Min Cost Flow problem*) rappresenta un noto ed importante modello la cui generalità e flessibilità d'uso ne consente una vasta applicazione a molti problemi reali. La descrizione del problema di MCF risulta di immediata comprensione: indichiamo con $G = (N, A)$ un grafo orientato definito da un insieme N di nodi e da un insieme A di archi orientati (in generale, in una rete gli archi sono interpretabili come canali attraverso cui fluiscono beni rappresentati per mezzo di grandezze discrete o continue). Assumiamo inoltre che siano associati pesi sia agli archi che ai nodi del grafo. I pesi degli archi rappresentano usualmente distanze, costi o tempi unitari relativi al trasferimento da un nodo all'altro; in particolare, ad ogni arco (i,j) sono associati un valore c_{ij} , che indica il costo per ogni unità del bene che percorre l'arco, una capacità inferiore l_{ij} e superiore u_{ij} che indicano, rispettivamente, il minimo ed il massimo numero di unità di bene ammissibili sull'arco. I pesi associati ai nodi rappresentano la domanda o l'offerta del bene in quei nodi. Più precisamente, ad ogni nodo $i \in N$ è associato un valore reale $b(i)$, che rappresenta l'offerta se $b(i) > 0$, la domanda se $b(i) < 0$. Nel primo caso il nodo i è un nodo di offerta del bene con offerta pari a $b(i)$, nel secondo caso il nodo i è un nodo di domanda del bene, con domanda pari a $-b(i)$. Se $b(i) = 0$ il nodo viene detto di trasferimento. Quindi, per ogni nodo di trasferimento, la somma dei flussi degli archi entranti, aumentata della produzione nel nodo e diminuita della domanda dello stesso, risulta uguale alla somma dei flussi degli archi uscenti dal nodo. Inoltre la domanda globale deve eguagliare l'offerta globale.

Le variabili decisionali rappresentano la quantità di flusso che deve passare su ciascun arco della rete, in modo da soddisfare i vincoli di capacità e i vincoli di conservazione del flusso, al minimo costo.

Più formalmente, indicata con x_{ij} la variabile decisionale rappresentante il flusso sull'arco (i,j) , il problema del flusso di costo minimo si rappresenta come segue:

$$\min \sum_{(i,j) \in A} c_{ij} x_{ij}$$

$$(1) \quad \sum_{j:(i,j) \in A} x_{ij} - \sum_{j:(j,i) \in A} x_{ji} = b(i) \quad \forall i \in N$$

$$(2) \quad l_{ij} \leq x_{ij} \leq u_{ij} \quad \forall (i,j) \in A$$

$$\text{con } \sum_{i \in N} b(i) = 0.$$

Il vincolo (1), noto come vincolo di conservazione del flusso, afferma che per ogni nodo i , la somma dei flussi degli archi uscenti (flusso totale in uscita), meno la somma dei flussi degli archi entranti (flusso totale in entrata) sia uguale alla offerta/domanda $b(i)$, mentre il vincolo (2) va sotto il nome di vincolo di capacità (in molte applicazioni la capacità inferiore viene assunta uguale a 0, e quindi viene fornita tra i parametri della rete solamente la capacità superiore).

Nell'ambito del trasporto marittimo, l'insieme dei porti e dei flussi di natanti tra essi possono essere efficacemente rappresentati mediante una rete orientata $G=(P, T)$ nella quale ciascun porto è rappresentato da un nodo dell'insieme P e ciascuna tratta da un arco orientato di T . Ad ogni arco (i,j) è possibile associare un costo per unità di flusso pari a $ct(i,j)$, capacità inferiore pari a 0 e capacità superiore $u(i,j) = N_o(i,j) + N_d(i,j)$. I pesi attribuiti ai singoli porti saranno proporzionali al differenziale tra il numero di natanti arrivati ed il numero di natanti partiti da ciascun porto nell'intervallo di tempo considerato.

Limitando l'analisi alla sola navigazione di cabotaggio, è ragionevole supporre rispettata l'assunzione di base che la domanda globale risulti uguale all'offerta globale, infatti, poiché ad ogni partenza corrisponde un arrivo, il numero complessivo delle partenze sarà uguale alla somma di tutti gli arrivi.

In generale, invece, non ci sono elementi certi per poter partizionare l'insieme dei porti in porti di domanda, di offerta o di trasferimento. In assenza di particolari

caratteristiche strutturali ogni porto dovrebbe comportarsi come un nodo di trasferimento. Nella pratica però, anche se il periodo di osservazione è abbastanza ampio, non è possibile escludere che alcuni natanti possano fermarsi più a lungo nel porto a causa di lavori di riparazione, condizioni metereologiche avverse, ritardi accumulati nelle operazioni di carico e scarico delle merci ecc. Inoltre, più sporadicamente, può avvenire il varo di natanti nuovi o la dismissione di quelli obsoleti. L'insieme di questi fattori, imprevedibili e di varia natura, comporta la possibilità di riscontrare piccoli scostamenti in eccesso o in difetto nel numero di natanti in ingresso ed in uscita di cui non si è in grado di valutare l'entità.

Il modello può essere formulato come un problema di programmazione lineare per mete multiple. Infatti, l'obiettivo desiderabile del pareggiamento tra i flussi in ingresso ed in uscita per ciascun porto, può essere espresso per mezzo di n mete soft mediante l'introduzione di due vettori $\alpha(l)$ e $\beta(l)$, le cui componenti rappresentano rispettivamente gli scostamenti positivi o negativi che realizzano il contestuale raggiungimento degli n valori target tutti pari a 0. Formuliamo quindi un problema di minimo vincolato associato al problema di programmazione per mete soffici mediante la tecnica del Goal Programming:

$$\min \sum_{(i,j) \in T} ct(i,j) \cdot (M_o(i,j) + M_d(i,j)) + \sum_{l \in P} cp(l) \cdot (\alpha(l) + \beta(l))$$

$$\sum_j (N_o(l,j) + M_o(l,j)) - \sum_i (N_d(i,l) + M_d(i,l)) + \alpha(l) - \beta(l) = 0 \quad \forall l$$

$$N_o(i,j) + M_o(i,j) = N_d(i,j) + M_d(i,j) \quad \forall (i,j) \in T$$

$$0 \leq M_o(i,j) \leq N_d(i,j) \quad \textit{intere} \quad \forall (i,j) \in T$$

$$0 \leq M_d(i,j) \leq N_o(i,j) \quad \textit{intere} \quad \forall (i,j) \in T$$

$$\alpha(l), \beta(l) \geq 0 \quad \textit{intere} \quad \forall l$$

in cui $cp(l)$ e $ct(i,j)$ sono descritti mediante le stesse espressioni fornite in precedenza (§ 2.2.1).

Il modello può essere ulteriormente articolato considerando l'insieme degli arrivi (o delle partenze) non come un unico aggregato, ma decomponendolo sia in base alla

periodicità temporale, il trimestre, sia in base alla classificazione del natante per tipologia e classe dimensionale:

$$\min \sum_t \sum_n \sum_d \left(\sum_{(i,j) \in T} ct^{t,n,d}(i,j) \cdot (M_o^{t,n,d}(i,j) + M_d^{t,n,d}(i,j)) + \sum_{l \in P} cp^{t,n,d}(l) \cdot (\alpha^{t,n,d}(l) + \beta^{t,n,d}(l)) \right)$$

$$\sum_j (N_o^{t,n,d}(l,j) + M_o^{t,n,d}(l,j)) - \sum_i (N_d^{t,n,d}(i,l) + M_d^{t,n,d}(i,l)) + \alpha^{t,n,d}(l) - \beta^{t,n,d}(l) = 0 \quad \forall l, n, t, d$$

$$N_o^{t,n,d}(i,j) + M_o^{t,n,d}(i,j) = N_d^{t,n,d}(i,j) + M_d^{t,n,d}(i,j) \quad \forall (i,j) \in T \text{ e } \forall t, n, d$$

$$0 \leq M_o^{t,n,d}(i,j) \leq N_d^{t,n,d}(i,j) \quad \textit{intere} \quad \forall (i,j) \in T \text{ e } \forall t, n, d$$

$$0 \leq M_d^{t,n,d}(i,j) \leq N_o^{t,n,d}(i,j) \quad \textit{intere} \quad \forall (i,j) \in T \text{ e } \forall t, n, d$$

$$\alpha^{t,n,d}(l), \beta^{t,n,d}(l) \geq 0 \quad \textit{intere} \quad \forall l \text{ e } \forall t, n, d$$

Questo modello, pur rimanendo sostanzialmente invariato rispetto al precedente nella formulazione e nell'interpretazione delle variabili, mediante il partizionamento del flusso di natanti in sottoinsiemi omogenei consente di ottenere una stima più accurata delle mancate risposte e quindi del flusso "reale".

