

n. 3/2007

**La nuova procedura di controllo e correzione
degli errori e delle mancate risposte parziali
nell'indagine sui Risultati Economici delle
Aziende agricole (REA)**

U. Guarnera, O. Luzi e I. Tommasi

Le collane esistenti presso l'ISTAT - *Rivista di Statistica Ufficiale*, *Contributi ISTAT* e *Documenti ISTAT* - costituiscono strumenti per promuovere e valorizzare l'attività di ricerca e per diffondere i risultati degli studi svolti, in materia di statistica ufficiale, all'interno dell'ISTAT, del SISTAN, o da studiosi esterni.

La *Rivista di Statistica Ufficiale* accoglie lavori che hanno come oggetto la misurazione dei fenomeni economici, sociali, demografici e ambientali, la costruzione di sistemi informativi e di indicatori, le questioni di natura metodologica, tecnologica o istituzionale connesse al funzionamento dei sistemi statistici e al perseguimento dei fini della statistica ufficiale.

I lavori pubblicati in *Contributi ISTAT* sono diffusi allo scopo di stimolare il dibattito intorno ai risultati preliminari di ricerca in corso.

I *Documenti ISTAT* forniscono indicazioni su linee, progressi e miglioramenti di prodotto e di processo che caratterizzano l'attività dell'Istituto.

Il Comitato di redazione esamina le proposte di lavori da pubblicare nelle tre collane sopra indicate. Quelli pubblicati nei *Contributi ISTAT* e nei *Documenti ISTAT* sono valutati preventivamente dai dirigenti dell'Istituto, mentre i lavori pubblicati nella *Rivista di Statistica Ufficiale* sono subordinati al giudizio di referee esterni.

Direttore responsabile della Rivista di Statistica Ufficiale: Patrizia Cacioli

Comitato di Redazione delle Collane Scientifiche dell'Istituto Nazionale di Statistica

Coordinatore: Giulio Barcaroli

Membri:	Corrado C. Abbate	Rossana Balestrino	Giovanni A. Barbieri
	Giovanna Bellitti	Riccardo Carbini	Giuliana Coccia
	Fabio Crescenzi	Carla De Angelis	Carlo M. De Gregorio
	Gaetano Fazio	Saverio Gazzelloni	Antonio Lollobrigida
	Susanna Mantegazza	Luisa Picozzi	Valerio Terra Abrami
	Roberto Tomei	Leonello Tronti	Nereo Zamaro

Segreteria: Gabriella Centi, Carlo Deli e Antonio Trobia

Responsabili organizzativi per la *Rivista di Statistica Ufficiale*: Giovanni Seri e Carlo Deli

Responsabili organizzativi per i *Contributi ISTAT* e i *Documenti ISTAT*: Giovanni Seri e Antonio Trobia

n. 3/2007

**La nuova procedura di controllo e correzione
degli errori e delle mancate risposte parziali
nell'indagine sui Risultati Economici delle
Aziende agricole (REA)**

U. Guarnera(), O. Luzzi(*) e I. Tommasi(**)*

(*) ISTAT - Servizio Metodologie, tecnologie e software per la produzione dell'informazione statistica
(**) ISTAT - Servizio Statistiche sull'agricoltura

Contributi e Documenti Istat 2007

Istituto Nazionale di Statistica
Servizio Produzione Editoriale

Produzione libraria e centro stampa:
Carla Pecorario
Via Tuscolana, 1788 - 00173 Roma

Sommario

1. Introduzione	5
2. Individuazione dei valori anomali	6
2.1. Individuazione dei valori anomali mediante macroediting	6
2.2. Individuazione dei valori anomali mediante uso di rapporti:	7
3. Trattamento degli errori di coerenza tramite il software generalizzato Banff	7
3.1. Metodologie per il controllo e la correzione dei dati disponibili in Banff	7
3.2. Strategia di localizzazione e correzione automatica della struttura economica aziendale	9
4. Trattamento della sezione Occupazione e Redditi	12
4.1. Metodologie per l'imputazione di errori e mancate risposte nella sezione Occupazione e Redditi	12
4.1.1. Il donatore di distanza minima disponibile in Quis	12
4.1.2. Il metodo di imputazione con regressione mediante EM implementato in Quis	12
4.2. Trattamento della sezione Manodopera familiare	13
4.2.1. Preparazione del dataset	13
4.2.2. Individuazione degli altri errori e delle mancate risposte	14
4.2.3. Preparazione delle variabili di strato e di matching	14
4.2.4. Imputazione	14
4.3. Trattamento della sezione Altra Manodopera Aziendale	15
4.3.1. Trattamento dell'Altra manodopera aziendale a tempo indeterminato	15
4.3.2. Trattamento dell'Altra manodopera aziendale a tempo determinato	15
5. Conclusioni	16
Riferimenti bibliografici	16
Appendice A: Vincoli di coerenza utilizzati per il controllo probabilistico	18
Appendice B: Manodopera familiare e Altra manodopera aziendale: Indicatori sugli effetti del processo di controllo e correzione dati	23

1. Introduzione

A partire dall'anno contabile di riferimento 2003 l'ISTAT e l'INEA, di concerto con le Regioni e le Province autonome di Trento e Bolzano, hanno avviato il progetto di integrazione delle indagini RICA e REA (*Risultati Economici delle Aziende Agricole*), precedentemente svolte separatamente dai due istituti: la REA, rilevazione statistica campionaria effettuata con intervista diretta al conduttore d'azienda, relativa ad un insieme di variabili selezionate per la stima degli aggregati del settore agricolo; la RICA, rilevazione contabile avente ad oggetto tutte le variabili aziendali necessarie per analisi di tipo microeconomico. Operativamente, ISTAT e INEA hanno collaborato alla progettazione di un campione casuale stratificato unico, estratto dall'universo di aziende definito dal V Censimento dell'Agricoltura (relativo al 2000), che permetta di soddisfare contemporaneamente le esigenze conoscitive dell'INEA, per quanto riguarda obiettivi e utilizzo della rete contabile agricola, e dell'ISTAT ai fini della contabilità nazionale. Il nuovo campione ha una numerosità totale di circa 20.000 aziende di cui circa l'80%, comune alle due indagini, con dimensione economica superiore o uguale a 4 UDE (Unità di Dimensione Economica). Il restante 20%, con UDE inferiore a 4, completa le esigenze informative della REA e della contabilità nazionale. La raccolta dei dati, gestita dall'INEA, avviene attraverso la rilevazione contabile combinata con l'intervista retrospettiva; in particolare, per le aziende con UDE maggiore o uguale a 4 ci si avvale del software CONTINEA per la rilevazione contabile, per le restanti unità si ricorre ad intervista retrospettiva con ausilio del questionario cartaceo. L'indagine RICA-REA rileva costi, giacenze e scorte, acquisti e vendite di capitali fissi, reimpieghi, ricavi, autoconsumo, contributi alle aziende, costo del lavoro e redditi delle famiglie agricole.

Tra le innovazioni metodologiche e operative, l'indagine RICA-REA ha subito una completa riprogettazione della procedura di trattamento degli errori e delle mancate risposte in fase di controllo e correzione (C&C nel seguito) dei dati.

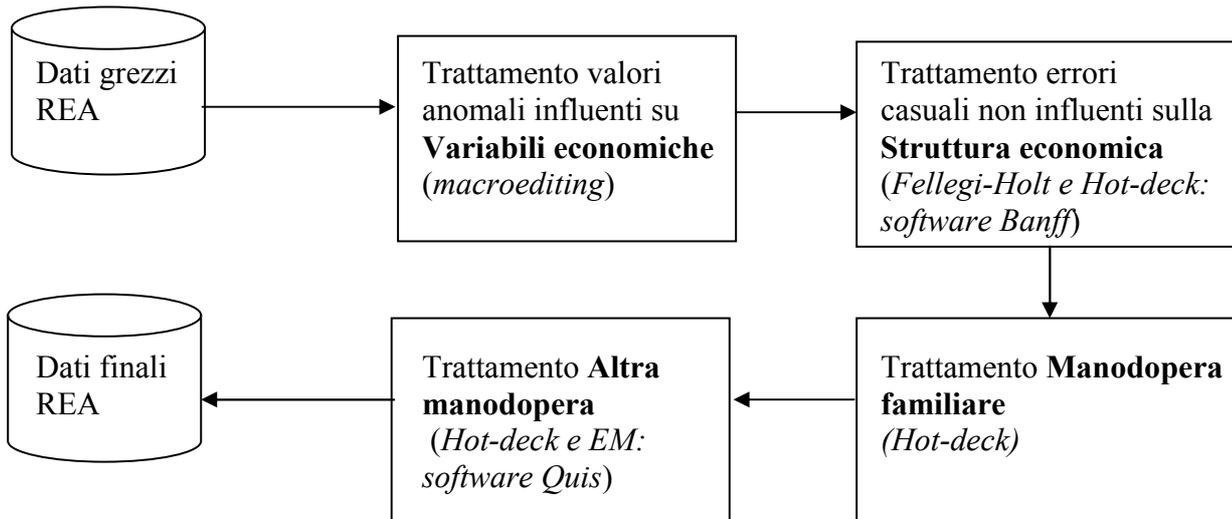
In particolare, la nuova procedura di C&C consiste in una combinazione di metodi e algoritmi integrati fra loro, ciascuno dei quali applicato ad uno specifico insieme di unità e di variabili.

In particolare, risultano applicate ai dati sia tecniche di tipo macroediting e di tipo grafico per l'individuazione dei valori anomali influenti, sia tecniche di tipo probabilistico per la localizzazione degli errori di coerenza (o di *compatibilità*) casuali, sia tecniche di imputazione parametriche e non parametriche (a seconda della tipologia di variabili oggetto di trattamento) per la ricostruzione dei valori errati o mancanti. Queste ultime hanno interessato le sole aziende con dimensione economica inferiore alle 4 UDE.

La strategia generale di C&C è illustrata nella Figura 1.

In questo documento, relativamente alle sole aziende REA, sono descritte le metodologie adottate per il controllo e il trattamento degli errori non campionari (incluse le mancate risposte parziali). In particolare, nella sezione 2 è descritto l'approccio adottato per l'individuazione dei valori anomali influenti sulle variabili principali del questionario. Il paragrafo 3 contiene la descrizione della strategia generale adottata per il trattamento delle variabili relative alla struttura economica delle aziende, in cui sono state adottate le metodologie probabilistiche disponibili nel software generalizzato Banff. Nel paragrafo 4 sono descritte le metodologie di individuazione e di imputazione degli errori e delle mancate risposte parziali relativamente alla sezione *Occupazione e Redditi* del questionario (separatamente per le sezioni *Manodopera Familiare* e *Altra Manodopera Aziendale*).

Figura 1: strategia complessiva di controllo e correzione dell'indagine REA



2. Individuazione dei valori anomali

Nel primo passo della procedura di C&C l'attenzione è stata concentrata sull'analisi di quei valori che possono avere maggiori effetti distorsivi sulle stime finali dei parametri di interesse (i totali delle variabili rilevate).

E' su questi dati, infatti, che è necessario concentrare le attività di controllo manuale/interattivo, riservando agli altri errori (che diremo *non influenti*) un trattamento di tipo automatico meno costoso in termini di risorse e tempi impiegati.

L'individuazione degli errori rilevanti è stata effettuata in due passi (illustrati nei paragrafi 2.1 e 2.2):

1. Individuazione delle stime anomale mediante tecniche di macroediting
2. Individuazione dei valori anomali rispetto a distribuzioni bivariate dei dati

Le variabili sottoposte a controllo in questa fase sono: *costi, giacenze e scorte, acquisto e vendita di capitali fissi, reimpieghi, ricavi, contributi, giornate di lavoro familiare e non familiare, costo del lavoro familiare e non familiare.*

2.1. Individuazione dei valori anomali mediante macroediting

Nell'approccio del macroediting (Granquist, 1992; Barcaroli *et al*, 1995), partendo da opportune aggregazioni (stratificazioni) dei dati, la revisione manuale/interattiva è limitata alle sole unità che contribuiscono maggiormente alla determinazione di stime anomale (*aggregati anomali*) rispetto a opportune stime di riferimento. L'obiettivo è quello di ridurre i tempi e i costi del controllo interattivo dei dati, mantenendo prefissati livelli di accuratezza delle stime.

Nel caso dell'indagine REA, per ogni variabile Y, le stime di interesse, per ogni dominio di analisi h, sono i totali

$$T_h(Y) = \sum_{i \in h} y_i w_i \quad (h=1, \dots, H)$$

mentre le stime di riferimento sono gli stessi totali calcolati nella precedente ripetizione dell'indagine, che indicheremo con $T_h(Y)'$. I pesi w_i utilizzati sono quelli che corrispondono all'inverso delle probabilità di inclusione iniziali.

Dall'analisi grafica dei valori $Th(Y)$ e $Th(Y)'$, vengono selezionati come aggregati anomali i casi in cui $Th(Y)$ e $Th(Y)'$ risultano "distanti" fra loro. Per ogni aggregato anomalo, vengono pertanto analizzati in ciascuno strato i contributi delle aziende ad esso appartenenti.

La revisione interattiva viene effettuata sui questionari corrispondenti alle aziende che contribuiscono maggiormente agli aggregati anomali.

I casi di errore vengono corretti interattivamente. I casi anomali corrispondenti a valori esatti vengono trattati modificando opportunamente il peso corrispondente, al fine di ridurre l'impatto sulle stime finali.

