

Consiglio Nazionale delle Ricerche



**METODO PER LA VALUTAZIONE E LA COMPARAZIONE
QUALI-QUANTITATIVA DELLA PRODUTTIVITÀ
SCIENTIFICA DEGLI ISTITUTI DI RICERCA EUROPEI**

Sara Berselli e Luca Tiberi

Direzione Centrale Supporto alla Programmazione e alle Infrastrutture

Maggio 2010

NOTA DEGLI AUTORI: il presente rapporto tecnico ha il solo scopo di illustrare una metodologia per la valutazione quali-quantitativa della produttività scientifica degli Istituti di ricerca. Gli autori non intendono esprimere un giudizio sulla qualità degli Istituti a cui tale metodo è stato applicato. Si precisa inoltre che lo studio è frutto della collaborazione dei due autori in ogni sua parte. Tuttavia la stesura materiale del testo è da suddividersi come segue: capitoli 1, 4-8 di Sara Berselli; capitoli 2-3 di Luca Tiberi

INDICE

Abstract	4
1. Introduzione.....	4
2. Banche dati attualmente disponibili per il reperimento di dati bibliometrici	7
2.1 ISI Web of knowledge: caratteristiche principali e procedura di interrogazione	7
2.1.1 Interrogazione: comandi fondamentali.....	8
2.1.2 Interrogazione: esempi pratici.....	10
2.1.3 Interrogazione: funzionalità aggiuntive.....	13
2.1.4 Visualizzazione ed estrazione dei dati.....	14
2.1.5 Analisi dei dati.....	15
2.2 Altri strumenti.....	16
3 Principali indici bibliometrici	16
3.1 Numero di citazioni.....	16
3.2 H-index	17
3.3 G-index.....	18
3.4 IF	19
4. Applicazione del metodo per la comparazione della produttività scientifica di istituti operanti nel campo dell'inquinamento atmosferico.....	20
5. Confronto fra i 37 Istituti considerati e relativa rappresentazione grafica	22
6. Esercizio per la comparazione della produttività scientifica di Paesi diversi	26
7. Considerazioni conclusive.....	29
8. Bibliografia.....	30

Abstract

Questo studio ha lo scopo di presentare un metodo per la valutazione e la comparazione degli istituti di Ricerca Europei, individuandone i parametri utili per la stima sia quantitativa che qualitativa. Lo studio, basato sui dati disponibili nella banca dati ISI Web of Knowledge, è stato applicato a 37 Istituti di ricerca europei che conducono attività di ricerca in campo ambientale. La loro attività scientifica è stata confrontata nel quinquennio 2003-2007, attraverso parametri attestanti la produttività individuale del personale operante nell'Istituto in esame e l'impatto del singolo paper in termini di citazioni. Lo studio è stato completato attraverso un primo tentativo di elaborazione di parametri bibliometrici che, a partire dai dati di ciascun Istituto, consentissero di fare una stima della produttività scientifica del Paese di appartenenza in uno specifico settore scientifico. Il report illustra in modo dettagliato il metodo di interrogazione della banca dati ISI Web of Knowledge, utilizzata dagli autori, ed espone brevemente i principali indici bibliometrici di norma utilizzati in questo ambito.

1. Introduzione

Negli ultimi anni, la valutazione della ricerca scientifica ha attratto notevole e crescente interesse e, in questo ambito, particolare attenzione è stata rivolta alla valutazione della ricerca condotta negli enti pubblici. Ciò è stato determinato in gran parte dalle sempre minori risorse economiche che i Governi Europei hanno destinato alla ricerca scientifica e alla leadership di Stati Uniti e Giappone, che rappresentano, in questo settore, i principali concorrenti per l'Europa. Molti autori, sia Europei che non, hanno proposto metodi e indici per la valutazione della produttività scientifica, sia in termini quantitativi che qualitativi.

Questo studio ha lo scopo di proporre un metodo per la valutazione della produttività scientifica di un singolo Istituto di ricerca e per la comparazione della sue performance con quelle di altri Istituti dislocati in differenti paesi. Tale metodo si basa sull'utilizzo di indici bibliometrici elaborati sulla base dei dati forniti dalla banca dati ISI Web of Knowledge.

Il metodo di valutazione elaborato è stato applicato a 37 Istituti di ricerca Europei, presenti in 12 diversi Paesi, che conducono attività di ricerca in campo ambientale. La scelta del settore ambientale, per la messa a punto del metodo, è stata dettata dalla crescente attenzione che i Governi Europei hanno dedicato negli ultimi anni allo sviluppo sostenibile e alla salvaguardia dell'ambiente, dedicandovi risorse economiche e promuovendone la ricerca scientifica. Oggi l'attività di ricerca è soprattutto rivolta alla gestione sostenibile del territorio e delle sue risorse, attraverso lo studio dell'impatto che le attività umane hanno sull'ecosistema. La ricerca è allo stesso tempo indirizzata alla messa a punto di nuove tecnologie e nuovi metodi per il monitoraggio, la prevenzione e la mitigazione dei rischi ambientali.

I 37 Istituti sono stati selezionati con il supporto del Panel E.2 "Scienze dell'Ambiente", che ha operato nell'ambito del processo di valutazione degli Istituti del Consiglio Nazionale delle Ricerche. Questi, come si può chiaramente desumere, non rappresentano la totalità degli Istituti che si occupano di ricerca in campo ambientale in Europa, ma rappresentano alcuni fra quelli più attivi e noti a livello internazionale. Essi svolgono attività di ricerca nell'ambito della mitigazione e del monitoraggio dell'inquinamento atmosferico, nella bonifica di siti contaminati, nella decontaminazione di acque e nella ricerca in ambito geologico e marino.

Per realizzare il presente studio, per ciascun Istituto sono state prese in considerazione le pubblicazioni presenti nella banca dati ISI Web of Knowledge nel quinquennio 2003-2007 e il numero totale di citazioni. E' stato inoltre preso in considerazione l'H index relativo al singolo Istituto, calcolato al 2009. Il metodo, che tiene conto del personale impegnato nella ricerca, permette di paragonare le performance dei diversi Istituti, anche se il loro numero di ricercatori è differente.

Il metodo proposto è stato utilizzato inizialmente per il confronto di Istituti Europei appartenenti a Paesi diversi che conducono attività di ricerca all'interno di una stessa area tematica (in questo caso l'ambiente) e più in dettaglio su uno stesso argomento. Tale confronto è utile qualora sia necessario attestare la produttività scientifica di Istituti appartenenti a Paesi diversi su un determinato

filone di ricerca. In questo caso sono stati presi in considerazione 7 Istituti di 3 diversi Paesi Europei, che conducono ricerca nel campo dell'inquinamento atmosferico. Le loro performance sono state confrontate prendendo in esame il numero di pubblicazioni realizzate dal 2003 al 2007 (P '03-'07); il numero di citazioni al 2007 (CIT '07); il numero medio di citazioni che ogni paper ha ricevuto fino al 2007 (citazioni per paper - CPP); il numero di paper pubblicati nella banca dati ISI da un singolo ricercatore dell'Istituto nel quinquennio 2003-2007 (paper per ricercatore - PPR); Hirsch Index al 2009 (HI '09). Come sottolineato anche da altri autori, PPR rappresenta il parametro che consente di confrontare, in termini di produttività scientifica, Istituti di ricerca con diverso numero di ricercatori, dal momento che il dato di produzione viene riferito alla singola unità di personale. Per ciascun Istituto i dati relativi al personale sono stati desunti mediante l'analisi dei report annuali presenti nei siti web di ciascuno di essi.