CAPITOLO 3

SPERIMENTAZIONE DEI MODELLI PROPOSTI

3.1 - Applicazione e confronto dei modelli

In questa sezione vengono presentati, e posti a confronto, i risultati ottenuti da una sperimentazione dei due modelli precedentemente esposti.

Il test è stato eseguito, a scopo esemplificativo, utilizzando una struttura di dati fittizi (Tabella 3.1) applicata ad un insieme di porti, le relazioni tra i quali sono rappresentate dal grafo qui di seguito riportato.

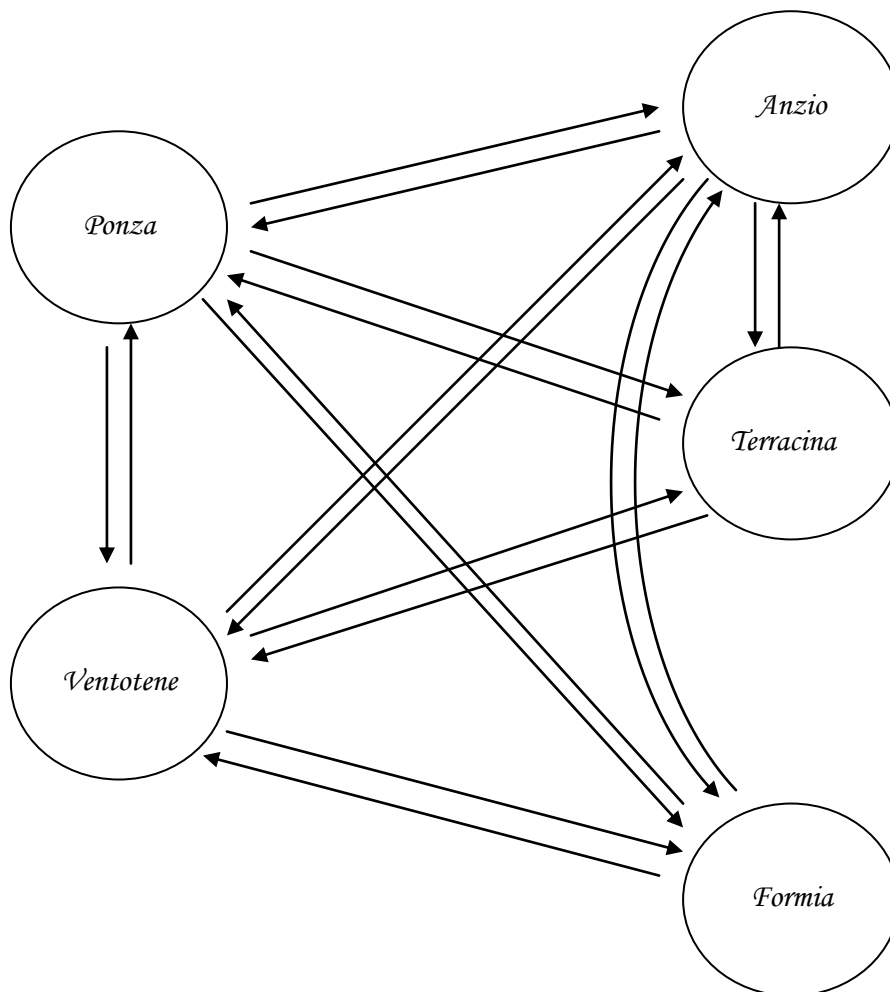


Tabella 3.1 – Numero di partenze e di arrivi registrati per porto di origine e destinazione e coefficienti di costo associati alle tratte.

Porto di origine	N_o	Porto di destinazione	N_d	$ct(i,j)$
Anzio	0	Formia	89	0,01
Anzio	870	Ponza	1.250	0,70
Anzio	50	Terracina	65	0,77
Anzio	1.760	Ventotene	1.756	1,00
Formia	3	Anzio	0	0,01
Formia	750	Ponza	750	1,00
Formia	40	Ventotene	40	1,00
Ponza	1.500	Anzio	1.500	1,00
Ponza	800	Formia	821	0,97
Ponza	200	Terracina	190	0,95
Ponza	103	Ventotene	20	0,19
Terracina	50	Anzio	10	0,20
Terracina	0	Ponza	222	0,01
Terracina	65	Ventotene	2	0,03
Ventotene	1.470	Anzio	1.200	0,82
Ventotene	68	Formia	70	0,97
Ventotene	196	Ponza	100	0,51
Ventotene	80	Terracina	88	0,91
Totale	8.005	Totale	8.173	

Come è possibile rilevare dalla tabella precedente, il data set su cui effettuare la sperimentazione dei modelli è stato immaginato, considerando la possibile difformità di casi analizzabili: tratte caratterizzate da totale assenza di dichiarazioni di arrivo ($N_o = 0$) o di partenza ($N_d = 0$), tratte caratterizzate da pareggi su valori numerici bassi e su valori numerici elevati di N_o ed N_d , tratte caratterizzate da scostamenti di diversa ampiezza tra N_o ed N_d . Verosimilmente, alcuni casi possono apparire poco realistici, ma la loro inclusione è stata intenzionalmente considerata al fine di osservare il comportamento del modello su un caso di sperimentazione il più possibile variegato.

Per quanto riguarda i pesi applicati alle tratte, è possibile osservare come il valore massimo, pari ad 1,00, sia attribuito a tutte le tratte (i,j) per le quali si realizza l'uguaglianza tra il numero delle dichiarazioni di arrivo ed il numero delle dichiarazioni di partenza, indipendentemente dalla loro numerosità. In tutti gli altri casi, i coefficienti $ct(i,j)$ assumono valori compresi tra 1,00 e 0,01, (soglia minima prefissata) decrescendo all'aumentare della differenza percentuale tra le due quantità N_o ed N_d . Una stessa differenza in valore assoluto infatti può avere un significato molto

diverso a seconda dell'ordine di grandezza delle quantità a cui va riferito. Inoltre, si può evidenziare come il sistema di pesi adottato attribuisca la stessa importanza a squilibri numerici di pari entità in favore degli arrivi o delle partenze.

Nella Tabella 3.2 sono riportati, per ciascun porto, il totale degli arrivi e delle partenze dichiarate, lo scostamento in valore assoluto e percentuale tra queste due quantità ed i valori dei pesi risultanti dall'applicazione della formula determinata in precedenza (§2.1.1).

Tabella 3.2 – Complesso degli arrivi e delle partenze per ciascun porto, scostamenti in termini assoluti e percentuali e coefficienti di costo.

Porto	Arrivi	Partenze	Scostamento	Scostamento %	cp
Anzio	2.710	2.680	30	1,1	2,02
Formia	980	793	187	23,6	2,47
Ponza	2.322	2.603	-281	-10,8	2,24
Terracina	343	115	228	198,3	5,97
Ventotene	1.818	1.814	4	0,2	2,00
Totale	8.173	8.005	168	-2,1	

I pesi *cp* assumono valore minimo pari a 2,00, nel caso di bilanciamento tra i flussi in ingresso ed i flussi in uscita da ogni porto, ed aumentano progressivamente all'aumentare dello scostamento percentuale in valore assoluto. Infatti, come per i pesi assegnati alle tratte, anche i coefficienti *cp* non operano distinzioni tra casi in cui la numerosità degli arrivi eccede quella delle partenze e viceversa.