2.2 Individuazione dei valori anomali mediante uso di rapporti:

In questo caso, l'individuazione dei valori anomali viene effettuata mediante analisi grafica delle distribuzioni campionarie, in alcuni casi per strati, dei seguenti rapporti caratteristici :

1. *Costi totali/Ricavi totali*;
2. *Spese per acquisto di beni e servizi per le coltivazioni/(Ricavi per vendita di prodotti vegetali + Ricavi per vendita di prodotti vegetali trasformati)*;
3. *Spese per acquisto beni e servizi per allevamenti/(Ricavi per vendita di animali + Ricavi per Vendita prodotti zootecnici + Ricavi per vendita prodotti zootecnici trasformati)*;
4. *Retribuzioni lorde dipendenti tempo indeterminato/ Giornate di lavoro corrispondenti*;
5. *Contributi sociali dipendenti tempo indeterminato/ Retribuzioni lorde corrispondenti*;
6. *Retribuzioni lorde dipendenti tempo determinato/ Giornate di lavoro corrispondenti*;
7. *Contributi sociali dipendenti tempo determinato/ Retribuzioni lorde corrispondenti*.

I casi con rapporto anomalo rispetto alla distribuzione osservata vengono selezionati interattivamente per il controllo manuale, ed eventualmente corretti.

3. Trattamento degli errori di coerenza tramite il software generalizzato Banff

In questa fase vengono trattate tutte le variabili rilevate nelle seguenti sezioni del questionario:

- *Struttura dei costi*,
- *Giacenze e scorte*,
- *Acquisti e vendite di capitali fissi nell'esercizio*,
- *Reimpieghi*,
- *Ricavi e autoconsumo*,
- *Contributi*.

In questa fase del processo, la localizzazione e la correzione degli errori, e l'imputazione delle mancate risposte parziali è stata effettuata utilizzando le metodologie disponibili nel software generalizzato Banff, che rappresenta la versione in ambiente SAS del software generalizzato GEIS (Kovar *et al.*, 1988).

3.1. Metodologie per il controllo e la correzione dei dati disponibili in Banff

Il controllo e la correzione automatica di variabili numeriche continue avviene nel software Banff (Kovar *et al.*, 1988; Cotton, 1991) sulla base della metodologia Fellegi-Holt (*FH*) e della tecnica *hot deck* nota come *donatore di distanza minima* (*Nearest-Neighbor Donor*, *NND*). Questi metodi

sono particolarmente adatti al trattamento di errori di origine completamente casuale non influenti in termini di impatto sulle stime dei parametri obiettivo.

L'algoritmo probabilistico FH (Fellegi *et al.*, 1976) può essere utilizzato per identificare errori casuali in dati che devono risultare coerenti, a livello micro, rispetto a prefissati vincoli (edit) fra variabili osservate. Per ogni unità che viola almeno un edit, l'algoritmo FH identifica il minimo numero di valori (variabili) da modificare in modo da poter riportare il record alla situazione di coerenza (soluzione di minimo cambiamento). In altri termini, l'algoritmo FH assegna ad ogni variabile che compare in almeno un edit una probabilità di essere errata proporzionale al numero di edit violati che coinvolgono tale variabile. La selezione delle variabili selezionate per la correzione può essere influenzata dallo statistico attraverso l'uso di pesi, che possono essere associati alle variabili e che ne misurano il grado di affidabilità (tanto maggiore è il peso di una variabile, tanto minore è la sua probabilità di essere inclusa nella soluzione di minimo cambiamento). Nel caso di utilizzo dei pesi, la soluzione di minimo cambiamento per un record errato è quella che coinvolge il sottoinsieme di variabili aventi la minima somma dei pesi.

Dal momento che gli errori sono individuati sulla base di un prefissato insieme di edit, quando si adotta la metodologia FH è necessaria la massima cautela nella definizione delle regole di controllo. Infatti, se viene utilizzato un insieme con regole poco numerose ma soprattutto poco connesse (in termini di variabili coinvolte in esse), l'algoritmo potrebbe determinare una scelta sostanzialmente casuale delle variabili da modificare. In tali casi, un approccio di tipo deterministico potrebbe risultare preferibile. D'altra parte, troppi vincoli di coerenza potrebbero dare origine ad eccessiva complessità del problema di localizzazione dell'errore, con conseguente impossibilità per l'algoritmo a determinare una soluzione. Per questo motivo, in alcune applicazioni particolarmente complesse si ricorre alla suddivisione degli edit in due o più sottoinsiemi, ed alla loro applicazione sequenziale ai dati in una prefissata gerarchia. Solo se nessuna variabile è coinvolta in edit appartenenti ai diversi sottogruppi, l'applicazione dei diversi sottoinsiemi di regole (e quindi dell'algoritmo probabilistico) produce risultati che non dipendono dalla sequenza con cui essi sono verificati sui dati. In caso contrario, se cioè una o più variabili compaiono in diversi sottoinsiemi di vincoli, la sequenza di applicazione influenza il risultato, e le variabili trattate nell'ambito di un sottoinsieme di edit devono essere rese non modificabili (attraverso l'uso dei pesi) nei passi successivi di controllo.

Una volta localizzati, gli errori ed i valori originariamente mancanti (mancate risposte parziali) devono essere sostituiti (*imputati*) con valori ammissibili. A questo fine, diverse tecniche sono disponibili in Banff: un metodo di imputazione con soluzione analitica (o *deduttiva*) basato sulla ricerca dell'unico valore (se esiste) che sostituito ai valori errati soddisfa tutti i vincoli di compatibilità; una tecnica di imputazione con NND.

Per quanto riguarda la tecnica del NND, per ogni unità i che viola almeno un edit (unità *ricevente*), i valori mancanti o classificati come errati per una o più variabili sono sostituiti congiuntamente con i valori delle stesse variabili osservati nell'unità più vicina d_i (*donatore*). Il donatore d_i è selezionato da un *serbatoio* di unità complete (cioè prive di valori mancanti) e coerenti (cioè che soddisfano tutti gli edit) in base alla funzione di distanza *minmax* calcolata rispetto a un insieme di *variabili di accoppiamento* (o *matching*) trasformate mediante la funzione *rango*. Una importante caratteristica del metodo adottato consiste nel fatto che l'imputazione avviene nel rispetto dei vincoli: per ogni unità *ricevente*, il donatore d_i viene utilizzato per l'imputazione solo se l'unità risultante diviene coerente rispetto a tutti gli edit. Nella pratica, nel metodo NND disponibile in Banff il donatore d_i viene utilizzato per l'imputazione solo se l'unità risultante diviene coerente rispetto a tutti gli edit *di post-imputazione*: questi edit (detti post-edit) corrispondono a vincoli originali opportunamente resi meno stringenti (ad esempio, vincoli di quadratura rilassati in modo da ammettere come validi valori in un intorno del valore esatto di confronto - il totale). L'uso dei post-edit ha l'obiettivo di allargare la rosa dei potenziali donatori. Infatti, dato un record i che fallisce una certa uguaglianza, può accadere che il sistema:

- non riesca a trovare un donatore che fornisca ad i un valore tale da riportarlo nella condizione di soddisfare l'uguaglianza;
- scarti record donatori "vicini" ad i , ma che non gli garantiscono il rispetto dell'uguaglianza, e seleziona un donatore più "distante" da i , ma che gli fornisce il valore richiesto.

E' evidente che se da un lato l'uso dei *post-edit* facilita l'individuazione di un donatore appropriato, dall'altro esso rende necessaria una successiva verifica dei dati imputati per verificare se qualcuno di essi viola i vincoli di uguaglianza originali.

In generale, le tecniche di imputazione (incluso il *NND*) risultano più efficienti se esse vengono applicate all'interno di *celle di imputazione*: si tratta di sotto-popolazioni definite sulla base di variabili (*covariate*) statisticamente associate alle variabili oggetto di correzione/imputazione, e quindi ritenute maggiormente omogenee in termini del/dei fenomeno/i oggetto di imputazione.

I metodi da donatore (o *hot deck*), sono riconosciuti fra i più adatti alla ricostruzione delle mancate risposte in termini di preservazione della variabilità delle distribuzioni semplici delle variabili imputate e delle associazioni fra esse, se di tali associazioni si tiene conto nel modello di imputazione. Ciò può essere fatto in fase di stratificazione e/o nel calcolo della distanza e/o adottando modalità di imputazione congiunta delle variabili con uno stesso donatore (Kalton e Kasprzyk, 1982; Chen e Shao, 2000).

3.2. Strategia di localizzazione e correzione automatica della struttura economica aziendale

Obiettivo di questa fase è risolvere in modo automatico (quindi poco costoso in termini di tempo e risorse) tutte le incoerenze logico-matematiche presenti nei dati, ma che sono considerate poco influenti sulle stime finali avendo superato la fase di controllo precedente. L'algoritmo utilizzato è appunto quello FH implementato in Banff. Le regole di controllo utilizzate (quadrature, esistenze incrociate fra variabili collegate fra loro nel questionario, uguaglianze fra valori osservati o ricavabili in più punti del modello, disuguaglianze) sono contenute nell'Allegato A.

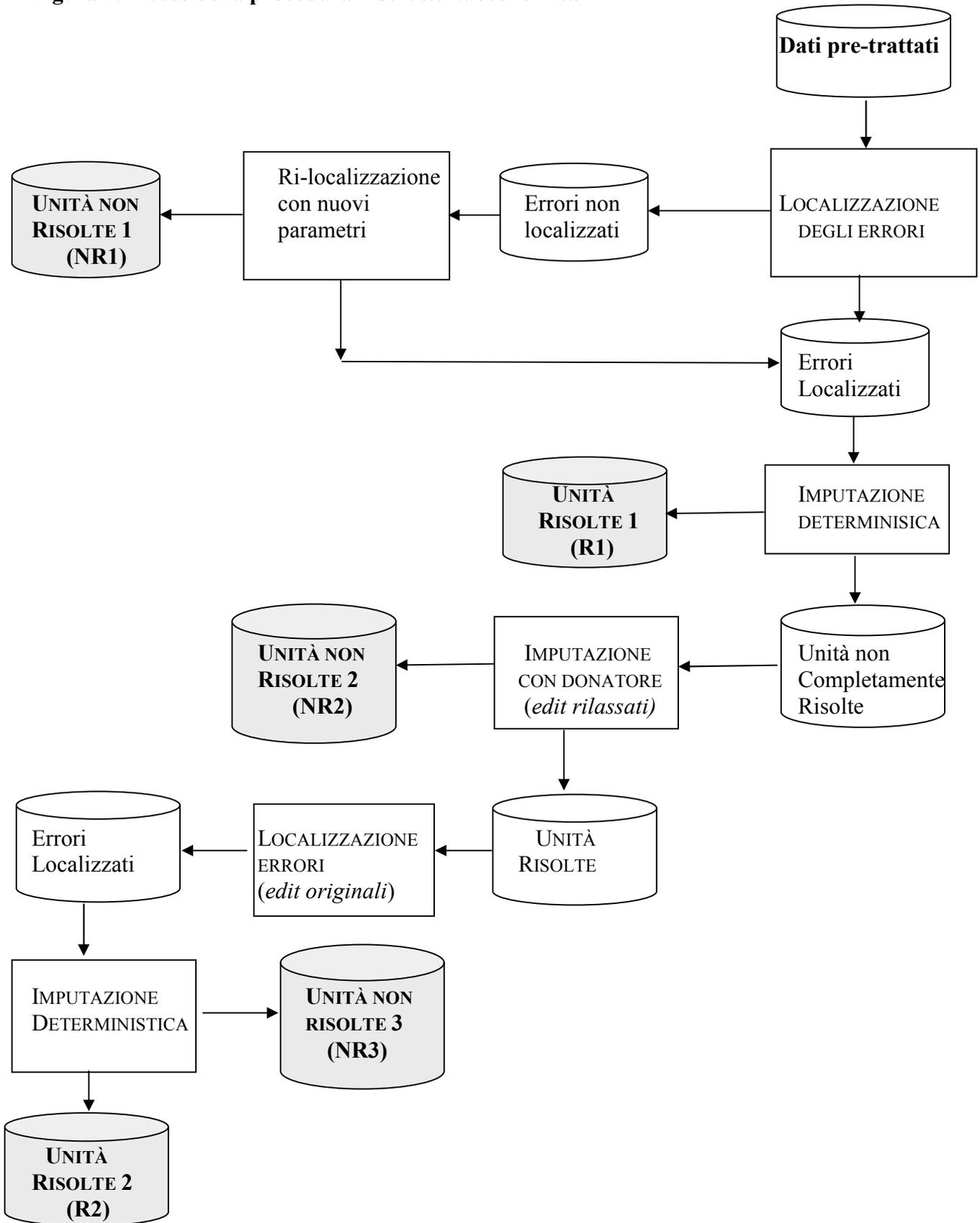
La procedura, che lavora su 9 variabili quantitative descrittive della struttura aziendale (superfici e consistenza allevamenti) e su tutte le 117 variabili economiche delle sezioni 12-17 del questionario, è strutturata in più passi, come illustrato schematicamente nella Figura 2.

Prima di procedere al C&C degli errori, è stata effettuata una fase di analisi e ottimizzazione del piano di vincoli (o edit) rispetto ai quali un'Azienda è considerata completa e coerente (*esatta*) oppure errata. In questa fase, sono state effettuate le seguenti operazioni:

- individuazione di errori nella specificazione dei vincoli;
- eliminazione dei vincoli ridondanti e/o contraddittori;
- eliminazione di vincoli che non identificano reali situazioni di errore.

Inoltre, dall'analisi delle frequenze di attivazione dei vincoli (procedura **Proc Editstats** di Banff, vedi Statistics Canada, 2003 e 2005) è stato possibile individuare alcuni errori sistematici che sono stati corretti deterministicamente prima di passare alla fase automatica di individuazione probabilistica degli errori e correzione /imputazione mediante le procedure predefinite in Banff. E' questo il caso della rilevazione di costi per allevamenti alla voce "Spese per alimenti di origine animale" in aziende prive di allevamenti suinicoli riscontrata in oltre 100 unità.

Figura 2. Flusso della procedura – Struttura economica



Sui dati così pre-elaborati (*dati pretrattati*), è stata effettuata la prima fase di localizzazione probabilistica degli errori, utilizzando l'insieme di vincoli lineari riportato nell'Appendice A. Si tratta di 70 vincoli, di cui 23 quadrature e 44 disuguaglianze, che coinvolgono le 126 variabili rilevate nelle sezioni del questionario oggetto del controllo.