I parametri sopra indicati, desunti, per i 37 Istituti esaminati, dalla banca dati, sono stati raccolti in una tabella riepilogativa e sono stati rappresentati in un grafico a bolle. Tale rappresentazione grafica consente di confrontare le performance degli istituti sulla base di parametri bibliometrici prestabiliti. Un parametro che è stato considerato, perché ritenuto interessante da parte degli autori, è il numero di ricercatori all'interno dei singoli Istituti. Mettendo in correlazione il numero di ricercatori dei singoli Istituti ai parametri bibliometrici PPR e CPP, è possibile verificare se esiste o meno una correlazione fra la massa di un Istituto e le sue performance in termini di numero di pubblicazioni e di citazioni.

Lo studio si conclude con un esercizio volto alla verifica dell'applicabilità di questo metodo alla determinazione della produttività scientifica dei diversi Paesi Europei, sia in termini qualitativi che quantitativi. In particolare, a partire dai dati dei singoli Istituti di un singolo Paese, ne viene proposta una possibile elaborazione per poter individuare degli indici bibliometrici rappresentativi del Paese stesso. In questa occasione, quindi, non sono stati considerati solo gli Istituti che si occupano di una sola tematica di ricerca, ma tutti quelli che nel Paese

operano nel campo ambientale. In particolare, sono stati esaminati 6 istituti francesi e 8 istituti tedeschi e i dati rappresentativi delle loro produttività scientifiche sono stati rielaborati così da poter essere rappresentativi della produttività scientifica di Francia e Germania in quello specifico ambito. Tramite questa operazione è stato possibile fare un primo tentativo per capire quanto la ricerca scientifica abbia prodotto nei cinque anni considerati nei due Paesi, in termini di numero di pubblicazioni e, qualitativamente, in termini di numero di citazioni. Resta inteso che questo confronto andrebbe completato avendo a disposizione tutti i dati di tutti gli Istituti che in quel Paese conducono attività di ricerca in campo ambientale. Pertanto in questa sede, data la mancanza di alcuni dati, si vuole solo proporre il metodo e i parametri bibliometrici ad esso legati.

Nei paragrafi successivi, oltre all'illustrazione del metodo proposto, verrà illustrata in modo dettagliato la procedura di interrogazione della banca dati ISI Web of Knowledge e verrà dedicato spazio ai parametri bibliometrici più utilizzati a livello internazionale.

2. Banche dati attualmente disponibili per il reperimento di dati bibliometrici

2.1 ISI Web of knowledge: caratteristiche principali e procedura di interrogazione

L'ISI Web of Knowledge (d'ora in avanti WoK) è una piattaforma multidisciplinare composta da più basi dati al suo interno, acquistabili separatamente, il cui nocciolo è costituito dal Web of Science (d'ora in avanti WoS), che è propriamente un repertorio citazionale, cioè uno strumento capace di collegare ad un articolo scientifico recensito nel database tutti i contributi che posteriormente l'hanno citato. Si capisce quindi la rilevanza e la specificità di

questo strumento rispetto al tema qui indagato di fronte alla più ampia categoria dei tradizionali repertori bibliografici, comunemente presenti nelle biblioteche di ricerca; è naturale cioè che un articolo che riceve moltissime citazioni produca una più profonda impressione nel mondo della ricerca in un determinato settore rispetto ad un altro che ne riceva di meno. E' infatti solo sulla base di questa semplice constatazione che si può parlare di impatto sulla ricerca e non certo sulla base di analisi contenutistiche che riguardino il merito scientifico di quanto viene pubblicato.

Il WoS si compone di cinque database tematici al suo interno che sono il Science citation index expanded (SCI-EXPANDED: circa 6650 testate indicizzate), il Social sciences citation index (SSCI: circa 2470 testate indicizzate), l'Arts & humanities citation index (A&HCI: circa 1390 testate indicizzate), il Conference proceedings citation index - Science (CPCI-S) e il Conference proceedings citation index - Social sciences & humanities (CPCI-SSH), i quali ultimi recensiscono i contributi di oltre 10900 congressi; a ognuna delle testate si attribuisce annualmente un Impact Factor (d'ora in avanti IF), che fotografa, come si è detto, l'impatto della pubblicazione sulla ricerca in un determinato settore.

2.1.1 Interrogazione: comandi fondamentali

L'interrogazione del WoS prevede un comando SEARCH, un CITED REFERENCE SEARCH e un ADVANCED SEARCH; è interessante notare come il CITED REFERENCE SEARCH fornisca il numero totale delle citazioni di un articolo a prescindere dal tipo di contratto che si è sottoscritto, laddove i due comandi di ricerca (semplice ed avanzata) restituiscono soltanto le citazioni che rientrano nell'arco cronologico sotteso all'abbonamento. I campi su cui possono insistere le queries del comando SEARCH sono quelli consueti: Topic (cui sottostanno i campi Title, Abstract, Author Keywords e Keywords Plus), Author, Publication Name, ecc., combinati tra loro con gli opportuni operatori booleani; il filtro più comunemente usato in questo tipo di ricerche è l'anno di pubblicazione, ma è possibile anche scegliere su quale dei database sottoscritti, indicati qui sopra, condurre la ricerca.

Il CITED REFERENCE SEARCH è volto al ritrovamento di un lavoro tramite la sua citazione e prevede l'inserimento del nome dell'autore, del titolo abbreviato della testata su cui è apparso il lavoro e l'anno di pubblicazione; è, a questo scopo, disponibile una comoda lista d'autorità delle abbreviazioni delle testate, da cui ricavare la forma esatta da inserire e, anche in questo caso, è possibile selezionare il database o i databases su cui condurre la query.

Una funzionalità molto utile per il singolo ricercatore è l'AUTHOR FINDER, disponibile nel SEARCH del WoS, attraverso il quale è possibile avere una lista di papers legati ad un determinato autore: si tratta di una serie di passaggi successivi, eventualmente suscettibili di essere avviati a conclusione ad ogni fase (FINISH NOW), durante i quali viene richiesto in primo luogo cognome e lettera iniziale del nome dell'autore; a questo punto vengono mostrati tutti i record autore corrispondenti ai criteri di ricerca (INCREASE RESULTS, FOCUS RESULTS) e l'utente è chiamato a spuntare tutte le occorrenze che hanno a che fare con la sua ricerca e per le quali è fornito il numero di records collegati; il tutto è corredato, inoltre, dalla possibilità di aggiungere un secondo nome, da unire al primo nella query. Si passa poi alla selezione dell'ambito disciplinare e infine all'ente di appartenenza, spuntato il quale si arriva finalmente alla lista di lavori collegati. L'impressione generale che si ricava dall'uso e dal funzionamento di questo comando è che esso si dimostra effettivamente valido soltanto nel caso in cui sia un autore a cercare i propri lavori: infatti da una parte l'abbreviazione del nome di battesimo, dall'altra il problema dell'affiliazione fa sì che si determini una certa confusione sia tra eventuali omonimi, circostanza da tenere sempre nella debita considerazione, sia tra gli enti di appartenenza, sui quali, appunto, può essere informato con precisione solo l'autore stesso; in alternativa tale procedura può essere utile soltanto nella misura in cui si disponga di un dettagliato *curriculum vitae*, dell'autore, di cui ci si vuole occupare.

Con l'ADVANCED SEARCH si entra nel vivo della problematica qui esaminata, poiché si ha la possibilità di indirizzare le queries sui campi che identificano gli enti di appartenenza dei singoli studiosi, in modo tale da valutarne l'attività: questi

campi, chiaramente elencati sulla pagina, sono essenzialmente l'AD (Address), inteso come l'indirizzo della struttura di afferenza di un autore, e l'OG (Organization), inteso come il nome dell'ente nell'ambito del quale l'autore opera; anche in questo caso il filtro più comunemente usato è l'anno di pubblicazione, dal momento che non è tanto interessante sapere quanti contributi di un determinato ente di ricerca vengano indicizzati in generale da questo strumento, quanto piuttosto quale sia la produttività in un certo anno o gruppi di anni; sono inoltre disponibili dei limiti per lingua e per tipo di pubblicazione e naturalmente, come già visto per le altre modalità di ricerca, è possibile scegliere su quale database indirizzare la query.