Nell'esempio presentato si è cercato di riproporre una situazione che comprendesse quanto più possibile le varie casistiche ipotizzabili.

3.1.1 - Sperimentazione del modello non lineare

L'applicazione del modello non lineare è stata effettuata utilizzando il software di modellazione AMPL. Il modellatore AMPL permette di risolvere problemi di programmazione matematica usando diversi risolutori, quali il solver MINOS specifico per problemi di programmazione non lineare ed il solver CPLEX, generalmente indicato per la soluzione di problemi di programmazione lineare e di programmazione lineare intera o mista, che possiede, tra le diverse opzioni, un'estensione dell'algoritmo Barrier per la soluzione di una classe particolare di problemi non lineari: i problemi di programmazione quadratica (QPs).

Cplex automaticamente riconosce un problema di programmazione quadratica attraverso la verifica del soddisfacimento delle seguenti condizioni: a) tutti i vincoli devono essere lineari, b) la funzione obiettivo deve potersi esprimere come somma di termini ognuno dei quali è una espressione lineare o il prodotto di due espressioni lineari, c) semidefinita positività (in caso di minimizzazione) o semidefinita negatività (in caso di massimizzazione) della matrice Q dei coefficienti quadratici della funzione obiettivo.

Il data set sperimentale ed il codice, mediante il quale è stato descritto il modello, sono stati sottoposti al software utilizzando ambedue i solver. I risultati forniti nei due casi sono risultati uguali.

In realtà, i valori di M_o e M_d trovati sono risultati essere soluzione del problema rilassato, ottenuto cioè eliminando i vincoli di interezza sulle variabili. In generale la soluzione del problema rilassato costituisce solo un'approssimazione del valore ottimo del problema di programmazione intera, ma in alcuni casi è possibile, a partire dalla soluzione del problema rilassato, determinare una soluzione intera ammissibile approssimando all'intero più vicino.

Il valore ottimo della funzione obiettivo è pari a 209.452. I valori delle variabili decisionali M_o e M_d trovati sono riportati in Tabella 3.3.

Tabella 3.3 – Numero di partenze e di arrivi, mancate risposte in partenza ed in arrivo e flusso finale per porto di origine e destinazione.

Porto di origine	No	Mo	Flusso finale sull'arco	Md	Nd	Porto di destinazione
Anzio	0	89	89	0	89	Formia
Anzio	870	380	1.250	0	1.250	Ponza
Anzio	50	15	65	0	65	Terracina
Anzio	1.760	0	1.760	4	1.756	Ventotene
Formia	3	0	3	3	0	Anzio
Formia	750	126	876	126	750	Ponza
Formia	40	34	74	34	40	Ventotene
Ponza	1.500	37	1.537	37	1.500	Anzio
Ponza	800	21	821	0	821	Formia
Ponza	200	0	200	10	190	Terracina
Ponza	103	0	103	83	20	Ventotene
Terracina	50	10	60	50	10	Anzio
Terracina	0	222	222	0	222	Ponza
Terracina	65	2	67	65	2	Ventotene
Ventotene	1.470	56	1.526	326	1.200	Anzio
Ventotene	68	2	70	0	70	Formia
Ventotene	196	100	296	196	100	Ponza
Ventotene	80	8	88	0	88	Terracina
Totale	8.005	1.102	9.107	934	8.173	Totale

Esaminando i risultati ottenuti è possibile rilevare come, per ogni tratta, il rispetto dei vincoli conduce ad uno stato di pareggiamento dei viaggi realizzato distribuendo unità “decisionali” in modo tale da:

- aumentare il numero delle dichiarazioni in tutte le tratte inizialmente in equilibrio in misura inversamente proporzionale alla numerosità iniziale;
- riportare sempre il porto di partenza/destinazione, che presenta dichiarazioni di viaggio totalmente mancanti, al valore registrato dal porto di destinazione/partenza della tratta corrispondente;
- in alcune occasioni, aumentare il numero delle dichiarazioni del porto di partenza/ destinazione con numerosità inferiore di registrazioni, fino al valore registrato dal porto di destinazione/partenza della tratta corrispondente;
- in altri casi invece, aumentare il numero delle dichiarazioni di entrambi i porti che hanno una diversa numerosità di registrazioni, quasi fino al massimo consentito (dato dalla somma delle registrazioni dei due porti).

Tabella 3.4 – Scostamento iniziale e finale tra flusso in ingresso e uscita di natanti per porto.

Porto	Arrivi iniziali Nd	Partenze iniziali No	Arrivi finali Nd+Md	Partenze finali No+Mo	Differenze iniziali	Differenze finali
Anzio	2.710	2.680	3.126	3.164	30	-38
Formia	980	793	980	953	187	27
Ponza	2.322	2.603	2.644	2.661	-281	-17
Terracina	343	115	353	349	228	4
Ventotene	1.818	1.814	2.004	1.980	4	24
Totale	8.173	8.005	9.107	9.107	168	0

Passando ad esaminare la Tabella 3.4, inoltre, è possibile rilevare come il bilanciamento di tutte le tratte avvenga riducendo la differenza tra natanti in ingresso ed in uscita per i porti di Formia, Ponza e Terracina ma non per i porti di Anzio e Ventotene. Il modello ha mostrato dunque di non assicurare, per ogni singolo porto, la minimizzazione dello scostamento tra il flusso in ingresso ed in uscita di natanti.

3.1.2 - Sperimentazione del modello lineare

Il modello lineare è stato implementato utilizzando il modulo OR (Operational Reserch) di SAS. Il software SAS/OR mette a disposizione diverse procedure per permettere di risolvere una grande varietà di problemi di ottimizzazione, classificati come problemi di programmazione lineare e non lineare. Nell'ambito dei primi, tra le diverse alternative proposte dal SAS/OR è stata selezionata la PROC LP, procedura che consente di risolvere problemi di programmazione lineare e di programmazione lineare intera e mista.

In generale, per la soluzione di problemi di programmazione lineare la PROC LP utilizza il metodo del simplesso riveduto, il quale opera alternando iterativamente due fasi: nella prima si individua una soluzione basica ammissibile del problema, nella seconda si determina la soluzione ottima. Se, come nel nostro caso, ad almeno una delle variabili decisionali è imposto il vincolo di interezza, allora la procedura si avvale della tecnica di ottimizzazione del Branch and Bound. Il Branch and Bound è una metodologia di ricerca della soluzione ottima che effettua una esplorazione parziale dell'insieme delle soluzioni ammissibili. Il problema iniziale dà luogo ad un insieme di sottoproblemi, secondo una struttura ad albero, ognuno dei quali deriva dal precedente mediante il partizionamento dell'insieme delle soluzioni ammissibili. Nella PROC LP viene risolto inizialmente il rilassamento lineare del problema di PLI, cioè il problema ottenuto eliminando il vincolo di interezza delle variabili, utilizzando l'algoritmo del simplesso riveduto, quindi vengono immessi i vincoli di interezza nei vari passi dell'algoritmo per l'individuazione dei sottoproblemi. Nessuna limitazione è imposta dalla PROC LP per la dimensione del problema, cioè sul numero di variabili e sul numero di vincoli presenti, se non quelli relativi alla quantità di memoria e di spazio fisico sul disco rigido disponibili.

La procedura accetta in input data set secondo due diversi formati: il formato "denso" ed il formato "sparso". Mentre nel formato denso il modello è espresso in modo molto simile rispetto alla sua formulazione matematica, ad ogni variabile del data set corrisponde una colonna del modello e ad ogni osservazione corrisponde una riga del modello, il formato sparso consente di specificare solamente i coefficienti non nulli delle variabili facenti parte del modello. La nostra scelta è caduta su quest'ultimo perché fornisce una rappresentazione più efficiente del modello, in prospettiva di una applicazione su grandi moli di dati.

Il data set sperimentale, sottoposto all'applicazione, ha restituito il valore ottimo della funzione obiettivo pari a 1254,579.