Le variabili hanno tutte uguale peso (1 per *default*), il che equivale a dire che, dato un certo sottoinsieme di vincoli violati, tutte le variabili in esse coinvolte hanno la stessa probabilità di essere incluse nella soluzione di minimo cambiamento.

Gli altri parametri utilizzati in questa fase sono:

- la massima cardinalità C ammessa per la soluzione di minimo cambiamento, cioè il massimo numero di variabili selezionabili come errate nella soluzione di minimo cambiamento dell'algoritmo;
- il massimo tempo disponibile T , per ognuno dei record errati, perché l'algoritmo individui una possibile soluzione (entro il limite di massima cardinalità assegnato).

Nella prima fase di localizzazione con metodologia FH sono stati posti $C=nessuna\ limitazione$, $T=50\ secondi$.

Delle 3.675 Aziende attive REA, con *Forma giuridica* = Azienda individuale o Società (di persone o di capitali), 519 sono risultate errate, violando almeno uno dei vincoli del piano di compatibilità.

A causa del limite imposto su C o su T , per un sottoinsieme di unità errate (21 unità) nessuna soluzione di minimo cambiamento è stata individuata. Un secondo passo di localizzazione è stato pertanto effettuato su questo sottoinsieme di dati con nuovi parametri $C=nessuna\ limitazione$, $T=150\ secondi$. Le 19 unità non risolte in questo secondo passo (*Unità non risolte 1 – NRI*) rappresentano il primo gruppo di aziende da sottoporre a revisione manuale per questo gruppo di variabili.

Tutte le unità risolte al primo o al secondo passo di localizzazione degli errori sono state sottoposte a un passo di *correzione con soluzione analitica*: per ogni unità errata, sulla base dei vincoli e delle variabili etichettate come errate dall'algoritmo di localizzazione degli errori, un algoritmo verifica se per qualche variabile errata esiste uno ed un solo valore possibile. Al termine di tale fase alcune unità saranno state completamente risolte (*Unità risolte – RI*), cioè non presenteranno valori errati residui, mentre per altre osservazioni sarà necessario procedere ad un'ulteriore passo di imputazione mediante *NND*. Imputazioni distinte sono state fatte all'interno di strati (*celle di imputazione*) definiti dalla variabile OTE riclassificata. Dalle 9 classi originarie sono state ottenute 5 nuove classi accorpando in un'unica classe OTE 1 (*Aziende specializzate nei seminativi*), 2 (*Aziende specializzate in ortofloricoltura*), 3 (*Aziende specializzate nelle coltivazioni permanenti*), 6 (*Aziende con policoltura*) e 9 (*aziende non classificabili*).

Per ogni strato, il metodo di imputazione è stato applicato secondo la seguente strategia (vedi Figura 2):

1. imputazione con donatore utilizzando: un insieme di edit uguale a quello iniziale, ma con i vincoli di quadratura rilassati del 20%¹. Questo per evitare che la necessità di imputare valori che garantiscano il rispetto delle quadrature possa impedire del tutto l'individuazione di donatori adatti. Le unità per le quali risulta comunque impossibile individuare un donatore costituiscono il secondo gruppo di aziende (*Unità non risolte 2 - NR2*) su cui effettuare una revisione di tipo manuale/interattivo.
2. Per le unità con donatore individuato, nuovo passo di localizzazione rispetto agli edit iniziali per garantire la coerenza dei dati rispetto ai vincoli nella loro forma originaria (in particolare, le quadrature).

¹ Ogni quadratura del tipo $x+y=z$ è stata sostituita dalle due disuguaglianze $x+y \leq 1.8*z$ e $x+y \geq 0.2*z$, che definiscono un intorno di z di ampiezza $\pm 0.4z$.

3. Correzione con soluzione analitica utilizzando l'insieme iniziale di vincoli. Le 31 unità non completamente risolte costituiscono il terzo gruppo di aziende (*Unità non risolte 3 – NR3*) su cui effettuare una revisione di tipo manuale/interattivo.

Al termine di questa procedura, delle 519 unità errate, 469 (*Unità risolte 1-R1 e Unità risolte 2-R2* nella Figura 2) risultano complete e coerenti rispetto all'insieme dei vincoli, mentre le 50 aziende per le quali non è stato possibile completare il processo di controllo (*NR1*) o di imputazione (*NR2 e NR3*) si è intervenuti interattivamente.

4. Trattamento della sezione Occupazione e Redditi

La sezione *Occupazione e Redditi* del questionario è suddivisa in due sotto-sezioni: *Manodopera Familiare* e *Altra Manodopera Aziendale*. Le due sotto-sezioni sono state trattate separatamente, come descritto nei paragrafi 4.2 e 4.3. Nel paragrafo 4.1 sono descritte sinteticamente le metodologie impiegate.

4.1. Metodologie per l'imputazione di errori e mancate risposte nella sezione Occupazione e Redditi

L'imputazione delle mancate risposte parziali e dei valori considerati non accettabili nella sezione occupazione e redditi è stata effettuata utilizzando alcune metodologie implementate nel software *Quis* (Guarnera, 2004; Di Zio *et al.*, 2004). Il software, che opera in ambiente SAS, è stato recentemente sviluppato in Istat al fine di sopperire alla mancanza di un software generalizzato in Istituto che consentisse di applicare alcuni comuni metodi di imputazione delle mancate risposte parziali per variabili numeriche. Tra i metodi disponibili in *Quis* sono stati applicati alla presente indagine il *NND* e la *regressione mediante algoritmo EM*.

4.1.1. Il donatore di distanza minima disponibile in *Quis*

In questo metodo i valori errati o mancanti della variabile obiettivo di una unità i vengono imputati utilizzando il corrispondente valore osservato nella unità j più vicina rispetto a una funzione di distanza scelta fra: *Euclidea*, *Manhattan*, *Minimax*.

La prossimità (o *similitudine*) fra le unità i e j è definita mediante una funzione di distanza calcolata su un prefissato insieme di variabili (dette di *accoppiamento* o di *matching*) scelte dallo statistico e automaticamente standardizzate dal software. Per una certa unità, il metodo effettua l'imputazione simultanea di tutti i valori mancanti delle variabili selezionate. Qualora in un dato record incompleto, alcune variabili di *matching* non siano osservate, il calcolo della distanza viene effettuato sulla base delle sole variabili disponibili. A differenza del metodo del donatore disponibile in *Banff*, quello implementato in *Quis* non tiene conto di eventuali vincoli di coerenza tra le variabili di interesse. Il suo utilizzo è quindi appropriato nei casi in cui non siano presenti edit o nei casi in cui la coerenza del serbatoio dei donatori rispetto ai vincoli garantisce automaticamente la coerenza dei dati imputati. Notiamo inoltre che in *Quis* è possibile includere nel calcolo della funzione di distanza anche variabili categoriche, cosa non possibile nel donatore di *Banff*.

4.1.2. Il metodo di imputazione con regressione mediante EM implementato in *Quis*

Il metodo di imputazione per regressione disponibile in *Quis* è basato sull'assunzione di normalità dei dati. I parametri del modello sono stimati mediante l'algoritmo *Estimation-Maximization (EM)* (Dempster *et al.* 1977). Questo algoritmo consente di effettuare le stime di massima verosimiglianza dei parametri di un modello applicando in modo iterativo i seguenti due passi:

- 1) *E-step*: sono calcolati i valori attesi della funzione di verosimiglianza a dati completi, rispetto alla distribuzione di probabilità dei dati mancanti condizionata ai dati osservati e alle stime correnti dei parametri.
- 2) *M-step*: i valori aggiornati dei parametri sono determinati massimizzando il valore atteso della verosimiglianza calcolato nell' *E-step*.

Nel caso del modello normale l'algoritmo è facilmente implementabile mediante l'uso degli *sweep operator* (Shafer, 1999), una famiglia di operatori che consente, nel caso multinormale, di trasformare i parametri delle distribuzioni congiunte in quelli delle distribuzioni condizionate (e viceversa). I parametri iniziali dell'EM (*starting point*) sono calcolati sui soli dati completamente osservati e l'algoritmo si arresta quando la differenza tra due stime consecutive è sotto una soglia prefissata per tutti i parametri.

Una volta che i parametri sono stati stimati, l'imputazione può essere effettuata in due diversi modi: imputando direttamente i valori attesi dei dati mancanti condizionatamente ai dati osservati (medie predittive), o, alternativamente, aggiungendo alle medie predittive un residuo normale.

4.2. Trattamento della sezione Manodopera familiare

Nella sezione relativa alla **Manodopera familiare** vengono rilevate le seguenti informazioni sulla manodopera familiare impiegata nell'azienda, distintamente per *Tipologia di addetto* (*Conduttore, Coniuge del conduttore, Familiari del conduttore che lavorano in azienda, Parenti del conduttore che lavorano in azienda*):

- *Sesso*,
- *Totale di giornate lavorate nell'anno*,
- *Ore medie lavorate in azienda nella giornata* (in classi),
- *Classi di redditi extraziendali dei componenti la famiglia agricola*,
- *Contributi sociali per la sola attività agricola aziendale*.

In questo ambito, la procedura di C&C ha riguardato la variabile *Totale giornate di lavoro*, (*GL* nel seguito), sulla base delle altre informazioni disponibili per ciascuna azienda.

L'individuazione degli errori e delle mancate risposte su *GL* è stata effettuata sulla base di regole di compatibilità definite dagli esperti di indagine. La tecnica di imputazione adottata è il *NND*, descritto nel paragrafo 4.1.1. Nel caso delle aziende REA, il metodo è stato utilizzato previa stratificazione dei dati. Gli strati (o *celle di imputazione*) sono stati ottenuti utilizzando alcune variabili strutturali o ausiliarie (*covariate* o *variabili di strato*) disponibili per tutto il campione. All'interno di ogni strato, i valori errati o mancanti della variabile obiettivo di una unità *i* sono stati imputati utilizzando il corrispondente valore osservato nella unità *j* più vicina appartenente allo stesso strato di *i*.

La procedura di individuazione degli errori e la loro imputazione si articola nei seguenti passi principali:

1. Preparazione del data set
2. Individuazione degli errori di coerenza e delle mancate risposte
3. Preparazione delle variabili di strato e di matching
4. Imputazione degli errori e delle mancate risposte per le aziende in cui la variabile *Ore medie lavorate in azienda nella giornata* (*HM*) è osservata ($HM \neq missing$)
5. Imputazione degli errori e delle mancate risposte per le aziende in cui la variabile *Ore medie lavorate in azienda nella giornata* non è osservata ($HM = missing$)

4.2.1. Preparazione del dataset

Al fine di applicare la tecnica del *NND*, per ogni azienda si è proceduto alla segmentazione del record fisico corrispondente ad ogni Azienda-campione in più record individuali, corrispondenti

ciascuno ad una unità di manodopera effettivamente impiegata in quell'azienda e contenente tutte le informazioni ad essa relative. In questo modo è stato ottenuto un data set di circa 6.607 unità. Di queste unità sono state selezionate 41 aziende individuali (*Forma Giuridica*=Azienda individuale), in cui mancano tutti i valori relativi al conduttore.

Si noti infine che la procedura ha riguardato le sole Aziende aventi *Forma Giuridica* = Azienda individuale e *Forma Giuridica* = Società di persone e di capitali.

4.2.2. Individuazione degli altri errori e delle mancate risposte

In questa fase sono individuati i valori della variabile *GL* da sottoporre a imputazione. Sulla variabile *GL* sono state ricercate due tipologie di errore: assenza di valore in situazioni in cui il valore dovrebbe essere fornito (sulla base di altre informazioni presenti nella sezione), errore nel valore osservato della variabile.

I valori mancanti di *GL* da sottoporre a ricostruzione mediante imputazione vengono individuati come segue:

- se $HM \neq 0$ e $GL = 0$, allora $GL = missing$
- se [*Tipologia di addetto* = *Conduttore*] e $GL = 0$ allora $GL = missing$

I valori errati di *GL* da sottoporre a ricostruzione mediante imputazione vengono individuati, condizionatamente alla variabile *HM*, sulla base delle seguenti regole (definite tenendo conto dei giorni festivi e del fatto che le giornate lavorate sono da intendersi come giornate di 8 ore):

- se $GL > 335$, allora $GL = missing$
- se $HM < 4$ ore e $GL > 160$, allora $GL = missing$
- se $4 < HM < 8$ ore e $GL > 290$, allora $GL = missing$

In questo modo, sono stati individuati 57 valori mancanti e 763 valori errati, da sottoporre a imputazione.

Nel seguito, con il termine *valore mancante* si farà riferimento a entrambe le tipologie di valori da ricostruire, indipendentemente dalla loro origine (valore osservato ma non accettabile, valore dovuto non osservato).

4.2.3. Preparazione delle variabili di strato e di matching

L'imputazione dei valori classificati come mancanti è stata effettuata all'interno di sottoinsiemi di record individuali (*celle di imputazione*) omogenei rispetto alle variabili *Forma Giuridica* (2 classi, *Forma Giuridica*=1 e *Forma Giuridica*=3) e *Tipologia di addetto* (riclassificata). La variabile *Classe di ore medie lavorate (HM)* è stata utilizzata come ulteriore variabile di stratificazione solo per l'imputazione dei casi in cui essa era osservata.

La variabile di stratificazione *Tipologia di addetto* è stata riclassificata accorpando le tipologie *Familiari del conduttore che lavorano in azienda* e *parenti del conduttore che lavorano in azienda* in una unica classe di addetti: la nuova variabile così ottenuta ha 3 classi.

Per quanto riguarda il calcolo della distanza, le variabili selezionate come variabili di *matching* sono l'*Altimetria* (1=*Pianura*, 2=*Collina*, 3=*Montagna*), la *Superficie Aziendale Totale (SAT)* e la classe di *Unità di Dimensione Economica (UDE)*.