Se dunque l'intento da perseguire è quello di accertare la produzione scientifica di un determinato ente di ricerca, vanno chiariti alcuni punti: in primo luogo non esiste all'interno del database un authority file delle strutture di ricerca, ma i nomi e gli indirizzi vengono riportati così come vengono inseriti dagli autori nelle pubblicazioni; il che significa che quanto più si vogliono ottenere dei risultati vicini alla verità, tanto più bisognerà strutturare delle queries che tengano conto di tutte le forme varianti dei nomi degli enti di ricerca; in secondo luogo va prestata molta attenzione ai casi di omonimia soprattutto quando si lavora con gli acronimi dei nomi degli enti e, in questo specifico caso, per ottenere la disambiguazione di sigle identiche, si può fare ricorso al campo CI (City) che identifica la città, sede dell'istituto di ricerca o al campo CU (Country) che ne identifica la nazionalità; ma di questo si parlerà dettagliatamente più avanti.

2.1.2 Interrogazione: esempi pratici

Ad esemplificazione di quanto sopra esposto si può, quindi, riportare un modello efficace di query, effettuabile con l'ADVANCED SEARCH:

(AD=[nome dell'ente : acronimo] or [nome dell'ente : lingua del paese d'appartenenza : opportunamente troncato] or [nome dell'ente : Inglese : opportunamente troncato]) or (OG=[nome dell'ente : acronimo] or [nome dell'ente

: lingua del paese d'appartenenza : opportunamente troncato] or [nome dell'ente :
Inglese : opportunamente troncato]).

Questa query, associata ad un anno di pubblicazione restituisce il numero di lavoro fatti in un determinato anno da un determinato istituto di ricerca, identificato tramite il suo acronimo, il suo nome nella lingua del paese di appartenenza e il suo nome in lingua inglese; l'uso dell'OR sui campi AD e OG, benché ridondante, risponde ad un'esigenza di maggior accuratezza dei risultati, perché è teoricamente possibile che esista all'interno del database il nome di un'organizzazione senza il suo indirizzo.

Vediamone ora un'applicazione pratica, volta ad accertare la produzione scientifica del CNR nell'anno 2004 (databases: SCI-EXPANDED, SSCI e A&HCI); secondo quanto detto la query sarà strutturata in questo modo:

1 (ad=cnr or ad=cons* naz* ric* or ad=nat* res* coun* ital*) or (og=cnr or og=cons* naz* ric* or og=nat* res* coun* ital*)

e restituisce 5.048 risultati, laddove la stessa query condotta solo sul campo OG

2 (og=cnr or og=cons* naz* ric* or og=nat* res* coun* ital*)

ne restituisce solo 4.056 e sul campo AD

3 (ad=cnr or ad=cons* naz* ric* or ad=nat* res* coun* ital*)

di nuovo 5.048. Alla luce dei risultati va da sé che AD comprende OG. Se si prende in considerazione il fatto che spesso si potrebbe avere a che fare con casi di omonimia nella denominazione delle strutture di ricerca, non sarà inappropriato tentare di ottimizzare la query vista sopra con l'uso del campo CU (country), anziché sussumere l'indicazione di paese nella sua denominazione, come appena fatto; la query così rivista avrà dunque questa fisionomia:

4 (ad=cnr or ad=cons* naz* ric* or (ad=nat* res* coun* and cu=ital*))

che restituisce 5.140 records, cioè circa 100 in più di 3, che era la query ottima fino ad ora; la differenza è dovuta a tutti quei record a firma National research council senza Ital* (Italy, italian, ecc.) nell'Address, ma che sono indicizzati con Ital* nel

campo CU; questo, tuttavia non mette al riparo dal recuperare anche tutti quei records che, ad esempio, hanno National research council of Canada nell'AD e, su un'altra riga, nel caso di collaborazione tra più autori, una struttura italiana, che ha, quindi, Italy, nel campo CU. Non servirebbe nemmeno inserire National research council of Canada nella query generale con NOT, perché scarterebbe tutti i records frutto di collaborazione tra i due istituti di ricerca. Si capisce, dunque, che non è il caso di insistere su questa strada, tentando anche

5 ((ad=cnr and cu=ital*) or ad=cons* naz* ric* or (ad=nat* res* counc* and cu=ital*))

che produce 4.949 records, cioè circa 200 in meno della precedente, perché non risolve il problema, cui si è accennato sopra. Al contrario, se si intraprende la via opposta, eliminando l'indicazione di paese dalla query

6 (ad=cnr or ad=cons* naz* ric* or ad=nat* res* counc*)

otteniamo 6.214 records, dai quali andrebbero poi tolti tutti quelli, ad esempio, di paternità National research council of Canada, che, attraverso l'uso del REFINE, funzionalità che si esaminerà più sotto, a proposito della visualizzazione dei risultati, scopriamo essere ben 1.038. In linea teorica, quindi, si può affermare, sulla base degli esempi visti, che è preferibile effettuare delle queries generiche, riservandosi poi di raffinare i risultati, piuttosto che fare delle ricerche molto accurate, rischiando di tagliar fuori dati utili per il proprio lavoro. Va comunque in via definitiva ribadito che l'esperienza, intesa in senso strettamente etimologico, cioè l'imparare dal fare e soprattutto dal commettere errori, costituisce anche in questo ordine di cose il miglior modo per costruire gli algoritmi di ricerca: se infatti trasformiamo l'espressione 3 in

7 (ad=cnr or ad=cons* naz* ric* or ad=nat* res* counc* ital* or ad=ital* nat* res* counc*) or (og=cnr or og=cons* naz* ric* or og=nat* res* counc* ital* or og=ital* nat* res* counc*)

aggiungendo l'ennesima forma variante con ital* all'inizio del blocco e non alla fine, otteniamo un risultato ancora maggiore (5.100) con circa 50 records in più sul

confronto con la 3; vale la pena di ricordare che il database per ogni record tende a confondere la data di inserimento nel database con la data di pubblicazione, circostanza che rende obbligatorio un filtraggio dei dati ottenuti ancora una volta per data di pubblicazione, attraverso il comando REFINE, di cui si parlerà più sotto, che infatti ci indica che dei 5100 records solo 4555 appartengono all'anno 2004.

Da quanto detto, infine, è evidente che andrebbe compiuto uno sforzo di normalizzazione nelle forme identificative delle strutture di ricerca, da compiersi, da una parte a cura dei singoli ricercatori, dall'altra ad opera dell'ISI e degli editori commerciali; va detto che una iniziativa in questo senso da parte dell'ISI è rappresentata dal MY RESEARCHER ID, un codice numerico che identifica l'autore in modo univoco all'interno del database, riducendo al minimo gli errori dovuti a omonimie, strutture di appartenenza, ecc.

2.1.3 Interrogazione: funzionalità aggiuntive

Per le tre modalità di ricerca qui sopra descritte, è disponibile uno storico nella sezione SEARCH HISTORY, dove si trovano ricapitolate tutte le queries condotte nella sessione corrente, con possibilità di combinarle con l'opportuno operatore booleano oppure di cancellarle; è anche possibile salvare le queries fatte, in modo tale da non doverle ridigitare successivamente, sia in locale, sia, previa registrazione, sul server ISI, nonché creare degli alerts; entrambe le funzionalità sono particolarmente utili nel caso di interrogazioni periodiche aventi sempre lo stesso tema, quali possono essere, ad esempio, proprio quelle di cui qui si parla, cioè quelle volte a conoscere la produttività di un istituto di ricerca o di un singolo ricercatore. Tra i due sistemi, tuttavia, esistono delle differenze che sono riassumibili essenzialmente nel fatto che, salvando la query, essa poi va materialmente ripetuta sulla base dati, mentre l>alert consente all'utente l'autoindirizzamento di mail, non appena nella base dati venga inserito un record che soddisfa i criteri di ricerca; nel primo caso dunque si ha sincronicamente tutta la messe di risultati, nel secondo invece soltanto degli aggiornamenti.