Nella Tabella seguente è possibile osservare le soluzioni trovate dalla PROC LP nel caso in esame.

Tabella 3.5 – Numero di partenze e di arrivi, mancate risposte in partenza ed in arrivo e flusso finale per porto di origine e destinazione.

Porto di origine	Flusso finale sull'arco					Porto di destinazione
	No	Mo	Md	Nd		
Anzio	0	89	89	0	89	Formia
Anzio	870	380	1.250	0	1.250	Ponza
Anzio	50	15	65	0	65	Terracina
Anzio	1.760	0	1.760	4	1.756	Ventotene
Formia	3	0	3	3	0	Anzio
Formia	750	187	937	187	750	Ponza
Formia	40	0	40	0	40	Ventotene
Ponza	1.500	0	1.500	0	1.500	Anzio
Ponza	800	21	821	0	821	Formia
Ponza	200	0	200	10	190	Terracina
Ponza	103	0	103	83	20	Ventotene
Terracina	50	10	60	50	10	Anzio
Terracina	0	222	222	0	222	Ponza
Terracina	65	2	67	65	2	Ventotene
Ventotene	1.470	127	1.597	397	1.200	Anzio
Ventotene	68	2	70	0	70	Formia
Ventotene	196	19	215	115	100	Ponza
Ventotene	80	8	88	0	88	Terracina
Totale	8.005	1.082	9.087	914	8.173	Totale

Dall'osservazione della Tabella 3.5 si nota come tutte le tratte per le quali vi era un porto di destinazione/partenza con totale assenza di dichiarazioni, abbiano subito un incremento numerico tale da ottenere il pareggiamento con il valore registrato dal porto di destinazione/partenza della medesima tratta.

Un buon comportamento nell'allocazione delle mancate risposte si evidenzia in tutte le tratte non bilanciate. Infatti, indipendentemente dall'entità della differenza tra il numero degli arrivi e quello delle partenze, il valore finale del flusso sull'arco è pari al maggiore tra N_o ed N_d o a valori ad esso prossimi.

Per quello che riguarda le tratte inizialmente in equilibrio, nessuna unità in più è stata aggiunta, se non nel caso della tratta Formia-Ponza. Va notato però che quest'ultimo

porto presenta il più alto sbilanciamento in valore assoluto tra flussi in ingresso e flussi in uscita.

In Tabella 3.6 sono riportate le situazioni iniziale e finale per quel che concerne il bilanciamento dei flussi in ciascun porto.

Tabella 3.6 – Scostamento iniziale e finale tra flusso in ingresso e uscita di natanti per porto.

Porto	Arrivi iniziali Nd	Partenze iniziali No	Arrivi finali	Partenze finali	differenze iniziali	differenze finali
Anzio	2.710	2.680	3.160	3.164	30	-4
Formia	980	793	980	980	187	0
Ponza	2.322	2.603	2.624	2.624	-281	0
Terracina	343	115	353	349	228	4
Ventotene	1.818	1.814	1.970	1.970	4	0
Totale	8.173	8.005	9.087	9.087	168	0

E' possibile riscontrare come, in questo caso, le soluzioni numeriche trovate rispondano, nel rispetto dei vincoli, all'obiettivo del minimo scostamento tra i flussi in ingresso ed in uscita per ciascun porto.

Inoltre, eseguendo la somma algebrica dei termini presenti nell'ultima colonna, si vede chiaramente come sia rispettata l'assunzione di base dei problemi di flusso a

costo minimo $\sum_{i \in N} b(i) = 0$, ovvero che la domanda globale deve eguagliare l'offerta globale.

3.1.3 - Confronto dei risultati ottenuti

Ad entrambi i modelli si chiedeva di realizzare due obiettivi, bilanciare i flussi in ingresso ed in uscita da ciascun porto ed incrementare i dati registrati con il minor numero possibile di mancate risposte, soddisfacendo al tempo stesso i vincoli di pareggiamento sulle tratte.

Dal confronto tra le due Tabelle 3.4 e 3.6, relative ai flussi in ingresso ed in uscita da ciascun porto, si evince chiaramente come il modello lineare abbia conseguito un grado di bilanciamento in ciascun porto uguale o superiore a quello ottenuto dall'applicazione del modello non lineare.

Di seguito sono riportate in forma tabellare le differenze assolute e percentuali iniziali e finali, quest'ultime relativamente ai due casi.

Tabella 3.7 – Differenze assolute e percentuali iniziali e finali nei due modelli.

Porto	Differenze iniziali		Differenze finali mod. non lineare		Differenze finali mod. lineare	
	N.	%	N.	%	N.	%
Anzio	30	-1,1	-38	1,2	-4	0,1
Formia	187	-23,6	27	-2,8	0	0,0
Ponza	-281	10,8	-17	0,6	0	0,0
Terracina	228	-198,3	4	-1,1	4	-1,1
Ventotene	4	-0,2	24	-1,2	0	0,0

Per quanto riguarda il secondo obiettivo, il modello non lineare ha incrementato il complesso degli arrivi di 1.102 unità ed il complesso delle partenze di 934 unità, determinando un flusso finale di viaggi pari a 9.107. Il modello lineare invece ha aggiunto 1.082 unità in partenza e 914 in arrivo, individuando il valore finale del flusso di viaggi pari a 9.087, lievemente inferiore rispetto all'altro.

Infine, in Tabella 3.8 sono riportati per ciascuna tratta gli incrementi percentuali rispetto al valore massimo sulla tratta stessa.

Tabella 3.8 – Incrementi assoluti e percentuali tra flusso finale e massimo flusso dichiarato per tratta, determinati con il modello non lineare e con il modello lineare.

ORIGINE	max{No,Nd}	Modello non lineare			Modello lineare			DESTINAZIONE
		Flusso finale	Differenze		Flusso finale	Differenze		
			N.	%		N.	%	
Anzio	89	89	0	0,0	89	0	0,0	Formia
Anzio	1.250	1.250	0	0,0	1.250	0	0,0	Ponza
Anzio	65	65	0	0,0	65	0	0,0	Terracina
Anzio	1.760	1.760	0	0,0	1.760	0	0,0	Ventotene
Formia	3	3	0	0,0	3	0	0,0	Anzio
Formia	750	876	126	16,8	937	187	24,9	Ponza
Formia	40	74	34	85,0	40	0	0,0	Ventotene
Ponza	1.500	1.537	37	2,5	1.500	0	0,0	Anzio
Ponza	821	821	0	0,0	821	0	0,0	Formia
Ponza	200	200	0	0,0	200	0	0,0	Terracina
Ponza	103	103	0	0,0	103	0	0,0	Ventotene
Terracina	50	60	10	20,0	60	10	20,0	Anzio
Terracina	222	222	0	0,0	222	0	0,0	Ponza
Terracina	65	67	2	3,1	67	2	3,1	Ventotene
Ventotene	1.470	1.526	56	3,8	1.597	127	8,6	Anzio
Ventotene	70	70	0	0,0	70	0	0,0	Formia
Ventotene	196	296	100	51,0	215	19	9,7	Ponza
Ventotene	88	88	0	0,0	88	0	0,0	Terracina
Totale	8.742	9.107	365	4,2	9.087	345	3,9	Totale

Il modello lineare si è dimostrato nel caso sperimentale analizzato maggiormente rispondente ai requisiti desiderati e dunque preferibile rispetto al precedente.

3.2 - Applicazione ai dati del Trasporto Marittimo

Riportiamo, a conclusione del presente lavoro, un confronto tra i risultati prodotti dalla procedura LP del modulo SAS/OR ed i dati relativi alla rilevazione statistica dei Trasporti Marittimi con riferimento all'anno 2002.

Naturalmente il raffronto ha come unico scopo quello di indagare la bontà o meno del comportamento del modello lineare proposto in un caso reale.