4.2.4. Imputazione

Il passo di imputazione viene effettuato separatamente per le aziende in cui $HM \neq missing$ e quelle in cui $HM = missing$.

In particolare, per le aziende in cui $HM \neq missing$, questa variabile viene utilizzata come variabile di stratificazione nella ricerca del donatore di distanza minima (applicando il relativo modulo di selezione e imputazione del software *Quis*).

Nel caso invece in cui $HM = missing$, le uniche variabili di strato sono la *Forma Giuridica* e la *Tipologia di addetto*, per un totale di 6 celle di imputazione.

In entrambi i casi, le variabili di *matching* sono quelle illustrate nel paragrafo 4.2.3.

Nella Tabella 1 dell'Appendice B sono riportate alcune statistiche relative agli effetti del passo di C&C sulla manodopera familiare.

4.3. Trattamento della sezione Altra Manodopera Aziendale

In questa fase della procedura di C&C sono state trattate le variabili rilevate nella sezione *Altra manodopera Aziendale*, in particolare *Totale giornate di lavoro prestate in azienda*, *Retribuzioni lorde*, *Contributi sociali totali*, *Accantonamenti per il trattamento di fine rapporto*.

Queste variabili sono rilevate separatamente per due diverse categorie di manodopera: *Dipendenti a tempo indeterminato* e *Dipendenti a tempo determinato*. Queste due tipologie di lavoro sono state trattate separatamente, utilizzando diverse metodologie, tenuto conto del fatto che le relazioni statistiche fra le variabili obiettivo possono essere considerate maggiormente stabili (e quindi opportunamente modellizzate) nel caso del tempo indeterminato.

4.3.1. Trattamento dell'Altra manodopera aziendale a tempo indeterminato

Un'analisi esplorativa preliminare ha evidenziato l'esistenza di forti relazioni lineari tra le variabili rilevate in questa sezione. In particolare, utilizzando la scala logaritmica, è risultato ragionevole supporre che la distribuzione congiunta delle tre variabili summenzionate unitamente alla variabile *Accantonamenti per il trattamento di fine rapporto (TFRDTI)*, fosse bene approssimata da una distribuzione normale 4-variata. Si è supposto inoltre che ogni qualvolta almeno una delle variabili *Giornate lavorate (GLDTI)* e/o *Retribuzioni lorde (WDTI)* presentasse valori positivi, anche le restanti dovessero essere non nulle. L'individuazione delle mancate risposte parziali è stata pertanto effettuata in base a tale criterio.

L'assunzione di normalità ha suggerito di effettuare l'imputazione delle mancate risposte mediante algoritmo EM implementato nel software *Quis*. A tale scopo, per migliorare le stime del modello sono stati considerati anche i dati già corretti delle aziende con dimensione economica maggiore o uguale a 4 UDE (rilevate con metodologia RICA), escludendo tuttavia le unità più estreme rispetto alle distribuzioni dei rapporti *WDTI/GLDTI* e *CSDTI/WDTI*. Inoltre, nel serbatoio dei donatori sono state incluse soltanto quelle unità per le quali le 4 variabili di interesse assumono tutte valori maggiori di zero.

Per preservare le caratteristiche distribuzionali dei dati, si è preferita l'opzione che prevede l'imputazione mediante generazione casuale dalla distribuzione stimata (imputazione con residuo), anziché l'utilizzo delle medie predittive.

Nelle tabelle 2 e 2.1 dell'Appendice B sono riportate alcune statistiche relative agli effetti del passo di C&C sull'altra manodopera a tempo indeterminato.

4.3.2. Trattamento dell'Altra manodopera aziendale a tempo determinato

Le particolari caratteristiche del lavoro a tempo determinato implicano che relazioni tra le variabili di analisi si differenzino in modo sostanziale da quelle esistenti tra le stesse variabili per la manodopera a tempo indeterminato. In particolare, a differenza del caso del lavoro a tempo indeterminato, sono stati considerati validi quei record in cui, pur essendo presenti retribuzioni lorde (*WDTD*) e giornate di lavoro (*GLDTD*), risultano nulli gli importi relativi ai contributi sociali (*CSDTD*) e/o al trattamento di fine rapporto (*TFRDTD*). In tali casi cioè non si è ritenuto di dover procedere all'imputazione di queste ultime due variabili. Sono stati invece considerati come mancate risposte i valori nulli di una sola tra le due variabili *GLDTD* e *WDTD* oppure i valori nulli di entrambe queste variabili in presenza di valori positivi di entrambe le variabili *CSDTD* e *TFRDTD*.

La concentrazione di “zeri” sulle variabili *CSDTD* e *TFRDTD* in corrispondenza di valori non nulli delle variabili *GLTD* e *WDTD*, fa sì che la distribuzione congiunta delle variabili di interesse presenti un carattere “semicontinuo” che rende problematico l’utilizzo di un modello parametrico esplicito per l’imputazione delle mancate risposte parziali. Si è pertanto preferito utilizzare il metodo del donatore di minima distanza implementato in Quis. Per avere un serbatoio di donatori più numeroso si è di nuovo utilizzato il data-set delle aziende RICA, opportunamente depurato dalle unità estreme rispetto ai rapporti *WDTDT / GLDTDT* e *CSDTD / WDTDT*. Data l’esiguità del numero di record, si è ritenuto opportuno non stratificare le unità in celle di imputazione. Le variabili di *matching* sono le stesse quattro variabili *WDTDT*, *GLDTDT*, *CSDTD*, *TFRDTD*. Naturalmente, come già osservato in generale nel paragrafo 4.1.1., per ogni record ricevente, nel calcolo della distanza vengono utilizzate solo le variabili disponibili in quel record. Nelle tabelle 3 e 3.1 dell’Appendice B sono riportate alcune statistiche relative agli effetti del passo di C&C sull’altra manodopera aziendale a tempo indeterminato.

5. Conclusioni

La strutturazione di una procedura di controllo in fasi il più specializzate possibile rispetto alla tipologie di errori non campionari presenti in dati di indagine consente sia un maggior controllo del flusso dei dati, sia una maggiore possibilità di aggiornare/monitorare il processo di controllo e correzione.

In questo contesto, l’adozione in ogni fase delle tecniche più appropriate per la tipologia di errore e di variabile oggetto di controllo e correzione in quella fase consente un trattamento ottimale sia dal punto di vista della qualità attesa del risultato, sia dal punto di vista dei tempi e dei costi.

La procedura di controllo e correzione realizzata per l’indagine REA è un esempio di procedura integrata per indagini economiche. In essa, diversi approcci sono utilizzati a seconda non solo della tipologia di errore (errori influenti, valori anomali, errori di compatibilità casuali, mancate risposte parziali), ma anche e soprattutto della tipologia di fenomeni (variabili) oggetto del controllo/imputazione e delle relazioni esistenti fra essi.

Specifici vantaggi dell’utilizzo di software generalizzato sono la riduzione dei costi di sviluppo di nuovo codice, lo sfruttamento di algoritmi complessi altrimenti non utilizzabili, la facilità di documentazione del processo (grazie alla reportistica prodotta automaticamente dal software, e la semplicità di aggiornamento del processo di controllo e correzione dovuta ad eventuali cambiamenti strutturali dell’indagine (stratificazioni, aggiunta/eliminazione di variabili, modifica della struttura interna del questionario, ecc.).

Riferimenti bibliografici

Chen J., Shao J. (2000) Nearest Neighbour Imputation for Survey Data, *Journal of Official Statistics*, **16**, 113-131.

Cotton C. (1991), *Functional description of the generalized edit and imputation system*, Statistics Canada, Business Survey Methods Division, July.

Barcaroli, Ceccarelli, Luzi, 1995, Il Macroediting - Tecniche di correzione interattiva di variabili quantitative guidata dall’analisi degli aggregati: il caso del Sistema dei Conti delle Imprese, Quaderni di Ricerca ISTAT, N. 1/1998

- Di Zio M., Guarnera U., Luzi O., Manzari A. (2004) Methods and software for editing and imputation: Recent advancement at ISTAT, *Rivista di Statistica Ufficiale - Quaderni di Ricerca ISTAT*, N. 2/2004
- Di Zio M., Luzi O. (2002b), Combining Methodologies in a Data Editing Procedure: an Experiment on the Survey of Balance Sheets of Agricultural Firms, *Statistica Applicata*, **14**, 1, pp. 59-80.
- Fellegi I.P., Holt, T.D. (1976), A Systematic Approach to Edit and Imputation, *Journal of the American Statistical Association*, **71**, 17-35.
- Granquist, 1992, A Review of methods for rationalizing the editing of survey data, "Statistical Data Editing Methods and Techniques" - United Nations, Vol. I
- Guarnera U. (2004) Alcuni metodi di imputazione delle mancate risposte parziali per dati quantitativi. Il software QUIS, *Contributi di Ricerca ISTAT*, N. 5/2004
- Guarnera U., Luzi O., Tommasi I. (2006) "Metodi Parametrici e non parametrici per la ricostruzione dei valori mancanti nell'indagine RICA-REA", *Atti del Convegno "Le statistiche agricole verso il Censimento del 2010: valutazioni e prospettive"*, Università di Cassino, 26-27 Ottobre.
- Lessler J.T., Kalsbeek W.D. (1992), *Non Sampling Errors in Surveys*, New York: Wiley.
- Statistics Canada (2003) *Banff - Functional Description of the Banff System for Editing and Imputation, Version 1.02*, Generalized Systems Methods Section, Business Survey Methods Division, December 2003.
- Statistics Canada (2005) *Banff Users Guide, Version 1.04*, Generalized Systems Methods Section, Business Survey Methods Division.
- Kalton G., Kasprzyk D. (1982) Imputing for Missing Survey Responses, *Proceedings of the Section on Survey Research Methods*, American Statistical Association, 22-31.
- Kovar J.G., MacMillan J., Whitridge P. (1988), Overview and Strategy for the Generalized Edit and Imputation System. *Statistics Canada, Methodology Branch Working Paper No. BSMD-88-007E*
- Schafer J. L., (1999) *Analysis of Incomplete Multivariate Data*. Chapman & Hall, London.

Appendice A: Vincoli di coerenza utilizzati per il controllo probabilistico

Vincoli di quadratura

SEZIONE II – VALORI ECONOMICI DELLE AZIENDE AGRICOLE

12. Struttura dei costi

12.1 Spese per acquisto di beni e servizi per le coltivazioni

Cod. 009 = cod 010+cod 020+cod 030+cod 040

Cod 040 = cod 041+cod 042+cod 043

12.2 Spese per acquisto di beni e servizi per gli allevamenti

Cod. 049 = cod 050+cod 060+cod 070+cod 080+cod 090

Cod 050 = cod 051+cod 052+cod 053

Cod 080= cod 081+cod 082

Cod 090 = cod 091+cod 092+cod 093

12.3 Spese per la meccanizzazione

Cod. 099 = cod 100+cod 110+cod 120+cod 130

Cod. 120 = cod 121+cod 122+cod 123

12.4 Spese generali e varie

Cod. 139 = cod 140+cod 150+cod 160+cod 170+ cod 180+cod 190+cod 200+cod 210+cod
220+cod 230+cod 240+cod 250+cod 260+cod 270

Cod. 160 = cod 161+cod 162

Cod. 210 = cod 211+cod 212

Cod. 230 = cod 231+cod 232+cod 233+cod 234

Cod. 270 = cod 271+cod 272+cod 273+cod 274+cod 275+cod 276

Totale costi

Cod. 300 = cod 009+cod 049+cod 099+cod 139+cod 280+cod 290

13. Giacenze e scorte

Cod. 350 = cod 310+cod 320+cod 330+cod 340

Cod. 400 = cod 360+cod 370+cod 380+cod 390

14. Acquisti e vendite di capitali fissi nell'esercizio

14.1 Capitali fissi

Cod. 500 = cod 410+cod 420+cod 430+cod 440+ cod 450+cod 460+cod 470+cod 480+cod 490

Cod. 570 = cod 510+cod 520+cod 530+cod 540+ cod 550+cod 560

15. Reimpieghi

Cod 600 = cod 580+cod 590

16. Ricavi e autoconsumo

16.1 Ricavi

Cod 609 = cod 610+cod 620+cod 630+cod 640+cod 650+cod 670+cod 680+cod 690+cod 700+cod 710

Totale Ricavi

Cod 720 = cod 609+cod 660

17. Contributi

Cod 840 = cod 730+cod 770+cod 780+cod 790+cod 800+cod 810+cod 820+cod 830

Cod 730 = cod 740+cod 750+cod 760

Altri vincoli

SEZIONE I – AGGIORNAMENTO NOTIZIE GENERALI SULL'AZIENDA

5. Titolo di possesso della superficie totale

Cod 5.4>0

6. Titolo di possesso della SAU

Se cod 6.4>0 allora cod 6.4 <= cod 5.4 (sez. I punto 5)

7. Utilizzazione della superficie aziendale

Se cod. 7.1>0 deve esistere cod 6.4>0
Se cod. 7.1>0 allora cod7.1<= cod 6.4

8. Allevamenti

Se (cod. 8.1+cod. 8.2+cod. 8.3+cod. 8.4+cod 8.5+cod 8.6)>0 deve esistere:
e/o alla Sez. 12.2 il cod. 049;
e/o alla Sez. 13. cod. 310;
e/o alla Sez. 15. cod. 600;

SEZIONE II – VALORI ECONOMICI DELLE AZIENDE AGRICOLE

12. Struttura dei costi

Cod 300>0;

12.1 Spese per acquisto di beni e servizi per le coltivazioni

Se cod. 009>0 deve esistere:
e/o alla Sez. 13.1 il cod. 360;
e/o alla Sez. 13.4 il cod 390;
e/o alla Sez. 15.1 il cod. 580;
e/o alla Sez. 16.1 il cod. 610 e/o cod 620;
e/o alla Sez. 16.2 il cod. 660;