2.1.4 Visualizzazione ed estrazione dei dati

Passiamo ora all'esame della visualizzazione dei dati, raccolti con le queries sopra descritte. Ogni set di risultati è personalizzabile sia in termini di record per pagina (10, 25, 50), sia in termini di ordinamento (LATEST DATE, TIMES CITED, ecc.) e presenta i dati in una forma sintetica costituita da titolo, autore/i, referenza bibliografica, numero di citazioni ricevute ed eventualmente il link al full text; la visualizzazione completa, invece, presenta in modo dettagliato tutti i campi del database, compreso l'ADDRESS, di cui abbiamo parlato in precedenza, ed sono qui disponibili una serie di funzionalità aggiuntive, che consistono nel poter creare, previa registrazione, degli alerts, che avvertono l'utente ogni volta che un dato lavoro viene citato, controllare la lista dei riferimenti bibliografici citati dal lavoro di partenza e la lista comprendente i lavori che posteriormente l'hanno citato, controllare l'IF della rivista con link diretto al Journal of citation reports per una serie di anni che dipendono, naturalmente, dal tipo di sottoscrizione effettuata, e avere una lista dei lavori che condividono almeno una citazione con quello di partenza (RELATED RECORDS); a questo proposito è interessante notare come ai fini del grado di rilevanza dei dati così ottenuti, per ogni record, si abbia da una parte il numero totale di citazioni bibliografiche effettuate, dall'altra il numero relativo a quelle condivise con il lavoro di partenza, che ne rende il livello di affinità.

Per quanto riguarda ora l'export dei dati, è prevista, attraverso una maschera posta in fondo alla pagina, una funzione PRINT, che stampa i dati richiesti, una funzione MAIL, che li invia ad un destinatario di posta elettronica, una funzione MARK, che li memorizza temporaneamente per usi futuri dell'utente e due tipi di SAVE, il primo dei quali consente il salvataggio dei dati in un software per la gestione delle referenze bibliografiche (tipicamente viene usato Endnote Web, ma ne sono previsti anche altri), mentre il secondo permette il salvataggio dei dati in un comune formato di lavoro (HTML, ecc.). Vale la pena di accennare al software Endnote Web in modo un po' più dettagliato, perché, agendo da interfaccia tra il

WoS e, ad esempio, Microsoft Word, dal quale esso è accessibile tramite l'installazione di un plug-in, consente l'export immediato delle referenze bibliografiche considerate utili dall'utente nella redazione di un lavoro; tra le funzionalità principali offerte da questo software, si accennerà al fatto che è possibile gestire il proprio patrimonio citazionale, creando cartelle e sottocartelle (anche in condivisione con altri utenti), importare citazioni bibliografiche da database online (Pubmed, ecc.), manualmente o da file, cambiare stile citazionale ed esportare le liste in vari formati.

2.1.5 Analisi dei dati

Oltre alle funzionalità descritte qui sopra, sul set di risultati ottenuto è possibile condurre una serie di operazioni volte al restringimento del numero dei record (REFINE), all'analisi dei risultati secondo alcuni parametri previsti (ANALYZE RESULTS) e alla metrica delle citazioni (CREATE CITATION REPORT), comando con il quale successivamente si possono escludere anche le autocitazioni, ovvero tutti quei records che citano quello/i di partenza con un autore in comune (VIEW WITHOUT SELF-CITATIONS). Il REFINE occupa la parte sinistra della schermata ed è organizzato in campi, dei quali il primo permette l'inserimento di parole ed insiste essenzialmente sul campo titolo, mentre gli altri propongono delle check boxes con i valori ricavati dalla query corrente, che si vuole appunto raffinare; l'ordinamento dei dati di default è per numero di occorrenze, ma è anche disponibile l'ordine alfabetico; molto utili, nell'ottica del controllo periodico della produttività scientifica di un ente di ricerca, sono il REFINE per SOURCE TITLES e per SUBJECT AREAS: il primo consente di suddividere i risultati ottenuti per testata, in modo tale da vedere quali siano i periodici sui quali maggiormente si orienta l'attenzione dei ricercatori o come essa possa subire dei cambiamenti anno dopo anno, il secondo consente di capire quali siano gli ambiti disciplinari presi in esame e quindi di valutare il grado di attenzione a certe tematiche di un determinato ente. L'ANALYZE RESULTS equivale a un REFINE avanzato e consente l'ordinamento dei risultati per autore, paese di pubblicazione, ente

sovvenzionatore, numero di sovvenzione, tipologia di documento (articolo, review, ecc.), ente produttore, lingua, anno di pubblicazione, titolo del periodico e materia; per ogni valore viene mostrato il numero di occorrenze, la percentuale sul totale e una check box, attraverso la quale è possibile l'eliminazione o la visualizzazione dei records collegati.

2.2 Altri strumenti

Aggiungeremo qui per completezza che sul mercato, e più in generale sul web, esistono altri strumenti in grado di svolgere la stessa funzione, assolta dal WoS, cioè repertori citazionali che possono essere utili in sede di accertamento della produzione scientifica di un ente di ricerca. Il più importante di essi è senz'altro Scopus (www.scopus.com), che con i suoi 26.000 seriali indicizzati prevede meccanismi di ricerca analoghi a quelli propri del WoS, con l'esclusione del computo dell'IF, che è un marchio dell'ISI; citeremo anche come strumenti disponibili gratuitamente sul web Google Scholar (scholar.google.com) e la sua interfaccia Publish or Perish, che è più propriamente un software, CiteseerX (citeseerx.ist.psu.edu/), che tuttavia non è uno strumento multidisciplinare come quelli qui ricordati, ma ristretto all'ambito informatico, Citebase (www.citebase.org) e la Digital library dell'ACM (portal.acm.org/dl.cfm).

3 Principali indici bibliometrici

3.1 Numero di citazioni

Il numero delle citazioni ricevute da un articolo rappresenta il dato bibliometrico più immediato e tangibile, quando si affrontano temi riguardanti la valutazione della produzione scientifica di un ente di ricerca; il principio è quello secondo il quale un articolo molto citato ha avuto un impatto sul mondo della

ricerca in un determinato settore maggiore di un altro che ne ha avute meno o nessuna, anche se in termini di comparabilità assoluta va considerato anche l'anno di pubblicazione, dal momento che soltanto due articoli pubblicati nello stesso anno sono perfettamente equiparabili in termini citazionali. Il dato, che nell'ISI WoS è etichettato TC (Times Cited), si trova in calce alla notizia immediatamente dopo i dati bibliografici del lavoro e viene calcolato su tutte le edizioni del WoS, a prescindere dal tipo di contratto sottoscritto, che agirà in questo caso solo da filtro sulla visualizzazione dei records; esiste anche la possibilità di avere un'interfaccia grafica delle relazioni citazionali tra l'articolo di partenza, la sua bibliografia e le citazioni ottenute attraverso il comando CITATION MAP, che ha il compito di creare una vera e propria mappa, nella quale il lavoro di partenza si trova al centro, la sua bibliografia sulla sinistra e le citazioni ottenute sulla destra, consentendo così all'utente di avere un colpo d'occhio generale sull'impatto del lavoro. Analogamente alla visualizzazione dei risultati delle normali ricerche, di cui abbiamo parlato prima, vi sono sul frame di sinistra tutti i comandi relativi al REFINE, l'ANALYZE RESULTS e la parte dedicata all'output dei risultati con possibilità di salvataggio dei dati, stampa, export, ecc. Il problema delle autocitazioni, che hanno il potere di far lievitare artificialmente il TC totale di un articolo, può essere risolto, utilizzando il citato comando CITED REFERENCE SEARCH, che restituisce il lavoro ricercato corredato del numero di citazioni corrispondenti, che andranno successivamente depurate nella sezione REFINE RESULTS in base al nome dell'autore.