Il piano di diffusione dei dati a livello nazionale prevede la pubblicazione di dati aggregati per porto, ed in particolare per ciò che concerne i viaggi, non distingue tra navigazione di cabotaggio (tra porti italiani) e navigazione internazionale. Per questo motivo, benché siano stati sottoposti alla procedura tutti i dati di cabotaggio semilavorati, cioè i dati provenienti dalla registrazione a valle del processo di

correzione e controllo dei singoli codici e delle relazioni logiche tra variabili, la nostra analisi si limiterà a considerare i risultati ottenuti solo per quei porti caratterizzati esclusivamente o quasi da navigazione di cabotaggio. In tutti gli altri casi l'entità numerica del dato risentirebbe troppo del flusso di navi dirette verso (o provenienti da) paesi esteri, e dell'eventuale sbilanciamento prodotto da questo flusso.

Tabella 3.9 – Confronto tra dati relativi al Trasporto Marittimo (anno 2002) e risultati dell'applicazione del modello lineare.

Porto	Arrivi		Scostamento Arrivi		Partenze		Scostamento Partenze	
	N.	da modello	N.	%	N.	da modello	N.	%
Porto S. Stefano	2.951	2.953	2	0,1	2.949	2.953	4	0,1
Capri	22.084	21.152	-932	-4,2	22.068	21.152	-916	-4,2
Isole Tremiti	1.986	1.976	-10	-0,5	1.986	1.976	-10	-0,5
La Maddalena	25.040	25.038	-2	0,0	25.040	25.040	0	0,0
Anzio	900	852	-48	-5,3	900	852	-48	-5,3

Nella tabella 3.9 sono stati riportati i risultati ottenuti per alcuni porti.

In relazione all'obiettivo di bilanciamento dei flussi in ingresso ed in uscita, si rileva come per tutti i porti, ad eccezione de La Maddalena per la quale $b=2$, sia raggiunto lo stato di equilibrio.

Come si evidenzia i valori numerici risultanti dal modello sono molto prossimi ai valori reali. Va comunque osservato che non in ogni caso esaminato si sono riscontrati scostamenti così esigui, ma questi risultati costituiscono comunque una base di analisi fortemente motivante per ulteriori e successivi approfondimenti di ricerca.

BIBLIOGRAFIA

- R.K. Ahuja, T.L. Magnanti, J.B. Orlin - *"Network flows"* Prentice Hall 1993
- A. Coloni - *"Elementi di Ricerca Operativa"*, Edizioni Zanichelli 1992
- Istat, *"Statistiche dei trasporti marittimi"*, Collana Informazioni 2001
- Istat, *"Manuale per la rilevazione statistica sui trasporti marittimi"*, 2000
- P. Bertolazzi, M. Lucertini - *"Programmazione matematica: Formulazione di problemi e applicazioni"* F. Angeli 1991
- B. Simeone - *"Preferenze e Decisioni"* (Dispense del corso - ediz. 2001)
- D. Bertsimas, J.N. Tsitsiklis - *"Introduction to Linear Optimization"*
Athena Scientific 1997
- Sas System - Sas OnlineDoc versione 8

Note

Sono stati utilizzati i seguenti software:

Modulo SAS/OR – SAS System versione 6.12

AMPL Plus Student Edition 1.6

Appendice 1

Ampl modello non lineare

Ampl dati

Output Ampl-Minos modello non lineare

Output Ampl-Cplex modello non lineare

Appendice 2

Programma SAS per il modello lineare

Output SAS modello lineare

Appendice 1

Ampl modello non lineare

```
set PORTI;
set ARRIVO within PORTI;
set PARTENZA within PORTI;

param No {PARTENZA, ARRIVO};
param Nd {PARTENZA, ARRIVO};
param ct {i in PARTENZA, j in ARRIVO};
param cp {l in PORTI};
#param ct {i in PARTENZA, j in ARRIVO} := (No[i,j]+Nd[i,j])/(2*abs(No[i,j]-
Nd[i,j]));
#param cp {l in PORTI} := (sum {i in PARTENZA} Nd[i,l] - sum {j in ARRIVO}
No[l,j])/2 ;

var Mo{PARTENZA, ARRIVO} integer ;
var Md{PARTENZA, ARRIVO} integer ;

minimize obiettivo:sum{l in PORTI} (cp[l]*(sum{i in PARTENZA}
(Nd[i,l]+Md[i,l]) - sum{j in ARRIVO} (No[l,j]+Mo[l,j])))^2 + su
m {i in PARTENZA, j in ARRIVO} (ct[i,j]*Mo[i,j])^2 + sum {i in PARTENZA, j
in ARRIVO} (ct[i,j]*Md[i,j])^2;

subject to vincolo1 {i in PARTENZA, j in ARRIVO}: No[i,j]+Mo[i,j] =
Nd[i,j]+Md[i,j];
subject to vincolo2 {i in PARTENZA, j in ARRIVO}: 0 <= Mo[i,j] <= Nd[i,j];
subject to vincolo3 {i in PARTENZA, j in ARRIVO}: 0 <= Md[i,j] <= No[i,j];
```

Ampl dati

```
set PORTI      := Anzio   Ponza   Formia   Ventotene  Terracina;  
set PARTENZA  := Anzio   Ponza   Formia   Ventotene  Terracina;  
set ARRIVO    := Anzio   Ponza   Formia   Ventotene  Terracina;
```

```
param No:           Anzio   Ponza   Formia   Ventotene  Terracina  
:=  
    Anzio           0       870     0       1760       50  
    Ponza           1500    0       800     103        200  
    Formia           3       750     0       40         0  
    Ventotene       1470    196     68      0          80  
    Terracina       50      0       0       65         0   ;
```

```
param Nd:           Anzio   Ponza   Formia   Ventotene  Terracina :=  
    Anzio           0       1250    89      1756       65  
    Ponza           1500    0       821     20         190  
    Formia           0       750     0       40         0  
    Ventotene       1200    100     70      0          88  
    Terracina       10      222     0       2          0   ;
```

```
param ct:           Anzio   Ponza   Formia   Ventotene  Terracina  
:=  
    Anzio           0       0.70    0.01     1.00       0.77  
    Ponza           1.00    0       0.97     0.19       0.95  
    Formia           0.01    1.00    0        1.00       0  
    Ventotene       0.82    0.51    0.97     0          0.91  
    Terracina       0.20    0.01    0        0.03       0   ;
```

```
param cp:= Anzio    2.02  
           Ponza    2.24  
           Formia   2.47  
           Ventotene 2.00  
           Terracina 5.97 ;
```

Output Ampl-Minos modello non lineare

OBJECTIVES:

obiettivo = 209452

VARIABLES:

Md [*,*]

:	Anzio	Formia	Ponza	Terracina	Ventotene	:=
Anzio	0	0	0	0	4	
Formia	3	0	125.652	0	34.1023	
Ponza	36.6082	0	0	10	83	
Terracina	50	0	0	0	65	
Ventotene	325.598	0	196	0	0;	

Mo [*,*]

:	Anzio	Formia	Ponza	Terracina	Ventotene	:=
Anzio	0	89	380	15	0	
Formia	0	0	125.652	0	34.1023	
Ponza	36.6082	21	0	0	0	
Terracina	10	0	222	0	2	
Ventotene	55.5975	2	100	8	0;	

CONSTRAINTS (Dual Values):

vincolo1 [*,*]

:	Anzio	Formia	Ponza	Terracina	Ventotene	:=
Anzio	316.593	316.61	688.993	334.38	-204.038	
Formia	0	-332.448	-81.144	-332.448	-264.243	
Ponza	243.376	209.678	170.16	-303.177	-202.031	
Terracina	-284.327	-285.127	-285.083	-285.127	-285.124	
Ventotene	-121.271	-192.275	-144.018	-182.789	-196.038;	

vincolo2 [*,*]

:	Anzio	Formia	Ponza	Terracina	Ventotene	:=
Anzio	0	0	0	0	520.631	
Formia	-332.448	0	0	0	0	
Ponza	0	0	0	473.337	372.191	
Terracina	0	0	0	0	0	
Ventotene	0	0	-2.84217e-14	0	0;	

vincolo3 [*,*]