12.2 Spese per acquisto di beni e servizi per gli allevamenti

Se cod 049>0 deve esistere:
e/o alla Sez. 8 il cod. 8.1 e/o cod 8.2 e/o cod 8.3 e/o cod 8.4 e/o cod 8.5 e/o cod 8.6;

Se cod 049>0 deve esistere:
e/o alla Sez. 13.2 il cod. 370;
e/o alla Sez. 13.3 il cod 380;
e/o alla Sez. 13.4 i1 il cod. 390;
e/o alla Sez. 14.1 il cod. 420 e/o cod 520;
e/o alla Sez. 15.2 il cod. 590;
e/o alla Sez. 16.1 il cod. 630 e/o cod 640 e/o cod 650;
e/o alla Sez. 16.2 il cod. 660;

Se cod 052>0 deve esistere:
alla Sez. 8 il cod 8.2;

12.3 Spese generali e varie

Se (cod 140+cod 150)>0 deve esistere:
e/o alla Sez. 12.1 il cod. 009;
e/o alla Sez. 12.2 il cod 049;
e/o alla Sez. 14.1 il cod. 440 e/o cod 450 e/o cod 530 e/o cod 540;
e/o alla Sez. 15. il cod. 600;
e/o alla Sez. 16.1 il cod. 609;

Se cod 160>0 deve esistere:
e/o alla Sez. 12.4 il cod. 210 e/o cod 250;
e/o alla Sez. 13.4 il cod 390;
e/o alla Sez. 14.1 il cod. 420 e/o cod 430 e/o cod 440 e/o cod 450 e/o cod 490
e/o cod 520 e/o cod 530 e/o cod 540 e/o cod 550 e/o cod 560;
e/o alla Sez. 15. il cod. 600;
e/o alla Sez. 16.1 (cod. 609-cod 700-cod 710)>0;

Se cod 170>0 deve esistere:
e/o alla Sez. 12.1 il cod. 009;
e/o alla Sez. 12.2 il cod 049;
e/o alla Sez. 12.3 il cod 099;
e/o alla Sez. 14.1 il cod. 500 e/o cod 570;

Se cod 180>0 deve esistere:
e/o alla Sez. 12.1 il cod. 009;
e/o alla Sez. 12.2 il cod 049;
e/o alla Sez. 14.1 il cod. 410 e/o cod 420 e/o cod 460 e/o cod 470 e/o cod 490;

Se cod 190>0 deve esistere:
e/o alla Sez. 12.1 il cod. 009;
e/o alla Sez. 12.2 il cod 049;
e/o alla Sez. 12.5 il cod 280;
e/o alla Sez. 14.1 il cod. 500 e/o cod 570;

Se cod 200>0 deve esistere:
e/o alla Sez. 12.1 il cod. 009;
e/o alla Sez. 12.2 il cod 049;
e/o alla Sez. 12.3 il cod 099;
e/o alla Sez. 16.1 il cod. 609;

Se cod 210>0 deve esistere:
e/o alla Sez. 12.1 il cod. 009;
e/o alla Sez. 12.2 il cod 049;
e/o alla Sez. 12.4 il cod 230;
e/o alla Sez. 16.1 il cod. 609;

Se cod 220>0 deve esistere:
e/o alla Sez. 12.1 il cod. 009;
e/o alla Sez. 12.2 il cod 049;
e/o alla Sez. 12.4 il cod 140 e/o cod 150 e/o cod 160 e/o cod 230;
e/o alla Sez. 12.6 il cod 290;
e/o alla Sez. 13.1 il cod 310 e/o cod 360;
e/o alla Sez. 13.2 il cod 320 e/o cod 370;
e/o alla Sez. 14.1 il cod 430 e/o cod 440 e/o cod 490
e/o cod 530 e/o cod 540 e/o cod 560;
e/o alla Sez. 15 il cod 600;
e/o alla Sez. 16.1 il cod. 610 e/o cod 620 e/o cod 650 e/o cod 670;
e/o alla Sez. 16.2 il cod 660;

12.4 Interessi passivi

Se cod 280>0 deve esistere:
e/o alla Sez. 12.1 il cod. 009;
e/o alla Sez. 12.2 il cod 049;
e/o alla Sez. 12.3 il cod 099;
e/o alla Sez. 12.4 il cod 220 e/o cod 230 e/o cod 240 e/o cod 260;
e/o alla Sez. 14.1 il cod 500;

12.5 Imposte indirette

Se cod 290>0 deve esistere:
e/o alla Sez. 12.3 il cod 100 e/o cod 110 e/o cod 120 e/o cod 130;
e/o alla Sez. 12.4 il cod 210 e/o cod 220 e/o cod 260;
e/o alla Sez. 14.1 il cod 420 e/o cod 430 e/o cod 440 e/o cod 450 e/o cod 470
e/o cod 520 e/o cod 530 e/o cod 540 e/o cod 550;
e/o alla Sez. 16.1 il cod. 610 e/o cod 620 e/o cod 630 e/o cod 640 e/o cod 650 e/o cod 670;

13. Giacenze e scorte

Se cod 310>0 deve esistere:
e/o alla Sez. 13.1 il cod 360;
e/o alla Sez. 15.1 il cod 580;
e/o alla Sez. 16.1 il cod. 610 e/o cod 620 e/o cod 670;
e/o alla Sez. 16.2 cod 660;

Se cod 320>0 deve esistere:
e/o alla Sez. 13.2 il cod 370;
e/o alla Sez. 15.2 il cod 590;
e/o alla Sez. 16.1 il cod. 640 e/o cod 650 e/o cod 670;
e/o alla Sez. 16.2 cod 660;

Se cod 330>0 deve esistere:
e/o alla Sez. 13.3 il cod 380;
e/o alla Sez. 16.1 il cod. 630;

Se cod 340>0 deve esistere:
e/o alla Sez. 12.1 il cod. 009;
e/o alla Sez. 12.2 il cod 049;
e/o alla Sez. 12.3 il cod 099;
e/o alla Sez. 13.4 il cod 390;

Se cod 360>0 deve esistere:
e/o alla Sez. 12.1 il cod. 009;
e/o alla Sez. 12.3 il cod 099;
e/o alla Sez. 12.4 il cod 140 e/o cod 150 e/o cod 160 e/o cod 220 e/o cod 230;
e/o alla Sez. 13.1 il cod 310;

Se cod 370>0 deve esistere:
e/o alla Sez. 12.2 il cod. 049;
e/o alla Sez. 12.3 il cod 099;
e/o alla Sez. 12.4 il cod 140 e/o cod 150 e/o cod 160 e/o cod 220 e/o cod 230;
e/o alla Sez. 13.2 il cod 320;

Se cod 380>0 deve esistere:
e/o alla Sez. 12.2 il cod. 049;
e/o alla Sez. 12.4 il cod 140 e/o cod 150 e/o cod 160 e/o cod 170 e/o cod 230;
e/o alla Sez. 13.3 il cod 330;

Se cod 390>0 deve esistere:
e/o alla Sez. 12.1 il cod. 009;

e/o alla Sez. 12.2 il cod 049;
e/o alla Sez. 12.3 il cod 099;
e/o alla Sez. 13.4 il cod 340;

14. Acquisti e vendite di capitali fissi nell'esercizio

Se cod 410>0 deve esistere:
e/o alla Sez. 12.1 il cod. 009;
e/o alla Sez. 12.3 il cod 099;
e/o alla Sez. 12.4 il cod 140 e/o cod 170 e/o cod 180 e/o cod 190 e/o cod 240;
e/o alla Sez. 12.5 il cod 280;
e/o alla Sez. 14.1 il cod 480;
e/o alla Sez. 15.1 il cod 580;
e/o alla Sez. 16.1 il cod. 610 e/o cod 620;

Se cod 420>0 deve esistere:
e/o alla Sez. 12.2 il cod. 049;
e/o alla Sez. 12.3 il cod 099;
e/o alla Sez. 12.4 il cod 140 e/o cod 150 e/o cod 160 e/o cod 170 e/o cod 180 e/o cod 190 e/o cod 230 e/o cod 240;
e/o alla Sez. 12.5 il cod 280;
e/o alla Sez. 14.1 il cod 430 e/o cod 480;
e/o alla Sez. 15.2 il cod 590;
e/o alla Sez. 16.1 il cod. 640 e/o cod 650;

Se (cod 430+cod 440 cod 450)>0 deve esistere:
e/o alla Sez. 12.1 il cod. 009;
e/o alla Sez. 12.2 il cod. 049;
e/o alla Sez. 12.3 il cod 099;
e/o alla Sez. 12.4 il cod 140 e/o cod 150 e/o cod 160 e/o cod 170 e/o cod 190 e/o cod 230;
e/o alla Sez. 12.5 il cod 280;
e/o alla Sez. 16. il cod. 720;

Se cod 460>0 deve esistere:
e/o alla Sez. 12.3 il cod 099;
e/o alla Sez. 12.4 il cod 180 e/o cod 190 e/o cod 240;
e/o alla Sez. 12.5 il cod 280;

Se cod 470>0 deve esistere:
e/o alla Sez. 12.3 il cod 099;
e/o alla Sez. 12.5 il cod 280;
e/o alla Sez. 14.1 il cod 410 e/o cod 430 e/o cod 440 e/o cod 450 e/o cod 460;
e/o alla Sez. 15.2 il cod 590;
e/o alla Sez. 16.1 il cod. 640 e/o cod 650;

Se cod 480>0 deve esistere:
e/o alla Sez. 12.4 il cod 190;
e/o alla Sez. 12.5 il cod 280;
e/o alla Sez. 14.1 il cod 410 e/o cod 420 e/o cod 430 e/o cod 440 e/o cod 450 e/o cod 490;

Se (cod 540+cod 550)>0 deve esistere:
e/o alla Sez. 12.3 il cod 099;
e/o alla Sez. 12.4 il cod 170 e/o cod 230;
e/o alla Sez. 14.1 il cod 470;

15. Reimpieghi

Se cod 580>0 deve esistere:
e/o alla Sez. 12.1 il cod. 009;
e/o alla Sez. 12.3 il cod 099;
e/o alla Sez. 12.4 il cod 140 e/o cod 150 e/o cod 160 e/o cod 220
e/o cod 230 e/o cod 240 e/o cod 250 e/o cod 260 e/o cod 270;

Se cod 590>0 deve esistere:
e/o alla Sez. 12.2 il cod. 049;
e/o alla Sez. 12.3 il cod 099;
e/o alla Sez. 12.4 il cod 140 e/o cod 150 e/o cod 160 e/o cod 220
e/o cod 230 e/o cod 240 e/o cod 250 e/o cod 260 e/o cod 270;

16. Ricavi e autoconsumo

Se (cod 610+cod 620)>0 deve esistere:
e/o alla Sez. 12.1 il cod. 009;
e/o alla Sez. 12.3 il cod 099;
e/o alla Sez. 12.4 il cod 140 e/o cod 150 e/o cod 160 e/o cod 170 e/o cod 180 e/o cod 190
e/o cod 200 e/o cod 210 e/o cod 220 e/o cod 230 e/o cod 240 e/o cod 250 e/o cod 260;
e/o alla Sez. 12.6 il cod 290;
e/o alla Sez. 13.1 il cod.310;
e/o alla Sez. 13.4 il cod 340;

Se (cod 630+cod 640+cod 650)>0 deve esistere:

e/o alla Sez. 12.2 il cod. 049;

e/o alla Sez. 12.3 il cod 099;

e/o alla Sez. 12.4 il cod 140 e/o cod 150 e/o cod 160 e/o cod 170 e/o cod 180
e/o cod 200 e/o cod 210 e/o cod 220 e/o cod 230 e/o cod 250 e/o cod 260;

e/o alla Sez. 12.6 il cod 290;

e/o alla Sez. 13.2 il cod.320;

e/o alla Sez. 13.3 il cod 330;

e/o alla Sez. 13.4 il cod 340;

Se cod 660>0 deve esistere:

e/o alla Sez. 12.1 il cod. 009;

e/o alla Sez. 12.2 il cod. 049;

e/o alla Sez. 12.3 il cod 099;

e/o alla Sez. 12.4 il cod 140 e/o cod 150 e/o cod 160 e/o cod 170 e/o cod 180
e/o cod 200 e/o cod 210 e/o cod 220 e/o cod 230 e/o cod 250 e/o cod 260;

Se cod 670>0 deve esistere:

e/o alla Sez. 12. il cod 300;

e/o alla Sez. 13. il cod 350;

Se cod 680>0 deve esistere:

e/o alla Sez. 12.3 il cod 099;

e/o alla Sez. 12.4 il cod 170 e/o cod 250;

Se cod 700>0 deve esistere:

e/o alla Sez. 12.1 il cod 009;

e/o alla Sez. 12.2 il cod 049;

e/o alla Sez. 12.3 il cod 099;

17. Contributi

Se cod 750>0 deve esistere:

e/o alla Sez. 8 il cod. 8.1 e/o cod 8.2 e/o cod 8.3 e/o cod 8.4 e/o cod 8.5 e/o cod 8.6;

Appendice B: Manodopera familiare e Altra manodopera aziendale: Indicatori sugli effetti del processo di controllo e correzione dati

Tab.1 - Effetti della procedura di controllo e correzione applicata alla sezione "Occupazione e redditi - Manodopera familiare"