3.2 H-index

Creato nel 2005 da J.E. Hirsch, si applica a singoli autori o a interi istituti di ricerca e serve a ricavare sul totale della produzione scientifica di questi ultimi i primi n articoli con almeno n citazioni ricevute, parametro ideale per valutare il livello di impatto di tale produzione sul settore cui appartiene. Si ottiene con il comando CREATE CITATION REPORT sulla query per autore o per ente e dipende, naturalmente, sia dalla data di decorrenza della sottoscrizione, sia dagli anni

selezionati per la query, che infatti vanno sempre indicati, per consentire la piena intellegibilità del dato. Si rappresenta, come esemplificato qui sotto, in un diagramma cartesiano che abbia sull'ascissa x gli articoli in ordine di citazioni decrescente e sull'ordinata y il numero di citazioni corrispondenti; disegnata la curva con questi dati, il punto di intersezione tra quest'ultima e la funzione $x = y$ darà l'H-index, che, elevato al quadrato fornirà il numero di citazioni, su cui esiste; per avere il numero di citazioni totali, infine, si rimanda alla formula $N_{c.tot.} = ah^2$, laddove a è un coefficiente, sperimentalmente determinato, che vale tra 3 e 5.

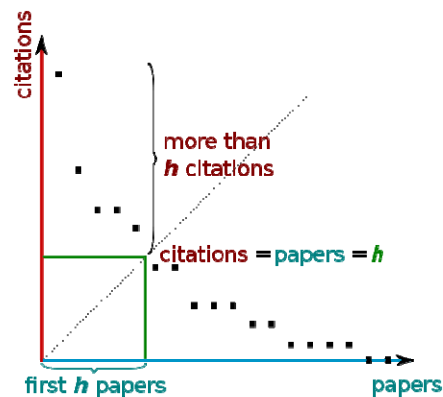


Figura 3.2.1: rappresentazione grafica dell'H index (fonte: Wikipedia)

3.3 G-index

Come si è visto, l'h-index, grazie alla sua semplicità sia in sede di calcolo, sia in sede concettuale, è un ottimo indicatore della produzione scientifica di un autore o di un istituto di ricerca, ma non tiene conto di una cosa per certi versi importantissima e cioè l'aumentare delle citazioni di tutti i papers con $TC > h$; in altre parole, se un articolo con il passare degli anni viene apprezzato sempre di più e vede aumentare cospicuamente le sue citazioni, ma esso si trova già nella parte alta della curva citazionale, cioè è già stato impiegato nel computo dell'h-index, tutte queste citazioni, che pure attestano il merito scientifico del paper, vanno, per così dire perdute. In questo senso nel 2006 L. Egghe ha ideato un g-index con questa definizione: una serie di papers ha un g-index g tale che g sia il numero più

alto di articoli, la somma delle cui citazioni sia uguale a g^2 . E' evidente come, con questa modalità di calcolo, le citazioni che ogni articolo con $TC > h$ riceve, siano impiegate per far scorrere verso il basso il nuovo g-index, includendo così nel computo generale un numero sempre maggiore di articoli.

3.4 IF

Altro dato bibliometrico importante, pertinente alle testate periodiche selezionate a questo scopo dall'ISI, i cui spogli costituiscono WoK/WoS, viene pubblicato nel Journal of citation reports (d'ora in avanti JCR), database distinto e separato dai primi due, ma raggiungibile attraverso apposito link, posto generalmente sul frame di destra all'interno della schermata di visualizzazione del record Wok/WoS, così da avere un riscontro immediato tra il singolo lavoro e il dato di impatto della testata. L'IF si ottiene annualmente sommando il numero di citazioni ottenute nell'anno di riferimento (ad esempio: 2004) dagli articoli pubblicati nei due anni precedenti (ad esempio: 2002-2003) da un determinato periodico e dividendolo per il numero di articoli pubblicati dallo stesso periodico nello stesso biennio; il numero così ottenuto costituisce un fattore di ranking per le testate che lo ricevono all'interno delle categorie scientifiche in cui è organizzato il JCR: ogni testata può essere assegnata a più categorie, all'interno delle quali occupa un posto più o meno alto, suscettibile di cambiamento anno dopo anno; la posizione assunta dalla testata nella graduatoria della sua categoria e quindi l'appartenenza ad un determinato quartile, dovrebbe garantire la comparabilità del numero IF tra i diversi settori scientifici, malgrado le diverse prassi citazionali di questi ultimi, che incidono in modo difforme sulla creazione del numero. In altre parole, per avere un quadro della situazione, è sufficiente dare un'occhiata alle categorie MEDICINE, GENERAL & INTERNAL e MATHEMATICS, nelle quali nell'anno 2008 la rivista più in alto in graduatoria ha rispettivamente $IF=50.017$ e 3.806 , numeri difficilmente confrontabili tra loro, dovuti appunto alle consuetudini citazionali dei rispettivi ambiti scientifici. Vale la pena infine di ricordare che esiste

un IF aggregato per categoria, che viene calcolato con analoghe modalità in base ai dati di tutti i periodici che vi sono inclusi.

4. Applicazione del metodo per la comparazione della produttività scientifica di istituti operanti nel campo dell'inquinamento atmosferico

Il metodo è stato inizialmente applicato a 7 istituti (3 francesi, 3 tedeschi e 1 norvegese), che conducono attività di ricerca nel campo dell'inquinamento atmosferico. Come sottolineato nel paragrafo introduttivo, questo tipo di analisi consente di confrontare fra di loro Istituti che, nell'ambito della ricerca scientifica in campo ambientale, si dedicano ad uno stesso tema di ricerca. Il confronto quindi è efficace sia per paragonare la produzione scientifica di Istituti diversi di un unico Paese, sia per confrontare questi con istituti stranieri che operano sulla stessa tematica di ricerca.

Per i 7 Istituti coinvolti sono stati desunti, mediante la banca dati ISI, il numero di pubblicazioni nel quinquennio 2003-2007 (P '03-'07), il numero di citazioni al 2007 (CIT '07) e l'H Index dell'istituto al 2009 (HI '09). Il parametro CPP è stato ottenuto dividendo il numero di citazioni per il numero di paper. Mediante l'analisi dettagliata dei report annuali di ciascun istituto sono stati ricavati i valori relativi al numero di ricercatori operante presso l'istituto. E' stato quindi possibile determinare il valore di PPR, dividendo il numero di paper per il numero di ricercatori dell'Istituto. I dati ottenuti sono riportati nella Tabella 4.1 e nelle Figure 4.1, 4.2, 4.3 sono riportati, in istogrammi, rispettivamente i valori di CPP, PPR e HI '09.

Questi tre parametri hanno il pregio di indicare rapidamente ed efficacemente l'operosità dell'Istituto e l'impatto del suo lavoro sulla comunità scientifica. Infatti, mentre PPR indica la produttività media nei cinque anni in esame del singolo

ricercatore, CPP e HI '09 indicano rispettivamente l'impatto della singola pubblicazione e del lavoro dell'Istituto, preso nella sua globalità, sul mondo scientifico.