:	Anzio	Formia	Ponza	Terracina	Ventotene	:=
Anzio	0	649.058	518.833	619.507	0	
Formia	-316.592	0	0	-47.3205	0	
Ponza	0	542.125	0	0	0	
Terracina	-596.92	47.3205	-455.243	0	-88.9681	
Ventotene	0	140.173	-212.219	102.338	0;	

Output Ampl-Cplex modello non lineare

OBJECTIVES:

obiettivo = 209452

VARIABLES:

Md [*,*]

:	Anzio	Formia	Ponza	Terracina	Ventotene
:=					
Anzio	0	0	2.52238e-08	2.11246e-08	4
Formia	3	0	125.652	0	34.1023
Ponza	36.6082	2.414e-08	0	10	83
Terracina	50	0	0	0	65
Ventotene	325.598	9.33629e-08	196	1.27872e-07	0;

Mo [*,*]

:	Anzio	Formia	Ponza	Terracina	Ventotene	:=
Anzio	0	89	380	15	2.51367e-08	
Formia	0	0	125.652	0	34.1023	
Ponza	36.6082	21	0	2.76479e-08	3.51618e-08	
Terracina	10	0	222	0	2	
Ventotene	55.5975	2	100	8	0;	

CONSTRAINTS (Dual Values):

vincolo1 [*,*]

:	Anzio	Formia	Ponza	Terracina	Ventotene	:=
Anzio	316.593	316.61	688.993	334.38	-204.038	
Formia	0	-332.448	-81.144	-332.448	-264.243	
Ponza	243.376	209.678	170.16	-303.177	-202.031	
Terracina	13.7968	-285.127	-285.083	-285.127	-251.498	
Ventotene	-121.271	-192.275	-63.2367	-182.789	-196.038;	

vincolo2 [*,*]

:	Anzio	Formia	Ponza	Terracina	Ventotene	:=
Anzio	0	0	0	0	0	
Formia	-332.448	0	0	0	0	
Ponza	0	0	0	0	0	
Terracina	0	0	0	0	0	
Ventotene	0	0	0	0	0;	

vincolo3 [*,*]

:	Anzio	Formia	Ponza	Terracina	Ventotene	:=
Anzio	0	649.058	0	0	0	
Formia	-316.592	0	0	-47.3205	0	
Ponza	0	0	0	0	0	
Terracina	0	47.3205	-455.243	0	0	
Ventotene	0	0	0	0	0;	

Appendice 2

Programma SAS per il modello lineare

```

/*****/
/*  Dati per l'esempio          */
/*****/
options source;

Data TFLOW;
Input PO $ PD $ _VAR_ $ _VALUE_  ;
CARDS;
ITFOM      ITANZ      ND      0
ITPNZ      ITANZ      ND      1500
ITTRC      ITANZ      ND      10
ITVTT      ITANZ      ND      1200
ITFOM      ITANZ      NO      3
ITPNZ      ITANZ      NO      1500
ITTRC      ITANZ      NO      50
ITVTT      ITANZ      NO      1470
ITANZ      ITFOM      ND      89
ITPNZ      ITFOM      ND      821
ITVTT      ITFOM      ND      70
ITANZ      ITFOM      NO      0
ITPNZ      ITFOM      NO      800
ITVTT      ITFOM      NO      68
ITANZ      ITPNZ      ND      1250
ITFOM      ITPNZ      ND      750
ITTRC      ITPNZ      ND      222
ITVTT      ITPNZ      ND      100
ITANZ      ITPNZ      NO      870
ITFOM      ITPNZ      NO      750
ITTRC      ITPNZ      NO      0
ITVTT      ITPNZ      NO      196
ITANZ      ITTRC      ND      65
ITPNZ      ITTRC      ND      190
ITVTT      ITTRC      ND      88
ITANZ      ITTRC      NO      50
ITPNZ      ITTRC      NO      200
ITVTT      ITTRC      NO      80
ITANZ      ITVTT      ND      1756
ITFOM      ITVTT      ND      40
ITPNZ      ITVTT      ND      20
ITTRC      ITVTT      ND      2
ITANZ      ITVTT      NO      1760
ITFOM      ITVTT      NO      40
ITPNZ      ITVTT      NO      103
ITTRC      ITVTT      NO      65
;
RUN;

/*****/
/*  Lettura dati da file esterno */
/*****/

/*
Data TFLOW;
infile "c:\dati.txt";
input PO $ PD $ _VAR_ $ _VALUE_ ;

```

```

run;
*/

PROC SORT DATA=TFLOW;
  By PO PD descending _VAR_;
RUN;

Data MATRIX;
  SET TFLOW;
  LAGVALUE=LAG(_VALUE_);

  IF Last.PD THEN DO;
    DIF=Sum(_VALUE_,-LAGVALUE);
    SOM=Sum(_VALUE_,LAGVALUE );
    appo=min(_VALUE_,LAGVALUE )/max(_VALUE_,LAGVALUE );
    if appo ge 0.01 then
      SS=appo;
      else SS=0.01;
    OUTPUT;
    _VALUE_=DIF;
    _VAR_='DF';
    OUTPUT;
    _VALUE_=SS;
    _VAR_='CT';
    _R_=0;

  End;

  OUTPUT;
  By PO PD;
  DROP LAGVALUE DIF APPO;
Run;

PROC SUMMARY DATA=MATRIX NOPRINT MISSING;
WHERE _VAR_ NE 'DF';
CLASS PO PD _VAR_;
VAR _VALUE_;
OUTPUT OUT=SPOPD(WHERE=( _TYPE_=5 AND _VAR_ EQ 'NO' OR _TYPE_=3 AND _VAR_ EQ 'ND' ) ) SUM=;
RUN;

DATA CS;
  SET SPOPD;
  IF PO='' then ID=PD; ELSE ID=PO;
  IF _VAR_='NO' THEN _VALUE_=-1*_VALUE_;
RUN;

Proc sort DATA=CS;
  by ID _VAR_;
RUN;

DATA CS1;
  SET CS;
  By ID _VAR_;
  LAGVALUE=LAG(_VALUE_);
  IF Last.ID THEN DO;
    _VAR_='CP';
    IF MIN(ABS(_VALUE_),ABS(LAGVALUE)) eq 0 THEN
      _VALUE_=2*max(ABS(_VALUE_),ABS(LAGVALUE));
    ELSE
      _VALUE_=2*max(ABS(_VALUE_),ABS(LAGVALUE))/min(ABS(_VALUE_),ABS(LAGVALUE));
  END;
  IF _VAR_ EQ 'CP';

```

```

KEEP ID _VALUE_;
RUN;

PROC SUMMARY DATA=CS NOPRINT MISSING NWAY;
CLASS ID;
VAR _VALUE_;
OUTPUT OUT=CS(DROP=_TYPE_ _FREQ_) SUM=_VALUE_;
RUN;

DATA MATRIX;
SET MATRIX
  SPOPD(IN=A DROP=_FREQ_)
  CS (IN=B)
  CS1 (IN=C);

IF A THEN DO;
  IF _TYPE_ EQ 3 THEN _VAR_='NDPD';
  IF _TYPE_ EQ 5 THEN _VAR_='NOPO';
END;

IF B THEN DO;
  _VAR_='CS';
  PO=ID;
  PD=ID;
END;

IF C THEN DO;
  _VAR_='CP';
  _R_ =1;
  PO=ID;
  PD=ID;
END;

DROP _TYPE_ ID SOM SS;
RUN;
*-----;

DATA LPFLOW;
LENGTH _TYPE_ $8.
  _ROW_ $17.
  _COL_ $17.
  _COEF_ 8.;

SET MATRIX(IN=A WHERE=( _VAR_ IN('CP','CT','NO','ND','CS','DF')));

*-----;
*Funzione Obiettivo;
*-----;

IF A AND _N_ EQ 1
THEN DO;
  _TYPE_='MIN';
  _COL_ ='';
  _ROW_ ='BILANC';
  _COEF_ =.;
  OUTPUT;
END;

IF _VAR_ EQ 'CT'
THEN DO;
  _TYPE_ ='';
  _COL_ =COMPRESS('MO' || '_' || PO || '_' || PD);