Regione	Aziende	Giornate di lavoro familiare rilevate	Giornate di lavoro familiare corrette	Differenza	%	Giornate di lavoro familiare stimate sui dati rilevati		Giornate di lavoro familiare stimate post-correzione		Differenza	%
						Giornate di lavoro familiare stimate sui dati rilevati	Differenza	Giornate di lavoro familiare stimate post-correzione	Differenza		
Piemonte	116	26.877	19.355	-7.522	-28,0	3.728.161	3.280.008	-448.153	-12,0		
V. Aosta	79	13.343	13.343	0	0,0	538.377	538.377	0	0,0		
Lombardia	54	12.529	8.238	-4.291	-34,2	4.769.638	3.805.757	-963.881	-20,2		
Trentino A.A.	205	31.347	25.091	-6.256	-20,0	3.234.709	2.584.618	-650.091	-20,1		
Veneto	498	46.243	38.699	-7.544	-16,3	5.753.502	5.518.952	-234.551	-4,1		
Friuli V.G.	122	20.803	18.229	-2.574	-12,4	1.260.130	1.392.402	132.272	10,5		
Liguria	176	45.376	38.094	-7.282	-16,0	4.547.141	3.738.964	-808.177	-17,8		
Emilia R.	184	72.860	57.930	-14.930	-20,5	8.852.460	7.155.147	-1.697.313	-19,2		
Toscana	110	19.159	16.868	-2.291	-12,0	6.811.759	5.798.289	-1.013.470	-14,9		
Umbria	86	13.531	10.990	-2.541	-18,8	2.469.988	1.899.319	-570.668	-23,1		
Marche	89	13.359	10.963	-2.396	-17,9	5.317.313	3.831.288	-1.486.025	-27,9		
Lazio	267	56.544	44.461	-12.083	-21,4	14.704.696	11.152.892	-3.551.805	-24,2		
Abruzzo	334	73.562	66.096	-7.466	-10,1	6.519.936	5.857.487	-662.449	-10,2		
Molise	80	14.891	10.631	-4.260	-28,6	1.535.660	1.195.760	-339.900	-22,1		
Campania	230	43.809	38.143	-5.666	-12,9	18.223.812	16.403.668	-1.820.144	-10,0		
Puglia	120	7.992	7.267	-725	-9,1	9.270.633	8.031.814	-1.238.819	-13,4		
Basilicata	137	26.156	19.987	-6.169	-23,6	6.923.177	6.453.414	-469.763	-6,8		
Calabria	193	21.112	17.618	-3.494	-16,5	10.675.652	9.836.293	-839.359	-7,9		
Sicilia	164	16.550	13.379	-3.171	-19,2	10.179.865	8.454.231	-1.725.635	-17,0		
Sardegna	431	144.960	118.931	-26.029	-18,0	6.418.564	6.141.716	-276.848	-4,3		
ITALIA	3675	721.003	594.313	-126.690	-17,6	131.735.174	113.070.395	18.664.779	-14,2		

Tab.2 - Effetti della procedura di controllo e correzione applicata alla sezione "Occupazione e redditi - Altra Manodopera a tempo indeterminato" (dati osservati)

Regione	Aziende	Giornate di lavoro manodopera tempo indeterminato rilevate		Retribuzioni lorde dipendenti tempo indeterminato rilevate		Giornate di lavoro manodopera tempo indeterminato corrette		Retribuzioni lorde dipendenti tempo indeterminato corrette		Differenza retribuzioni lorde	Differenza % retribuzioni lorde
		manodopera tempo indeterminato rilevate	tempo rilevate	lorde dipendenti tempo indeterminato rilevate	tempo rilevate	manodopera tempo indeterminato corrette	tempo corrette	lorde dipendenti tempo indeterminato corrette	tempo corrette		
Piemonte	116	18595	1103832	18595	1103832	0	0	0	0	0,0	0,0
V. Aosta	79	0	0	0	0	0	0	0	0	0,0	0,0
Lombardia	54	26880	339104	26880	339104	0	0	0	0	0,0	0,0
Trentino A.A.	205	1467	0	1467	73096	0	73096	0	73.096		
Veneto	498	37052	2232558	37052	2265458	0	32.900	0	32.900	1,5	
Friuli V.G.	122	2730	156938	2730	156938	0	0	0	0	0,0	0,0
Liguria	176	0	0	0	0	0	0	0	0	0,0	0,0
Emilia R.	184	74755	6204446	74755	6204446	0	0	0	0	0,0	0,0
Toscana	110	38597	3140936	38597	3140936	0	0	0	0	0,0	0,0
Umbria	86	0	0	0	0	0	0	0	0	0,0	0,0
Marche	89	0	0	0	0	0	0	0	0	0,0	0,0
Lazio	267	200	7000	200	7000	0	0	0	0	0,0	0,0
Abruzzo	334	3150	151461	3150	151461	0	0	0	0	0,0	0,0
Molise	80	0	0	0	0	0	0	0	0	0,0	0,0
Campania	230	90	2600	90	2600	0	0	0	0	0,0	0,0
Puglia	120	300	20814	300	20814	0	0	0	0	0,0	0,0
Basilicata	137	0	0	0	0	0	0	0	0	0,0	0,0
Calabria	193	162	2941	162	2941	0	0	0	0	0,0	0,0
Sicilia	164	0	0	0	0	0	0	0	0	0,0	0,0
Sardegna	431	87856	4863219	87856	5075807,309	0	212.588	0	212.588	4,4	
ITALIA	3675	291834	18225849	291834	18544433,31	0	318.584	0	318.584	1,7	

Tab.2.1 - Effetti della procedura di controllo e correzione applicata alla sezione "Occupazione e redditi - Altra Manodopera a tempo indeterminato" (dati stimati)

Regione	Aziende	Giornate di lavoro manodopera		Retribuzioni lorde dipendenti		Giornate di lavoro manodopera		Retribuzioni lorde dipendenti		Differenza % retribuzioni lorde
		indeterminato stimato su dati osservati	tempo	indeterminato stimato su dati osservati	tempo	indeterminato stimato su dati corretti	tempo	indeterminato stimato su dati corretti	Differenza retribuzioni lorde	
Piemonte	116	251.126	15.115.754	251.126	15.115.754	0	0	0	0	0,0
V. Aosta	79	0	0	0	0	0	0	0	0	0,0
Lombardia	54	29.450	505.075	29.450	505.075	0	0	0	0	0,0
Trentino A.A.	205	4.084	0	4.084	193.852	0	193.852	0	193.852	0,9
Veneto	498	128.767	8.887.864	128.767	8.964.785	0	76.921	0	76.921	0,0
Friuli V.G.	122	22.870	1.283.506	22.870	1.283.506	0	0	0	0	0,0
Liguria	176	0	0	0	0	0	0	0	0	0,0
Emilia R.	184	221.479	17.301.448	221.479	17.301.448	0	0	0	0	0,0
Toscana	110	190.970	15.025.516	190.970	15.025.516	0	0	0	0	0,0
Umbria	86	0	0	0	0	0	0	0	0	0,0
Marche	89	0	0	0	0	0	0	0	0	0,0
Lazio	267	23.199	811.950	23.199	811.950	0	0	0	0	0,0
Abruzzo	334	14.233	697.127	14.233	697.127	0	0	0	0	0,0
Molise	80	0	0	0	0	0	0	0	0	0,0
Campania	230	62.851	1.815.706	62.851	1.815.706	0	0	0	0	0,0
Puglia	120	1.763	122.338	1.763	122.338	0	0	0	0	0,0
Basilicata	137	0	0	0	0	0	0	0	0	0,0
Calabria	193	477.245	8.664.056	477.245	8.664.056	0	0	0	0	0,0
Sicilia	164	0	0	0	0	0	0	0	0	0,0
Sardegna	431	169.038	7.670.178	169.038	8.007.481	0	337.303	0	337.303	4,4
ITALIA	3675	1.597.074	77.900.518	1.597.074	78.508.594	0	608.076	0	608.076	0,8

Tab.3 - Effetti della procedura di controllo e correzione applicata alla sezione "Occupazione e redditi - Altra Manodopera a tempo determinato" (dati osservati)

Regione	Aziende	Giornate di lavoro manodopera tempo determinato rilevate		Retribuzioni dipendenti tempo determinato rilevate		Giornate di lavoro manodopera tempo determinato corrette		Retribuzioni dipendenti tempo determinato corrette		Differenza giornate di lavoro	Differenza % giornate di lavoro	Differenza retribuzioni lorde	Differenza % retribuzioni lorde
		manodopera tempo determinato rilevate	lavoro rilevate	Retribuzioni dipendenti tempo determinato rilevate	lavoro rilevate	manodopera tempo determinato corrette	lavoro corrette	Retribuzioni dipendenti tempo determinato corrette	lavoro corrette				
Piemonte	116	230	6.880	230	9.737	0	0,0	2.857	41,5				
V. Aosta	79	380	20.000	380	20.000	0	0,0	0	0,0				
Lombardia Trentino A.A.	54	2.700	18.000	2.700	163.000	0	0,0	145.000	805,6				
Veneto	205	788	0	788	23.640	0	0,0	23.640	0,0				
Friuli V.G.	498	6.444	514.400	6.444	514.400	0	0,0	0	0,0				
Liguria	122	2.867	111.859	2.867	111.859	0	0,0	0	0,0				
Emilia R.	176	0	0	0	0	0	0,0	0	0,0				
Toscana	184	112.641	8.006.910	113.271	8.006.910	630	0,6	0	0,0				
Umbria	110	30.690	1.406.590	30.690	1.406.590	0	0,0	0	0,0				
Marche	86	0	0	0	0	0	0,0	0	0,0				
Lazio	89	0	0	0	0	0	0,0	0	0,0				
Abruzzo	267	1.691	250	1.691	66.488	0	0,0	66.238	26495,2				
Molise	334	1.513	96.318	1.513	96.318	0	0,0	0	0,0				
Campania	80	2.330	25.270	2.330	25.270	0	0,0	0	0,0				
Puglia	230	1.043	36.654	1.043	38.694	0	0,0	2.040	5,6				
Basilicata	120	4.862	158.481	4.862	158.481	0	0,0	0	0,0				
Calabria	137	550	18.000	550	18.000	0	0,0	0	0,0				
Sicilia	193	734	22.200	734	22.200	0	0,0	0	0,0				
Sardegna	164	676	26.085	676	26.695	0	0,0	610	0,0				
ITALIA	431	52.201	2.287.477	52.201	2.294.327	0	0,0	6.850	0,3				
	3675	222.340	12.755.374	222.970	13.002.609	630	0,3	247.235	1,9				

Tab.3.1 - Effetti della procedura di controllo e correzione applicata alla sezione "Occupazione e redditi - Altra Manodopera a tempo determinato" (dati stimati)

Regione	Aziende	Giornate di lavoro manodopera		Retribuzioni lorde dipendenti		Giornate di lavoro manodopera		Retribuzioni lorde dipendenti		Differenza % giornate di lavoro	Differenza retribuzioni lorde	Differenza % retribuzioni lorde
		determinato stimato su dati osservati	tempo determinato stimato su dati osservati	determinato stimato su dati osservati	tempo determinato stimato su dati osservati	determinato stimato su dati osservati	tempo determinato stimato su dati osservati	Differenza giornate di lavoro	Differenza retribuzioni lorde			
Piemonte	116	4.406	116.738	4.406	189.194	0	0,0	0	72.457	0,0	72.457	62,1
V. Aosta	79	1.099	57.827	1.099	57.827	0	0,0	0	0	0,0	0	0,0
Lombardia Trentino A.A.	54	2.887	42.093	2.887	187.093	0	0,0	0	145.000	0,0	145.000	344,5
Veneto	205	27.363	0	27.363	820.881	0	0,0	0	820.881	0,0	820.881	0,0
Friuli V.G.	498	25.478	2.243.898	25.478	2.243.898	0	0,0	0	0	0,0	0	0,0
Liguria	122	24.315	941.612	24.315	941.612	0	0,0	0	0	0,0	0	0,0
Emilia R.	176	0	0	0	0	0	0,0	0	0	0,0	0	0,0
Toscana	184	313.960	23.324.264	314.275	23.324.264	315	0,1	315	0	0,0	0	0,0
Umbria	110	166.329	7.774.417	166.329	7.774.417	0	0,0	0	0	0,0	0	0,0
Marche	86	0	0	0	0	0	0,0	0	0	0,0	0	0,0
Lazio	89	0	0	0	0	0	0,0	0	0	0,0	0	0,0
Abruzzo	267	226.206	250.789	226.206	8.762.705	0	0,0	0	8.511.916	0,0	8.511.916	3394,0
Molise	334	13.607	704.733	13.607	704.733	0	0,0	0	0	0,0	0	0,0
Campania	80	14.567	314.496	14.567	314.496	0	0,0	0	0	0,0	0	0,0
Puglia	230	197.326	9.259.870	197.326	9.288.804	0	0,0	0	28.934	0,0	28.934	0,3
Basilicata	120	314.364	12.274.235	314.364	12.274.235	0	0,0	0	0	0,0	0	0,0
Calabria	137	12.991	402.768	12.991	402.768	0	0,0	0	0	0,0	0	0,0
Sicilia	193	699.186	19.205.989	699.186	19.205.989	0	0,0	0	0	0,0	0	0,0
Sardegna	164	202.259	7.964.669	202.259	8.319.129	0	0,0	0	354.460	0,0	354.460	0,0
ITALIA	431	381.808	12.480.956	381.808	12.717.539	0	0,0	0	236.583	0,0	236.583	1,9
	3675	2.628.152	97.359.354	2.628.467	107.529.584	315	0,0	315	10.170.230	0,0	10.170.230	10,4

Contributi ISTAT(*)