ISTITUTO	N. Ric.	P '03-'07	CIT '07	CPP	PPR	HI '09
MPIMET Max Plank Inst. for Meteorology (DE)	160	465	4641	9,98	2,91	46
IFT Leibniz Inst. for Tropospheric Research (DE)	84	359	1947	5,42	4,27	28
NILU Norwegian Inst. for Air Research (NO)	72	298	2253	7,56	4,14	32
LSCE Lab. des Sci. du Climat et de l'Environ. (FR)	187	804	6396	7,96	4,30	50
MPIC Max Plank Inst. for Chemistry (DE)	140	1027	8804	8,57	7,34	59
LA Lab. d'Aerologie (FR)	63	220	1291	5,87	3,49	23
LMD Lab. de Meteorologie Dynamique (FR)	80	367	2189	5,96	4,59	32

Tabella 4.1: dati bibliometrici dei 7 istituti Europei operanti nell'ambito dell'inquinamento atmosferico

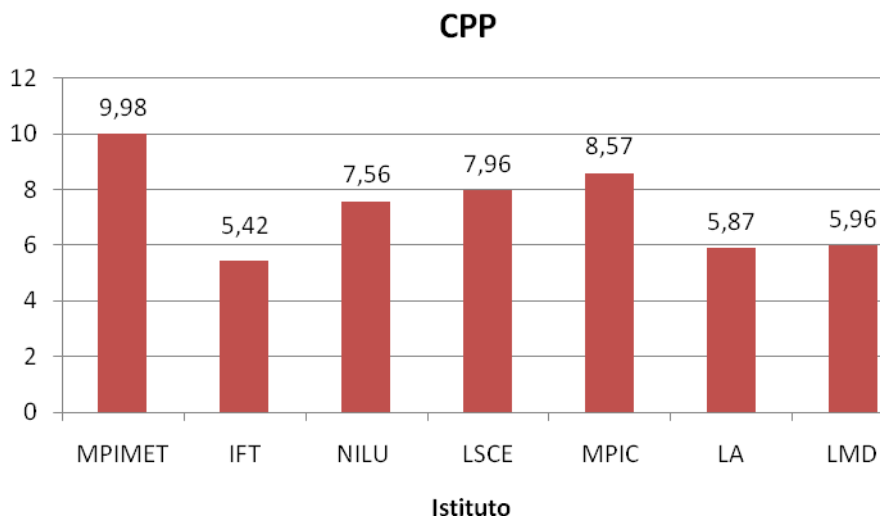


Figura 4.1: confronto fra i 7 Istituti Europei in termini di - CPP

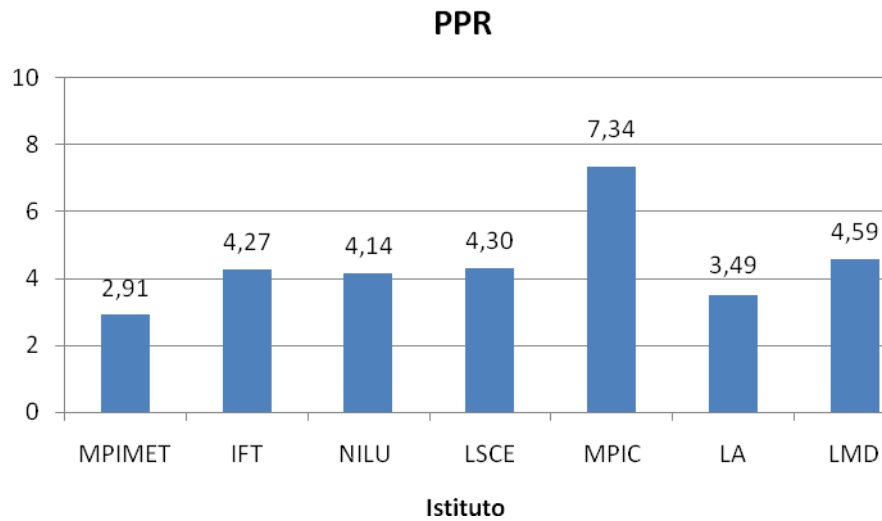


Figura 4.2: confronto fra i 7 Istituti Europei in termini di PPR

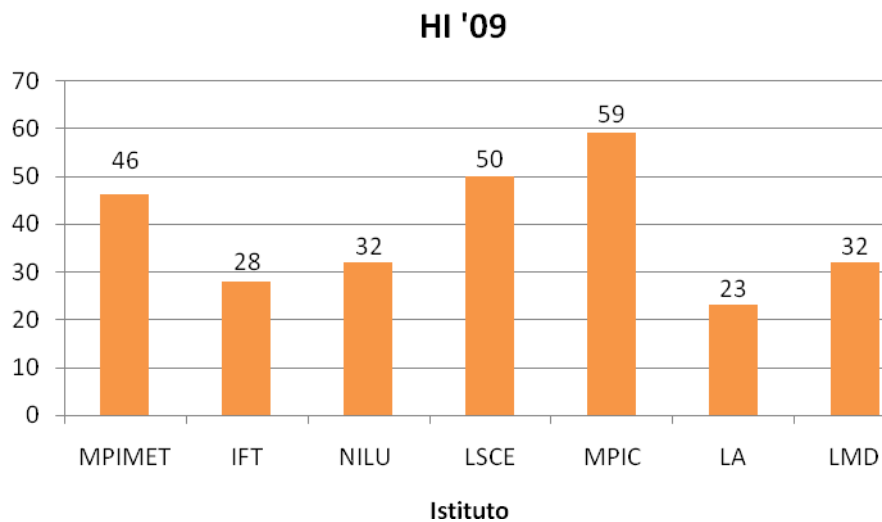


Figura 4.3: confronto fra i 7 Istituti Europei in termini H index, calcolato al 2009

5. Confronto fra i 37 Istituti considerati e relativa rappresentazione grafica

Nella tabella 5.1 vengono proposti in successione, ordinati per numero crescente di ricercatori, i dati bibliometrici relativi ai 37 Istituti esaminati. Come

nella precedenti tabelle, si riportano il numero di pubblicazioni dal 2003 al 2007 (P'03-07), il numero di citazioni (CIT '07), il numero di citazioni per pubblicazione (CPP), il numero di pubblicazioni per ricercatore (PPR) e H Index al 2009 (HI '09).

	ISTITUTO	N. Ric.	P '03-'07	CIT '07	CPP	PPR	HI '09
1	ENSTIB Ecole Nat. Sup. des Technol. des Industries du Bois (FR)	33	104	223	2,14	3,15	10
2	WKI Fraunhofer Inst. of wood res. - Wilhelm Klauditz Inst. (DE)	47	42	59	1,40	0,89	7
3	BGI Bayerisches Geoinstitut (DE)	60	455	2154	4,73	7,58	29
4	IVL Swedish Environ Research Inst (SWE)	60	111	1005	9,05	1,85	17
5	IBAF Ist. di Biologia Agro-ambientale e Forestale CNR (IT)	61	91	842	9,25	1,49	22
6	LA Lab. d'Aerologie (FR)	63	220	1291	5,87	3,49	23
7	ICG-2 Inst. for Chemistry and Dynamics of the Geosphere (DE)	70	236	891	3,78	3,37	23
8	NILU Norwegian Inst. for Air Research (NO)	72	298	2253	7,56	4,14	32
9	IDPA Ist. per la Dinamica dei Processi Ambientali CNR (IT)	79	138	482	3,49	1,75	15
10	LMD Lab. de Meteorologie Dynamique (FR)	80	367	2189	5,96	4,59	32
11	IMAA Ist. di metodologie per l'analisi ambientale CNR (IT)	82	217	583	2,69	2,65	16
12	IFT Leibniz Inst. for Tropospheric Research (DE)	84	359	1947	5,42	4,27	28
13	ISE Ist. per lo studio degli ecosistemi CNR (IT)	87	152	362	2,38	1,75	14
14	HCMR Hellenic Centre for Marine Research (GR)	90	274	591	2,16	3,04	17
15	NIVA Norwegian Inst. water Research (NO)	100	161	1511	9,39	1,61	19
16	IBED Inst. for Biodiversity and Ecosystem Dynamics (NL)	100	782	7168	9,17	7,82	40
17	IRSA Istituto per la ricerca sulle acque CNR (IT)	105	148	853	5,76	1,41	15
18	IGS Inst. of Geological Science, University of Bern (CH)	123	365	1648	4,52	2,97	26
19	MPIB Max Plank Inst. of Biogeochemistry (DE)	124	412	4955	12,0	3,32	50
20	IRNA Inst. de Recursos Naturales y Agrobiol. Sevilla CSIC (ES)	124	287	1293	4,51	2,31	25
21	IGG Ist. di Geoscienze e Georisorse CNR (IT)	133	529	1894	3,58	3,98	27
22	MPIC Max Plank Inst. for Chemistry (DE)	140	1027	8804	8,57	7,34	59
23	ICM Inst. de Ciències del Mar CSIC (ES)	150	578	2568	4,44	3,85	31
24	MPIMET Max Plank Inst. for Meteorology (DE)	160	465	4641	9,98	2,91	46
25	ISMAR Ist. di scienze del mare CNR (IT)	175	321	894	2,79	1,83	18
26	IE Inst. of Entomology, Academy of Science (CZ)	175	330	1519	4,60	1,89	23
27	ISAC Ist. di Scienze dell'Atmosfera e del Clima CNR (IT)	178	423	1885	4,46	2,38	28
28	SYKE Finnish Environ Inst (FI)	180	36	374	10,3	0,2	11
29	LSCE Lab. des Sciences du Climat et de l'Environ. (FR)	187	804	6396	7,96	4,30	50
30	EAWAG Swiss Fed Inst Aquat Sci & Technol (CH)	300	777	9968	12,8	2,59	41
31	EEZ Estacion Experimental del Zaidin CSIC (ES)	325	489	2212	4,52	1,50	27
32	CEH Centra for Ecology & Hydrology (UK)	500	326	3292	10,1	0,65	27
33	BGR Bundesanstalt fur Geowissenschaften und Rohstoffe (DE)	654	355	945	2,66	0,54	20
34	BGS British Geological Survey (UK)	700	792	2800	3,54	1,13	29
35	BRGM Bureau de Recherches Geologiques et Minieres (FR)	765	557	2031	3,65	0,73	21
36	PSI Paul Sherrer Inst. (CH)	1118	3540	19711	5,57	3,17	67
37	IFREMER Inst. Français de Rech. Exploit. de la Mer (FR)	1320	1716	7005	4,08	1,30	35