```



```

_ROW_ = 'BILANC';
_COEF_ = _VALUE_ ;
OUTPUT;
_TYPE_ = '';
_COL_ = COMPRESS('MD' || '_' || 'PO' || '_' || PD);
_ROW_ = 'BILANC';
_COEF_ = _VALUE_ ;
OUTPUT;
END;

```

```

IF _VAR_ EQ 'CP'
THEN DO;
_TYPE_ = '';
_COL_ = COMPRESS('A' || '_' || 'PO' || '_' || PD);
_ROW_ = 'BILANC';
_COEF_ = _VALUE_ ;
OUTPUT;
_TYPE_ = '';
_COL_ = COMPRESS('B' || '_' || 'PO' || '_' || PD);
_ROW_ = 'BILANC';
_COEF_ = _VALUE_ ;
OUTPUT;
END;

```

```

*-----;
*Analisi Sensibilità ;
*-----;

```

```

IF A AND _N_ EQ 1
THEN DO;
_TYPE_ = 'PRICESEN';
_COL_ = '';
_ROW_ = 'CHANGE';
_COEF_ = .;
OUTPUT;
END;

```

```

IF _VAR_ EQ 'CT'
THEN DO;
_TYPE_ = '';
_COL_ = COMPRESS('MO' || '_' || 'PO' || '_' || PD);
_ROW_ = 'CHANGE';
_COEF_ = _R_ ;
OUTPUT;
_TYPE_ = '';
_COL_ = COMPRESS('MD' || '_' || 'PO' || '_' || PD);
_ROW_ = 'CHANGE';
_COEF_ = _R_ ;
OUTPUT;
END;

```

```

IF _VAR_ EQ 'CP'
THEN DO;
_TYPE_ = '';
_COL_ = COMPRESS('A' || '_' || 'PO' || '_' || PD);
_ROW_ = 'CHANGE';
_COEF_ = _R_ ;
OUTPUT;
_TYPE_ = '';
_COL_ = COMPRESS('B' || '_' || 'PO' || '_' || PD);
_ROW_ = 'CHANGE';
_COEF_ = _R_ ;
OUTPUT;

```

```

END;

*-----;
*Bilanciamento Flusso;
*-----;

*-- M01_;
IF _VAR_ EQ 'NO'
THEN DO;
  _TYPE_='';
  _COL_=COMPRESS('MO' || '_' || 'PO' || '_' || 'PD');
  _ROW_=COMPRESS('CONS' || '_' || 'PO');
  _COEF_=1;
  OUTPUT;
END;

*-- MD_1;
IF _VAR_ EQ 'ND'
THEN DO;
  _TYPE_='';
  _COL_=COMPRESS('MD' || '_' || 'PO' || '_' || 'PD');
  _ROW_=COMPRESS('CONS' || '_' || 'PD');
  _COEF_=-1;
  OUTPUT;
END;

*-- A1/B1--*;
IF _VAR_ EQ 'CP'
THEN DO;
  _TYPE_='';
  _COL_=COMPRESS('A' || '_' || 'PO' || '_' || 'PD');
  _ROW_=COMPRESS('CONS' || '_' || 'PO');
  _COEF_=1      ;
  OUTPUT;
  _TYPE_='';
  _COL_=COMPRESS('B' || '_' || 'PO' || '_' || 'PD');
  _ROW_=COMPRESS('CONS' || '_' || 'PD');
  _COEF_=-1      ;
  OUTPUT;
END;

*-- _RHS_ cons --*;
IF _VAR_ EQ 'CS'
THEN DO;
  _TYPE_='EQ';
  _COL_='_RHS_';
  _ROW_=COMPRESS('CONS' || '_' || 'PO');
  _COEF_=_VALUE_;
  OUTPUT;
END;

*-----;
* Pareggiamento TRATTA ;
*-----;

*-- M0nm/MDnm --*;
IF _VAR_ EQ 'CT'
THEN DO;
  _TYPE_='';
  _COL_=COMPRESS('MO' || '_' || 'PO' || '_' || 'PD');
  _ROW_=COMPRESS('PAR' || '_' || 'PO' || '_' || 'PD');
  _COEF_=1;
  OUTPUT;

```

```

    _TYPE_='';
    _COL_=COMPRESS('MD' || '_' || PO || '_' || PD);
    _ROW_=COMPRESS('PAR' || '_' || PO || '_' || PD);
    _COEF_=-1;
    OUTPUT;
END;

*-- _RHS_ par --*;
IF _VAR_ EQ 'DF'
THEN DO;
    _TYPE_='EQ';
    _COL_='_RHS_';
    _ROW_=COMPRESS('PAR' || '_' || PO || '_' || PD);
    _COEF_=_VALUE_;
    OUTPUT;
END;

*-----;
* Upperbound      ;
*-----;

IF _VAR_ EQ 'ND'
THEN DO;
    _TYPE_='UPPERBD';
    _COL_=COMPRESS('MO' || '_' || PO || '_' || PD);
    _ROW_=COMPRESS('AVAILABLE');
    _COEF_=_VALUE_;
    OUTPUT;
END;

IF _VAR_ EQ 'NO'
THEN DO;
    _TYPE_='UPPERBD';
    _COL_=COMPRESS('MD' || '_' || PO || '_' || PD);
    _ROW_=COMPRESS('AVAILABLE');
    _COEF_=_VALUE_;
    OUTPUT;
END;

KEEP _TYPE_ _COL_ _ROW_ _COEF_;
RUN;

proc lp SPARSEDATA
    data=LPFLOW
    primalin= RISUL;
Run;

/*
data Trasport;
input _type_ $8. _col_ $14. _row_ $17. _coef_ ;
datalines;

* inizio FO -----;
min      .          bilanc          .
.        MO11       bilanc           &CT11
.        MO12       bilanc           &CT12
.        .          .                .
.        MOnm       bilanc           &CTnm
.        MD11       bilanc           &CT11

```

```

.      MD12      bilanc      &CT12
.      .          .          .
.      MDnm     bilanc      &CTnm
.      A1       bilanc      &Cp1
.      A2       bilanc      &Cp2
.      .          .          .
.      An       bilanc      &Cpn
.      B1       bilanc      &Cp1
.      B2       bilanc      &Cp2
.      .          .          .
.      Bn       bilanc      &Cpn

```

* Bilanciamento FLUSSO -----;

```

.      MO1_     cons1       1
.      MD_1     cons1      -1
.      A1       cons1       1
.      B1       cons1      -1
eq  _rhs_     cons1      &CS1 * (-No1_ +Nd_1)

```

```

.      MO2_     cons2       1
.      MD_2     cons2      -1
.      A2       cons2       1
.      B2       cons2      -1
eq  _rhs_     cons2      &CS2 * (-No2_ +Nd_2)

```

```

.      MOn_     consn       1
.      MD_n     consn      -1
.      An       consn       1
.      Bn       consn      -1
eq  _rhs_     consn      &CSn * (-Non_ +Nd_n)

```

* Pareggiamento TRATTA -----;

```

.      MO11     par11       1
.      MD11     par11       1
eq  _rhs_     par11      &Df11 * (-NO11+ND11)

```

```

.      MO12     par12       1
.      MD12     par12      -1
eq  _rhs_     par12      &Df12 * (-NO12+ND12)

```

```

.      MO1n     par1n       1
.      MD1n     par1n      -1
eq  _rhs_     par1n      &Df1n * (-NO1n+ND1n)

```

```

.      MOm1     parm1       1
.      MDm1     parm1      -1
eq  _rhs_     parm1      &Dfm1 * (-NOM1+NDm1)

```

```

.      MOmn     parmnm      1
.      MDmn     parmnm     -1
eq  _rhs_     parmnm     &Dfmn * (-NOMn+NDmn)

```

```

upperbd .      available      .