- 1/2002 - Francesca Biancani, Andrea Carone, Rita Pistacchio e Giuseppina Ruocco - *Analisi delle imprese individuali*
- 2/2002 - Massimiliano Borgese - *Proposte metodologiche per un progetto d'indagine sul trasporto aereo alla luce della recente normativa comunitaria sul settore*
- 3/2002 - Nadia Di Veroli e Roberta Rizzi - *Proposta di classificazione dei rapporti di lavoro subordinato e delle attività di lavoro autonomo: analisi del quadro normativo*
- 4/2002 - Roberto Gismondi - *Uno stimatore ottimale in presenza di non risposte*
- 5/2002 - Maria Anna Pennucci - *Le strategie europee per l'occupazione dal Libro bianco di Delors al Consiglio Europeo di Cardiff*
- 1/2003 - Giovanni Maria Merola - *Safety Rules in Statistical Disclosure Control for Tabular Data*
- 2/2003 - Fabio Bacchini, Pietro Gennari e Roberto Iannaccone - *A new index of production for the construction sector based on input data*
- 3/2003 - Fulvia Ceroni e Enrica Morganti - *La metodologia e il potenziale informativo dell'archivio sui gruppi di impresa: primi risultati*
- 4/2003 - Sara Mastrovita e Isabella Siciliani - *Effetti dei trasferimenti sociali sulla distribuzione del reddito nei Paesi dell'Unione europea: un'analisi dal Panel europeo sulle famiglie*
- 5/2003 - Patrizia Cella, Giuseppe Garofalo, Adriano Paggiaro, Nicola Torelli e Caterina Viviano - *Demografia d'impresa: l'utilizzo di tecniche di abbinamento per l'analisi della continuità*
- 6/2003 - Enrico Grande e Orietta Luzi - *Metodologie per l'imputazione delle mancate risposte parziali: analisi critica e soluzioni disponibili in Istat*
- 7/2003 - Stefania Fivizzani, Annalisa Lucarelli e Marina Sorrentino - *Indagine sperimentale sui posti di lavoro vacanti*
- 8/2003 - Mario Adua - *L'agricoltura di montagna: le aziende delle donne, caratteristiche agricole e socio-rurali*
- 9/2003 - Franco Mostacci e Roberto Sabbatini - *L'euro ha creato inflazione? Changeover e arrotondamenti dei prezzi al consumo in Italia nel 2002*
- 10/2003 - Leonello Tronti - *Problemi e prospettive di riforma del sistema pensionistico*
- 11/2003 - Roberto Gismondi - *Tecniche di stima e condizioni di coerenza per indagini infraannuali ripetute nel tempo*
- 12/2003 - Antonio Frenda - *Analisi delle legislazioni e delle prassi contabili relative ai gruppi di imprese nei paesi dell'Unione Europea*
- 1/2004 - Marcello D'Orazio, Marco Di Zio e Mauro Scanu - *Statistical Matching and the Likelihood Principle: Uncertainty and Logical Constraints*
- 2/2004 - Giovanna Brancato - *Metodologie e stime dell'errore di risposta. Una sperimentazione di reintervista telefonica*
- 3/2004 - Franco Mostacci, Giuseppina Natale e Elisabetta Pugliese - *Gli indici dei prezzi al consumo per sub popolazioni*
- 4/2004 - Leonello Tronti - *Una proposta di metodo: osservazioni e raccomandazioni sulla definizione e la classificazione di alcune variabili attinenti al mercato del lavoro*
- 5/2004 - Ugo Guarnera - *Alcuni metodi di imputazione delle mancate risposte parziali per dati quantitativi: il software Quis*
- 6/2004 - Patrizia Giaquinto, Marco Landriscina e Daniela Pagliuca - *La nuova funzione di analisi dei modelli implementata in Genesees v. 3.0*
- 7/2004 - Roberto Di Giuseppe, Patrizia Giaquinto e Daniela Pagliuca - *MAUSS (Multivariate Allocation of Units in Sampling Surveys): un software generalizzato per risolvere il problema dell'allocazione campionaria nelle indagini Istat*
- 8/2004 - Ennio Fortunato e Liana Verzicco - *Problemi di rilevazione e integrazione della condizione professionale nelle indagini sociali dell'Istat*
- 9/2004 - Claudio Pauselli e Claudia Rinaldelli - *La valutazione dell'errore di campionamento delle stime di povertà relativa secondo la tecnica Replicazioni Bilanciate Ripetute*
- 10/2004 - Eugenio Arcidiacono, Marina Briolini, Paolo Giuberti, Marco Ricci, Giovanni Sacchini e Giorgia Telloli - *Procedimenti giudiziari, reati, indagati e vittime in Emilia-Romagna nel 2002: un'analisi territoriale sulla base dei procedimenti iscritti nel sistema informativo Re.Ge.*
- 11/2004 - Enrico Grande e Orietta Luzi - *Regression trees in the context of imputation of item non-response: an experimental application on business data*
- 12/2004 - Luisa Frova e Marilena Pappagallo - *Procedura di now-cast dei dati di mortalità per causa*
- 13/2004 - Giorgio DellaRocca, Marco Di Zio, Orietta Luzi, Emanuela Scavalli e Giorgia Simeoni - *IDEA (Indices for Data Editing Assessment): sistema per la valutazione degli effetti di procedure di controllo e correzione dei dati e per il calcolo degli indicatori SIDI*
- 14/2004 - Monica Pace, Silvia Bruzzone, Luisa Frova e Marilena Pappagallo - *Review of the existing information about death certification practices, certificate structures and training tools for certification of causes of death in Europe*
- 15/2004 - Elisa Berntsen - *Modello Unico di Dichiarazione ambientale: una fonte amministrativa per l'Archivio delle Unità Locali di Asia*
- 16/2004 - Salvatore F. Allegra e Alessandro La Rocca - *Sintetizzare misure elementari: una sperimentazione di alcuni criteri per la definizione di un indice composto*
- 17/2004 - Francesca R. Pogelli - *Un'applicazione del modello "Country Product Dummy" per un'analisi territoriale dei prezzi*
- 18/2004 - Antonia Manzari - *Valutazione comparativa di alcuni metodi di imputazione singola delle mancate risposte parziali per dati quantitativi*
- 19/2004 - Claudio Pauselli - *Intensità di povertà relativa: stima dell'errore di campionamento e sua valutazione temporale*
- 20/2004 - Maria Dimitri, Ersilia Di Pietro, Alessandra Nuccitelli e Evelina Paluzzi - *Sperimentazione di una metodologia per il controllo della qualità di dati anagrafici*
- 21/2004 - Tiziana Pichiorri, Anna M. Sgamba e Valerio Papale - *Un modello di ottimizzazione per l'imputazione delle mancate risposte statistiche nell'indagine sui trasporti marittimi dell'Istat*

(*) ultimi cinque anni

- 22/2004 – Diego Bellisai, Piero D. Falorsi, Annalisa Lucarelli, Maria A. Pennucci e Leonello G. Tronti – *Indagine pilota sulle retribuzioni di fatto nel pubblico impiego*
- 23/2004 – Lidia Brondi – *La riorganizzazione del sistema idrico: quadro normativo, delimitazione degli ambiti territoriali ottimali e analisi statistica delle loro caratteristiche strutturali*
- 24/2004 – Roberto Gismondi e Laura De Sandro – *Provisional Estimation of the Italian Monthly Retail Trade Index*
- 25/2004 – Annamaria Urbano, Claudia Brunini e Alessandra Chessa – *I minori in stato di abbandono: analisi del fenomeno e studio di una nuova prospettiva d'indagine*
- 26/2004 – Paola Anzini e Anna Ciammola – *La destagionalizzazione degli indici della produzione industriale: un confronto tra approccio diretto e indiretto*
- 27/2004 – Alessandro La Rocca – *Analisi della struttura settoriale dell'occupazione regionale: 8° Censimento dell'industria e dei servizi 2001 7° Censimento dell'industria e dei servizi 1991*
- 28/2004 – Vincenzo Spinelli e Massimiliano Tancioni – *I Trattamenti Monetari non Pensionistici: approccio computazionale e risultati della sperimentazione sugli archivi INPS-DM10*
- 29/2004 – Paolo Consolini – *L'indagine sperimentale sull'archivio fiscale modd.770 anno 1999: analisi della qualità del dato e stime campionarie*
- 1/2005 – Fabrizio M. Arosio – *La stampa periodica e l'informazione on-line: risultati dell'indagine pilota sui quotidiani on-line*
- 2/2005 – Marco Di Zio, Ugo Guarnera e Orietta Luzi – *Improving the effectiveness of a probabilistic editing strategy for business data*
- 3/2005 – Diego Moretti e Claudia Rinaldelli – *EU-SILC complex indicators: the implementation of variance estimation*
- 4/2005 – Fabio Bacchini, Roberto Iannaccone e Edoardo Otranto – *L'imputazione delle mancate risposte in presenza di dati longitudinali: un'applicazione ai permessi di costruzione*
- 5/2005 – Marco Broccoli – *Analisi della criminalità a livello comunale: metodologie innovative*
- 6/2005 – Claudia De Vitiis, Loredana Di Consiglio e Stefano Falorsi – *Studio del disegno campionario per la nuova rilevazione continua sulle Forze di Lavoro*
- 7/2005 – Edoardo Otranto e Roberto Iannaccone – *Continuous Time Models to Extract a Signal in Presence of Irregular Surveys*
- 8/2005 – Cosima Mero e Adriano Pareto – *Analisi e sintesi degli indicatori di qualità dell'attività di rilevazione nelle indagini campionarie sulle famiglie*
- 9/2005 – Filippo Oropallo – *Enterprise microsimulation models and data challenges*
- 10/2005 – Marcello D' Orazio, Marco Di Zio e Mauro Scanu – *A comparison among different estimators of regression parameters on statistically matched files through an extensive simulation study*
- 11/2005 – Stefania Macchia, Manuela Murgia, Loredana Mazza, Giorgia Simeoni, Francesca Di Patrizio, Valentino Parisi, Roberto Petrillo e Paola Ungaro – *Una soluzione per la rilevazione e codifica della Professione nelle indagini CATI*
- 12/2005 – Piero D. Falorsi, Monica Scannapieco, Antonia Boggia e Antonio Pavone – *Principi Guida per il Miglioramento della Qualità dei Dati Toponomastici nella Pubblica Amministrazione*
- 13/2005 – Ciro Baldi, Francesca Ceccato, Silvia Pacini e Donatella Tuzi – *La stima anticipata OROS sull'occupazione. Errori, problemi della metodologia attuale e proposte di miglioramento*
- 14/2005 – Stefano De Francisci, Giuseppe Sindoni e Leonardo Tininini – *Da Winci/MD: un sistema per data warehouse statistici sul Web*
- 15/2005 – Gerardo Gallo e Evelina Palazzi – *I cittadini italiani naturalizzati: l'analisi dei dati censuari del 2001, con un confronto tra immigrati di prima e seconda generazione*
- 16/2005 – Saverio Gazzelloni, Mario Albisinni, Lorenzo Bagatta, Claudio Ceccarelli, Luciana Quattrociochi, Rita Ranaldi e Antonio Toma – *La nuova rilevazione sulle forze di lavoro: contenuti, metodologie, organizzazione*
- 17/2005 – Maria Carla Congia – *Il lavoro degli extracomunitari nelle imprese italiane e la regolarizzazione del 2002. Prime evidenze empiriche dai dati INPS*
- 18/2005 – Giovanni Bottazzi, Patrizia Cella, Giuseppe Garofalo, Paolo Misso, Mariano Porcu e Marianna Tosi – *Indagine pilota sulla nuova imprenditorialità nella Regione Sardegna. Relazione Conclusiva*
- 19/2005 – Fabrizio Martire e Donatella Zindato – *Le famiglie straniere: analisi dei dati censuari del 2001 sui cittadini stranieri residenti*
- 20/2005 – Ennio Fortunato – *Il Sistema di Indicatori Territoriali: percorso di progetto, prospettive di sviluppo e integrazione con i processi di produzione statistica*
- 21/2005 – Antonella Baldassarini e Danilo Birardi – *I conti economici trimestrali: un approccio alla stima dell'input di lavoro*
- 22/2005 – Francesco Rizzo, Dario Camol e Laura Vignola – *Uso di XML e WEB Services per l'integrazione di sistemi informativi statistici attraverso lo standard SDMX*
- 1/2006 – Ennio Fortunato – *L'analisi integrata delle esigenze informative dell'utenza Istat: Il contributo del Sistema di Indicatori Territoriali*
- 2/2006 – Francesco Altarocca – *I design pattern nella progettazione di software per il supporto alla statistica ufficiale*
- 3/2006 – Roberta Palmieri – *Le migranti straniere: una lettura di genere dei dati dell'osservatorio interistituzionale sull'immigrazione in provincia di Macerata*
- 4/2006 – Raffaella Amato, Silvia Bruzzone, Valentina Delmonte e Lidia Fagiolo – *Le statistiche sociali dell'ISTAT e il fenomeno degli incidenti stradali: un'esperienza di record linkage*
- 5/2006 – Alessandro La Rocca – *Fuzzy clustering: la logica, i metodi*
- 6/2006 – Raffaella Cascioli – *Integrazione dei dati micro dalla Rilevazione delle Forze di Lavoro e dagli archivi amministrativi INPS: risultati di una sperimentazione sui dati campione di 4 province*
- 7/2006 – Gianluca Brogi, Salvatore Cusimano, Giuseppina del Vicario, Giuseppe Garofalo e Orietta Patacchia – *La realizzazione di Asia Agricoltura tramite l'utilizzo di dati amministrativi: il contenuto delle fonti e i risultati del processo di integrazione*
- 8/2006 – Simonetta Cozzi – *La distribuzione commerciale in Italia: caratteristiche strutturali e tendenze evolutive*
- 9/2006 – Giovanni Seri – *A graphical framework to evaluate risk assessment and information loss at individual level*

- 10/2006 – Diego Bellisai, Annalisa Lucarelli, Maria Anna Pennucci e Fabio Rapiti – *Feasibility studies for the coverage of public institutions in sections N and O*
- 11/2006 – Diego Bellisai, Annalisa Lucarelli, Maria Anna Pennucci e Fabio Rapiti – *Quarterly labour cost index in public education*
- 12/2006 – Silvia Montagna, Patrizia Collesi, Florinda Damiani, Danila Fulgenzio, Maria Francesca Loporcaro e Giorgia Simeoni – *Nuove esperienze di rilevazione della Customer Satisfaction*
- 13/2006 – Lucia Coppola e Giovanni Seri – *Confidentiality aspects of household panel surveys: the case study of Italian sample from EU-SILC*
- 14/2006 – Lidia Brondi – *L'utilizzazione delle surveys per la stima del valore monetario del danno ambientale: il metodo della valutazione contingente*
- 15/2006 – Carlo Boselli – *Le piccole imprese leggere esportatrici e non esportatrici: differenze di struttura e di comportamento*
- 16/2006 – Carlo De Greogorio – *Il nuovo impianto della rilevazione centralizzata del prezzo dei medicinali con obbligo di prescrizione*
- 1/2007 – Paolo Roberti, Maria Grazia Calza, Filippo Oropallo e Stefania Rossetti – *Knowledge Databases to Support Policy Impact Analysis: the EuroKy-PIA Project*