Tabella 5.1: dati bibliometrici dei 37 Istituti esaminati

Per poter confrontare le performance di questi Istituti, si è pensato di farne una rappresentazione attraverso l'utilizzo di un grafico a bolle (Figura 5.1). Sull'asse delle ascisse è riportato il valore PPR, mentre su quello delle ordinate è riportato il valore CPP. La dimensione della bolla è rappresentativa del numero di ricercatori dell'Istituto.

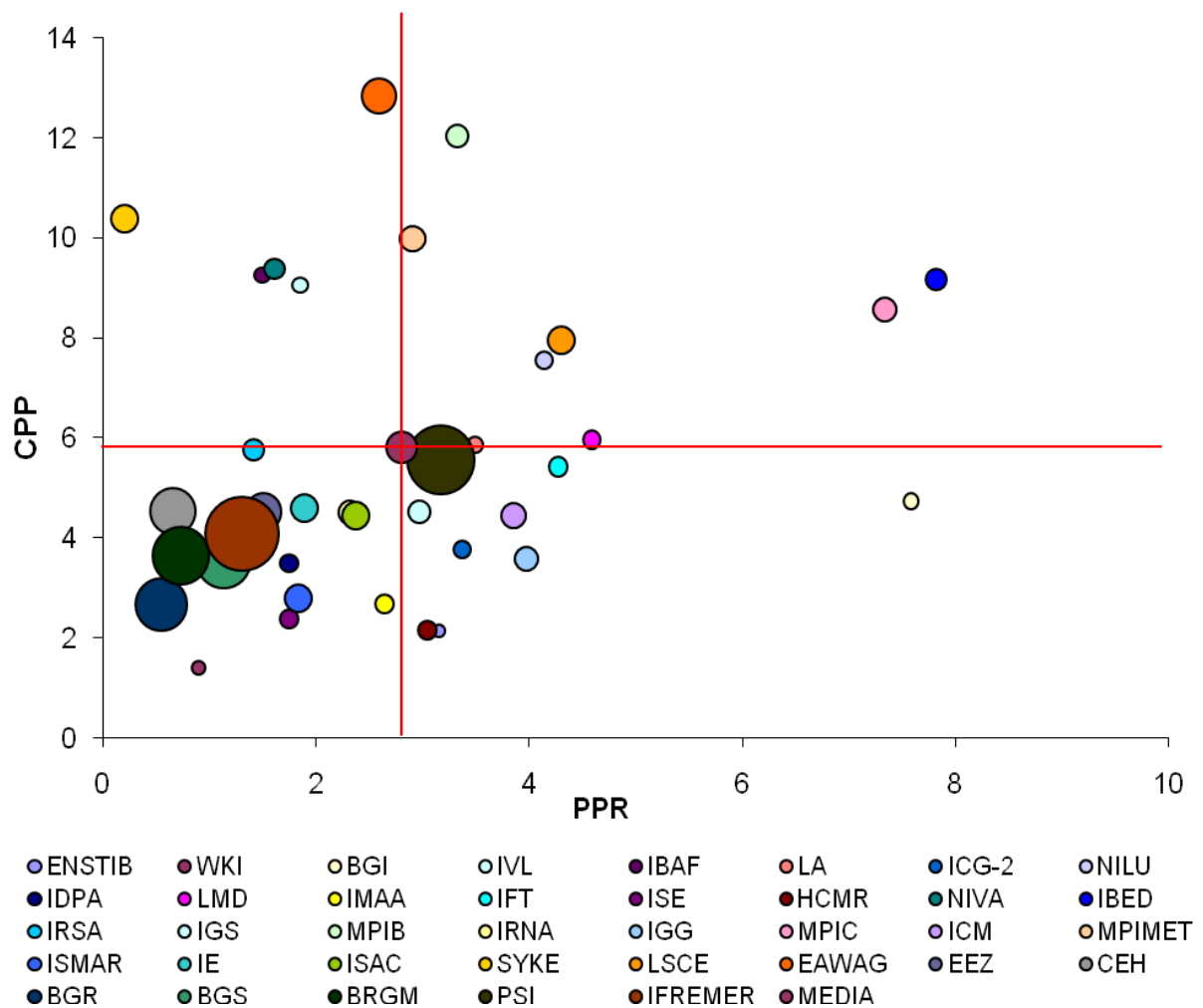


Figura 5.1: diagramma a bolle rappresentativo di PPR, di CPP e del numero di ricercatori

Le due linee rosse, che si intersecano in corrispondenza del centro della bolla rappresentativa della media calcolata sui 37 Istituti, dividono in quattro parti il grafico. Vedendo in quale dei quattro quadranti si collocano gli Istituti è possibile

vedere rapidamente come le loro performance in termini di CPP e PPR si pongono rispetto al valore medio. Introducendo anche la grandezza della bolla come terzo parametro da analizzare, si può collegare l'efficienza dell'istituto alla sua massa.

Nel quadrante in alto a destra si collocano gli Istituti che hanno un alto numero di PPR e contemporaneamente di CPP. Questi quindi rappresentano gli Istituti più performanti fra quelli esaminati. Se si considera il numero di ricercatori di questi Istituti, si può osservare che hanno al massimo 200 ricercatori. Nel quadrante in basso a sinistra invece si collocano gli istituti con i valori più bassi di PPR e CPP. Si può osservare che in generale sono presenti in questa parte del grafico gli Istituti con più alto numero di ricercatori. Nei restanti due quadranti invece si collocano gli istituti che risultano eccellere o rispetto al parametro PPR (quadrante in basso a destra) o rispetto a CPP (quadrante in alto a sinistra), presentando tuttavia performance più scarse rispettivamente in termini di numero di citazioni e numero di pubblicazioni.

Come si può osservare, in un'ottica di confronto delle performance degli istituti, il grafico a bolle rappresenta uno strumento molto utile perché consente di valutare tre parametri contemporaneamente e vedere se fra questi sussiste una qualche correlazione.