```

```

.      MO11      available      &ND11
.      MO12      available      &ND12
.      MO1n      available      &ND1n
.      MD11      available      &NO11
.      MD12      available      &NO12
.      MD1n      available      &NO1n
;
proc lp SPARSEDATA;
Run;

*/

```

Output SAS modello lineare

The SAS System

LINEAR PROGRAMMING PROCEDURE

PROBLEM SUMMARY

Min BILANC	Objective Function
RHS	Rhs Variable
TYPE	Type Variable
Problem Density	0.077505
Variable Type	
Non-negative	10
Upper Bounded	36
Total	46
Constraint Type	
EQ	23
Objective	1
Total	24

LINEAR PROGRAMMING PROCEDURE

SOLUTION SUMMARY

Terminated Successfully

Objective value	1254.579
Phase 1 iterations	21
Phase 2 iterations	6
Phase 3 iterations	0
Integer iterations	0
Integer solutions	0
Initial basic feasible variables	2
Time used (secs)	0
Number of inversions	3
Machine epsilon	1E-8
Machine infinity	1.7976931349E308
Maximum phase 1 iterations	100
Maximum phase 2 iterations	100
Maximum phase 3 iterations	99999999
Maximum integer iterations	100
Time limit (secs)	120

L I N E A R P R O G R A M M I N G P R O C E D U R E

VARIABLE SUMMARY

Variable Col Name	Status	Type	Price	Activity	Reduced Cost
1 A_ITANZ_ITANZ		NON-NEG	2.02238806	0.000000	4.044776
2 A_ITFOM_ITFOM		NON-NEG	2.47162673	0.000000	1.881770
3 A_ITPNZ_ITPNZ		NON-NEG	1.78409527	0.000000	3.194238
4 A_ITTRC_ITTRC	BASIC	NON-NEG	5.96521739	4.000000	0.000000
5 A_ITVTT_ITVTT		NON-NEG	2.00441014	0.000000	2.394145
6 B_ITANZ_ITANZ	BASIC	NON-NEG	2.02238806	4.000000	0.000000
7 B_ITFOM_ITFOM		NON-NEG	2.47162673	0.000000	3.061484
8 B_ITPNZ_ITPNZ		NON-NEG	1.78409527	0.000000	0.373952
9 B_ITTRC_ITTRC		NON-NEG	5.96521739	0.000000	11.930435
10 B_ITVTT_ITVTT		NON-NEG	2.00441014	0.000000	1.614675
11 MD_ITANZ_ITFOM		UPPERBD	0.01	0.000000	2.632245
12 MD_ITANZ_ITPNZ		UPPERBD	0.696	0.000000	2.004245
13 MD_ITANZ_ITTRC		UPPERBD	0.76923077	0.000000	9.526067
14 MD_ITANZ_ITVTT	BASIC	UPPERBD	0.99772727	4.000000	0.000000
15 MD_ITFOM_ITANZ		UPPERBD	0.01	3.000000	-2.012388
16 MD_ITFOM_ITPNZ	BASIC	UPPERBD	1	187.000	0.000000
17 MD_ITFOM_ITVTT		UPPERBD	1	0.000000	0.610265
18 MD_ITPNZ_ITANZ	DEGEN	UPPERBD	1	0.000000	0.000000
19 MD_ITPNZ_ITFOM		UPPERBD	0.97442144	0.000000	3.948843
20 MD_ITPNZ_ITTRC	BASIC	UPPERBD	0.95	10.000000	0.000000
21 MD_ITPNZ_ITVTT	BASIC	UPPERBD	0.19417476	83.000000	0.000000
22 MD_ITTRC_ITANZ	BASIC	UPPERBD	0.2	50.000000	0.000000
23 MD_ITTRC_ITPNZ		UPPERBD	0.01	0.000000	-1.400143
24 MD_ITTRC_ITVTT	BASIC	UPPERBD	0.03076923	65.000000	0.000000
25 MD_ITVTT_ITANZ	BASIC	UPPERBD	0.81632653	397.000	0.000000
26 MD_ITVTT_ITFOM		UPPERBD	0.97142857	0.000000	2.922449
27 MD_ITVTT_ITPNZ	BASIC	UPPERBD	0.51020408	115.000	0.000000
28 MD_ITVTT_ITTRC		UPPERBD	0.90909091	0.000000	8.173134
29 MO_ITANZ_ITFOM	BASIC	UPPERBD	0.01	89.000000	0.000000
30 MO_ITANZ_ITPNZ	BASIC	UPPERBD	0.696	380.000	0.000000
31 MO_ITANZ_ITTRC	BASIC	UPPERBD	0.76923077	15.000000	0.000000
32 MO_ITANZ_ITVTT		UPPERBD	0.99772727	0.000000	3.628108
33 MO_ITFOM_ITANZ		UPPERBD	0.01	0.000000	-0.579857
34 MO_ITFOM_ITPNZ	BASIC	UPPERBD	1	187.000	0.000000
35 MO_ITFOM_ITVTT		UPPERBD	1	0.000000	0.410143
36 MO_ITPNZ_ITANZ		UPPERBD	1	0.000000	1.387755
37 MO_ITPNZ_ITFOM	BASIC	UPPERBD	0.97442144	21.000000	0.000000
38 MO_ITPNZ_ITTRC		UPPERBD	0.95	0.000000	9.275361
39 MO_ITPNZ_ITVTT		UPPERBD	0.19417476	0.000000	1.408758
40 MO_ITTRC_ITANZ		UPPERBD	0.2	10.000000	-7.587605
41 MO_ITTRC_ITPNZ		UPPERBD	0.01	222.000	-5.955217
42 MO_ITTRC_ITVTT		UPPERBD	0.03076923	2.000000	-6.293414
43 MO_ITVTT_ITANZ	BASIC	UPPERBD	0.81632653	127.000	0.000000
44 MO_ITVTT_ITFOM	BASIC	UPPERBD	0.97142857	2.000000	0.000000
45 MO_ITVTT_ITPNZ	BASIC	UPPERBD	0.51020408	19.000000	0.000000
46 MO_ITVTT_ITTRC	BASIC	UPPERBD	0.90909091	8.000000	0.000000

L I N E A R P R O G R A M M I N G P R O C E D U R E

CONSTRAINT SUMMARY

Constraint Row Name	Type	S/S Col	Rhs	Activity	Dual Activity
1 BILANC	OBJECT			1254.579	.
2 PAR_ITANZ_ITFOM	EQ		89	89.000000	2.032388
3 PAR_ITANZ_ITPNZ	EQ		380	380.000	2.718388
4 PAR_ITANZ_ITTRC	EQ		15	15.000000	2.791619
5 PAR_ITANZ_ITVTT	EQ		-4	-4.000000	-0.607992
6 PAR_ITFOM_ITANZ	EQ		-3	-3.000000	0.000000
7 PAR_ITFOM_ITPNZ	EQ		0	0.000000	0.410143
8 PAR_ITFOM_ITVTT	EQ		0	0.000000	0.000000
9 PAR_ITPNZ_ITANZ	EQ		0	0.000000	1.022388
10 PAR_ITPNZ_ITFOM	EQ		21	21.000000	2.384565
11 PAR_ITPNZ_ITTRC	EQ		-10	-10.000000	-6.915217
12 PAR_ITPNZ_ITVTT	EQ		-83	-83.000000	0.195560
13 PAR_ITTRC_ITANZ	EQ		-40	-40.000000	1.822388
14 PAR_ITTRC_ITPNZ	EQ		222	222.000	0.000000
15 PAR_ITTRC_ITVTT	EQ		-63	-63.000000	0.358966

16	PAR_ITVTT_ITANZ	EQ	-270	-270.000	1.206062
17	PAR_ITVTT_ITFOM	EQ	2	2.000000	1.361164
18	PAR_ITVTT_ITPNZ	EQ	-96	-96.000000	0.899939
19	PAR_ITVTT_ITTRC	EQ	8	8.000000	1.298826
20	CONS_ITANZ	EQ	30	30.000000	-2.022388
21	CONS_ITFOM	EQ	187	187.000	0.589857
22	CONS_ITPNZ	EQ	-281	-281.000	-1.410143
23	CONS_ITTRC	EQ	228	228.000	5.965217
24	CONS_ITVTT	EQ	4	4.000000	-0.389735