Documenti ISTAT(*)

- 1/2002 – Paolo Consolini e Rita De Carli - *Le prestazioni sociali monetarie non pensionistiche: unità di analisi, fonti e rappresentazione statistica dei dati*
- 2/2002 – Stefania Macchia - *Sperimentazione, implementazione e gestione dell'ambiente di codifica automatica della classificazione delle Attività economiche*
- 3/2002 – Maria De Lucia - *Applicabilità della disciplina in materia di festività nel pubblico impiego*
- 4/2002 – Roberto Gismondi, Massimo Marciani e Mauro Giorgetti - *The italian contribution towards the implementation of an european transport information system: main results of the MESUDEMO project*
- 5/2002 – Olimpio Cianfarani e Sauro Angeletti - *Misure di risultato e indicatori di processo: l'esperienza progettuale dell'Istat*
- 6/2002 – Riccardo Carbinì e Valerio De Santis – *Programma statistico nazionale: specifiche e note metodologiche per la compilazione delle schede identificative dei progetti*
- 7/2002 – Maria De Lucia – *Il CCNL del personale dirigente dell'area 1 e la valutazione delle prestazioni dei dirigenti*
- 8/2002 – Giuseppe Garofalo e Enrica Morganti – *Gruppo di lavoro per la progettazione di un archivio statistico sui gruppi d'impresa*
- 1/2003 – Francesca Ceccato, Massimiliano Tancioni e Donatella Tuzi – *MODSIM-P: Il nuovo modello dinamico di previsione della spesa pensionistica*
- 2/2003 – Anna Pia Mirto – *Definizioni e classificazioni delle strutture ricettive nelle rilevazioni statistiche ufficiali sull'offerta turistica*
- 3/2003 – Simona Spirito – *Le prestazioni assistenziali monetarie non pensionistiche*
- 4/2003 – Maria De Lucia – *Approfondimenti di alcune tematiche inerenti la gestione del personale*
- 5/2003 – Rosalia Coniglio, Marialuisa Cugno, Maria Filmeno e Alberto Vitalini – *Mappatura della criminalità nel distretto di Milano*
- 6/2003 – Maria Letizia D'Autilia – *I provvedimenti di riforma della pubblica amministrazione per l'identificazione delle "Amministrazioni pubbliche" secondo il Sec95: analisi istituzionale e organizzativa per l'anno 2000*
- 7/2003 – Francesca Gallo, Pierpaolo Massoli, Sara Mastrovita, Roberto Merluzzi, Claudio Pauselli, Isabella Siciliani e Alessandra Sorrentino – *La procedura di controllo e correzione dei dati Panel Europeo sulle famiglie*
- 8/2003 – Cinzia Castagnaro, Martina Lo Conte, Stefania Macchia e Manuela Murgia – *Una soluzione in-house per le indagini CATI: il caso della Indagine Campionaria sulle Nascite*
- 9/2003 – Anna Pia Maria Mirto e Norina Salamone – *La classificazione delle strutture ricettive turistiche nella normativa delle regioni italiane*
- 10/2003 – Roberto Gismondi e Anna Pia Maria Mirto – *Le fonti statistiche per l'analisi della congiuntura turistica: il mosaico italiano*
- 11/2003 – Loredana Di Consiglio e Stefano Falorsi – *Alcuni aspetti metodologici relativi al disegno dell'indagine di copertura del Censimento Generale della Popolazione 2001*
- 12/2003 – Roberto Gismondi e Anna Rita Giorgi – *Struttura e dinamica evolutiva del comparto commerciale al dettaglio: le tendenze recenti e gli effetti della riforma "Bersani"*
- 13/2003 – Donatella Cangialosi e Rosario Milazzo – *Fabbisogni formativi degli Uffici comunali di statistica: indagine rapida in Sicilia*
- 14/2003 – Agostino Buratti e Giovanni Salzano – *Il sistema automatizzato integrato per la gestione delle rilevazioni dei documenti di bilancio degli enti locali*
- 1/2004 – Giovanna Brancato e Giorgia Simeoni – *Tesori del Sistema Informativo di Documentazione delle Indagini (SIDI)*
- 2/2004 – Corrado Peperoni – *Indagine sui bilanci consuntivi degli Enti previdenziali: rilevazione, gestione e procedure di controllo dei dati*
- 3/2004 – Marzia Angelucci, Giovanna Brancato, Dario Camol, Alessio Cardacino, Sandra Maresca e Concetta Pellegrini – *Il sistema ASIMET per la gestione delle Note Metodologiche dell'Annuario Statistico Italiano*
- 4/2004 – Francesca Gallo, Sara Mastrovita, Isabella Siciliani e Giovanni Battista Arcieri – *Il processo di produzione dell'Indagine ECHP*
- 5/2004 – Natale Renato Fazio e Carmela Pascucci – *Gli operatori non identificati nelle statistiche del commercio con l'estero: metodologia di identificazione nelle spedizioni "groupage" e miglioramento nella qualità dei dati*
- 6/2004 – Diego Moretti e Claudia Rinaldelli – *Una valutazione dettagliata dell'errore campionario della spesa media mensile familiare*
- 7/2004 – Franco Mostacci – *Aspetti Teorico-pratici per la Costruzione di Indici dei Prezzi al Consumo*
- 8/2004 – Maria Frustaci – *Glossario economico-statistico multilingua*
- 9/2004 – Giovanni Seri e Maurizio Lucarelli – *"Il Laboratorio per l'analisi dei dati elementari (ADELE): monitoraggio dell'attività dal 1999 al 2004"*
- 10/2004 – Alessandra Nuccitelli, Francesco Bosio e Luciano Fioriti – *L'applicazione RECLINK per il record linkage: metodologia implementata e linee guida per la sua utilizzazione*
- 1/2005 – Francesco Cuccia, Simone De Angelis, Antonio Laureti Palma, Stefania Macchia, Simona Mastroluca e Domenico Perrone – *La codifica delle variabili testuali nel 14° Censimento Generale della Popolazione*
- 2/2005 – Marina Peci – *La statistica per i Comuni: sviluppo e prospettive del progetto Sisco.T (Servizio Informativo Statistico Comunale. Tavole)*
- 3/2005 – Massimiliano Renzetti e Annamaria Urbano – *Sistema Informativo sulla Giustizia: strumenti di gestione e manutenzione*
- 4/2005 – Marco Broccoli, Roberto Di Giuseppe e Daniela Pagliuca – *Progettazione di una procedura informatica generalizzata per la sperimentazione del metodo Microstrat di coordinamento della selezione delle imprese soggette a rilevazioni nella realtà Istat*
- 5/2005 – Mauro Albani e Francesca Pagliara – *La ristrutturazione della rilevazione Istat sulla criminalità minorile*
- 6/2005 – Francesco Altarocca e Gaetano Sberno – *Progettazione e sviluppo di un "Catalogo dei File Grezzi con meta-dati di base" (CFG) in tecnologia Web*

- 7/2005 – Salvatore F. Allegra e Barbara Baldazzi – *Data editing and quality of daily diaries in the Italian Time Use Survey*
- 8/2005 – Alessandra Capobianchi – *Alcune esperienze in ambito internazionale per l'accesso ai dati elementari*
- 9/2005 – Francesco Rizzo, Laura Vignola, Dario Camol e Mauro Bianchi – *Il progetto "banca dati della diffusione congiunturale"*
- 10/2005 – Ennio Fortunato e Nadia Mignolli – *I sistemi informativi Istat per la diffusione via web*
- 11/2005 – Ennio Fortunato e Nadia Mignolli – *Sistemi di indicatori per l'attività di governo: l'offerta informativa dell'Istat*
- 12/2005 – Carlo De Gregorio e Stefania Fatello – *L'indice dei prezzi al consumo dei testi scolastici nel 2004*
- 13/2005 – Francesco Rizzo e Laura Vignola – *RSS: uno standard per diffondere informazioni*
- 14/2005 – Ciro Baldi, Diego Bellisai, Stefania Fivizzani, Annalisa Lucarelli e Marina Sorrentino – *Launching and implementing the job vacancy statistics*
- 15/2005 – Stefano De Francisci, Massimiliano Renzetti, Giuseppe Sindoni e Leonardo Tininini – *La modellazione dei processi nel Sistema Informativo Generalizzato di Diffusione dell'ISTAT*
- 16/2005 – Ennio Fortunato e Nadia Mignolli – *Verso il Sistema di Indicatori Territoriali: rilevazione e analisi della produzione Istat*
- 17/2005 – Raffaella Cianchetta e Daniela Pagliuca – *Soluzioni Open Source per il software generalizzato in Istat: il caso di PHPSurveyor*
- 18/2005 – Gianluca Giuliani e Barbara Boschetto – *Gli indicatori di qualità dell'Indagine continua sulle Forze di Lavoro dell'Istat*
- 19/2005 – Rossana Balestrino, Franco Garritano, Carlo Cipriano e Luciano Fanfoni – *Metodi e aspetti tecnologici di raccolta dei dati sulle imprese*
- 1/2006 – Roberta Roncati – www.istat.it (versione 3.0) *Il nuovo piano di navigazione*
- 2/2006 – Maura Seri e Annamaria Urbano – *Sistema Informativo Territoriale sulla Giustizia: la sezione sui confronti internazionali*
- 3/2006 – Giovanna Brancato, Riccardo Carbini e Concetta Pellegrini – *SIQual: il sistema informativo sulla qualità per gli utenti esterni*
- 4/2006 – Concetta Pellegrini – *Soluzioni tecnologiche a supporto dello sviluppo di sistemi informativi sulla qualità: l'esperienza SIDI*
- 5/2006 – Maurizio Lucarelli – *Una valutazione critica dei modelli di accesso remoto nella comunicazione di informazione statistica*
- 6/2006 – Natale Renato Fazio – *La ricostruzione storica delle statistiche del commercio con l'estero per gli anni 1970-1990*
- 7/2006 – Emilia D'Acunto – *L'evoluzione delle statistiche ufficiali sugli indici dei prezzi al consumo*
- 8/2006 – Ugo Guarnera, Orietta Luzi e Stefano Salvi – *Indagine struttura e produzioni delle aziende agricole: la nuova procedura di controllo e correzione automatica per le variabili su superfici aziendali e consistenza degli allevamenti*
- 9/2006 – Maurizio Lucarelli – *La regionalizzazione del Laboratorio ADELE: un'ipotesi di sistema distribuito per l'accesso ai dati elementari*
- 10/2006 – Alessandra Bugio, Claudia De Vitiis, Stefano Falorsi, Lidia Gargiulo, Emilio Gianicolo e Alessandro Pallara – *La stima di indicatori per domini sub-regionali con i dati dell'indagine: condizioni di salute e ricorso ai servizi sanitari*
- 11/2006 – Sonia Vittozzi, Paola Giacchè, Achille Zuchegna, Piero Crivelli, Patrizia Collesi, Valerio Tiberi, Alexia Sasso, Maurizio Bonsignori, Giuseppe Stassi e Giovanni A. Barbieri – *Progetto di articolazione della produzione editoriale in collane e settori*
- 12/2006 – Alessandra Coli, Francesca Tartamella, Giuseppe Sacco, Ivan Faiella, Marcello D'Orazio, Marco Di Zio, Mauro Scanu, Isabella Siciliani, Sara Colombini e Alessandra Masi – *La costruzione di un Archivio di microdati sulle famiglie italiane ottenuto integrando l'indagine ISTAT sui consumi delle famiglie italiane e l'Indagine Banca d'Italia sui bilanci delle famiglie italiane*
- 13/2006 – Ersilia Di Pietro – *Le statistiche del commercio estero dell'Istat: rilevazione Intrastat*
- 14/2006 – Ersilia Di Pietro – *Le statistiche del commercio estero dell'Istat: rilevazione Extrastat*
- 15/2006 – Ersilia Di Pietro – *Le statistiche del commercio estero dell'Istat: comparazione tra rilevazione Intrastat ed Extrastat*
- 16/2006 – Fabio M. Rapiti – *Short term statistics quality Reporting: the LCI National Quality Report 2004*
- 17/2006 – Giampiero Siesto, Franco Branchi, Cristina Casciano, Tiziana Di Francescantonio, Piero Demetrio Falorsi, Salvatore Filiberti, Gianfranco Marsigliesi, Umberto Sansone, Ennio Santi, Roberto Sanzo e Alessandro Zeli – *Valutazione delle possibilità di uso di dati fiscali a supporto della rilevazione PMI*
- 18/2006 – Mauro Albani – *La nuova procedura per il trattamento dei dati dell'indagine Istat sulla criminalità*
- 19/2006 – Alessandra Capobianchi – *Review dei sistemi di accesso remoto: schematizzazione e analisi comparativa*
- 20/2006 – Francesco Altarocca – *Gli strumenti informatici nella raccolta dei dati di indagini statistiche: il caso della Rilevazione sperimentale delle tecnologie informatiche e della comunicazione nelle Pubbliche Amministrazioni locali*
- 1/2007 – Giuseppe Stassi – *La politica editoriale dell'Istat nel periodo 1996-2004: collane, settori, modalità di diffusione*
- 2/2007 – Daniela Ichim – *Microdata anonymisation of the Community Innovation Survey data: a density based clustering approach for risk assessment*
- 3/2007 – Ugo Guarnera, Orietta Luzi e Irene Tommasi – *La nuova procedura di controllo e correzione degli errori e delle mancate risposte parziali nell'indagine sui Risultati Economici delle Aziende Agricole (REA)*