In questa sede l'unico metro di valutazione delle performance degli Istituti è rappresentato dalle pubblicazioni e dal numero di citazioni. Per un'analisi più approfondita, sarebbe sicuramente interessante associare ad ogni Istituto altri parametri che ne possano stimare l'attività scientifica, quali ad esempio il numero di brevetti e di spin off, il numero di progetti nazionali ed internazionali (con relativo importo) ai quali l'Istituto ha preso parte e i corsi universitari che hanno visti coinvolti ricercatori dell'Istituto. Questo approfondimento consentirebbe di capire se gli Istituti, che hanno dimostrato una performance più bassa in termini di CPP e PPR, concentrano maggiormente la loro attività in campi che danno luogo a prodotti scientifici diversi dalle pubblicazioni. E' questo il caso degli Istituti a più alto numero di ricercatori, che dalla nostra analisi sono risultati i meno performanti in termini di pubblicazioni.

6. Esercizio per la comparazione della produttività scientifica di Paesi diversi

In questa parte finale del lavoro, il metodo è stato utilizzato per il confronto di Istituti, operanti nell'ambito della ricerca scientifica in campo ambientale, appartenenti a due diversi Paesi Europei. Gli Istituti esaminati conducono attività di ricerca nel settore dell'inquinamento dell'atmosfera, delle acque, dei suoli, nel campo delle georisorse etc. Non vi è pertanto, come nel primo caso esaminato, un unico settore di ricerca.

Nelle Tabelle 6.1 e 6.2 sono riportati i dati bibliometrici desunti per i 6 Istituti francesi e gli 8 Istituti tedeschi considerati. In questo caso, oltre ai parametri già utilizzati precedentemente (CPP, PPR e HI'09), è stato estrapolato anche il parametro PPS, paper per staff, che indica il numero medio di pubblicazioni nei cinque anni considerati realizzati dalla singola unità di personale dell'istituto. Si comprende quindi anche il personale tecnico amministrativo.

FRANCIA - Istituti	N. Staff	N. Ric.	P '03-'07	CIT '07	CPP	HI '09	PPS	PPR
LSCE Lab. des Sciences du Climat et de l'Environnement	252	187	804	6396	7,96	50	3,19	4,30
LA Lab. d'Aerologie	96	63	220	1291	5,87	23	2,29	3,49
LMD Lab. de Meteorologie Dynamique	160	80	367	2189	5,96	32	2,29	4,59
IFREMER Inst. français de recherche pour l'exploitation de la mer	1482	1320	1716	7005	4,08	35	1,16	1,30
BRGM Bureau de Recherches Geologiques et Minieres	850	765	557	2031	3,65	21	0,66	0,73
ENSTIB Ec. Nationale Superieure des Technologies des Industries du Bois	54	33	104	223	2,14	10	1,93	3,15

Tabella 6.1: dati bibliometrici dei 6 istituti francesi

GERMANIA – Istituti	N. Staff	N. Ric.	P '03-'07	CIT '07	CPP	HI '09	PPS	PPR
MPIMET Max Plank Inst. for Meteorology	200	160	465	4641	9,98	46	2,33	2,91
IFT Leibniz Inst. for Tropospheric Research	154	84	359	1947	5,42	28	2,33	4,27
MPIC Max Plank Inst. for Chemistry	242	140	1027	8804	8,57	59	4,24	7,34
MPIB Max Plank Institute of Biogeochemistry	197	124	412	4955	12,03	50	2,09	3,32
BGR Federal Institute for Geosciences and Natural Resources	726	654	355	945	2,66	20	0,49	0,54
WKI Fraunhofer – Inst. of wood research - Wilhelm Klauditz institut	87	47	42	59	1,40	7	0,48	0,89
ICG2 Inst. for Chemistry and Dynamics of the Geosphere	71	70	236	891	3,78	23	3,32	3,37
BGI Bavarian Research Inst. of experimental geochemistry and geophysics	70	60	455	2154	4,73	29	6,5	7,58

Tabella 6.2: dati bibliometrici degli 8 istituti tedeschi

Per confrontare gli Istituti, i dati delle precedenti tabelle sono stati elaborati in modo tale da rendere il parametro bibliometrico riconducibile al Paese. In particolare, per ogni Paese sono stati sommati i papers realizzati dai suoi Istituti dal 2003 al 2007 ottenendo il parametro $\sum P '03-'07$ e le citazioni totali ottenendo il parametro $\sum CIT '07$. Dividendo $\sum CIT '07$ per $\sum P '03-'07$ è stato ottenuto il numero medio di citazioni per paper (CPP). Dividendo $\sum P '03-'07$ per il numero totale del personale (staff) e dei ricercatori di tutti gli Istituti francesi, è stato possibile ottenere PPS e PPR per la Francia. Ugualmente è stato fatto per gli Istituti tedeschi. Infine, per ottenere il dato relativo all'H Index, è stata fatta una media dei valori dell'H index dei singoli Istituti. Quindi, la media dell'H index degli Istituti francesi viene considerata come l'H Index rappresentativo della Francia, quello risultante dalla media degli H index degli Istituti tedeschi è il dato rappresentativo dell'H Index della Germania. Nella Tabella 5.3 vengono riportati, per i due Paesi considerati, i dati sopra citati.

	ΣP '03-'07	ΣCIT '07	CPP	PPS	PPR	HI '09 medio
FRANCIA	3768	19135	5,08	1,30	1,54	28,5
GERMANIA	3351	24396	7,28	1,92	2,50	32,75

Tabella 6.3: confronto fra dati bibliometrici degli istituti francesi e tedeschi

Per un confronto visivo delle performance, in Figura 6.1 sono riportati gli istogrammi rappresentativi dei valori di CPP, PPS e PPR.

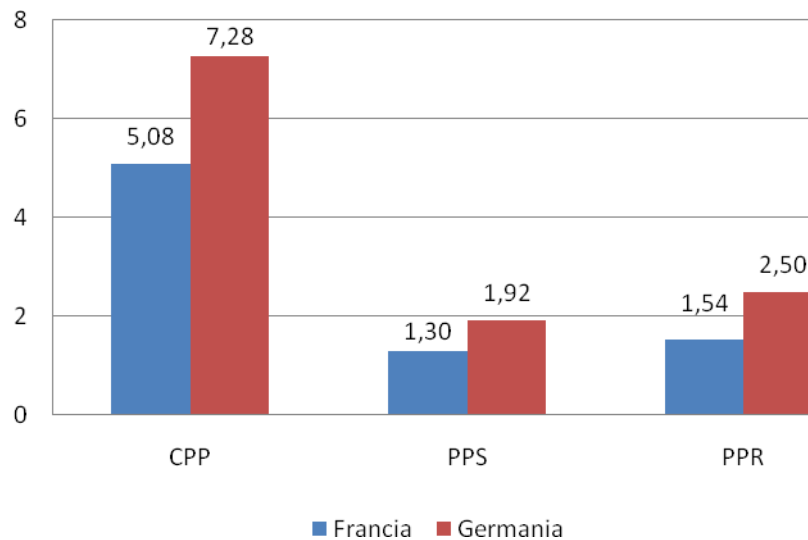


Figura 6.1: istogrammi di confronto in termini di CPP, il numero di paper per unità di staff (PPS) e PPR

Infine, come già anticipato nel paragrafo introduttivo, si vuole ribadire che per confrontare la produttività scientifica complessiva di un Paese in un dato settore, è necessario possedere i dati bibliometrici di tutti gli istituti che in quel Paese sono attivi in quello specifico ambito. Solo così si potrebbero elaborare dei parametri bibliometrici rappresentativi della produttività globale.

7. Considerazioni conclusive

Questo studio ha avuto lo scopo di illustrare una metodica per la determinazione della produttività scientifica degli istituti di ricerca Europei, sia dal punto di vista qualitativo che quantitativo. La metodica, basata sull'analisi dei dati presenti nella banca dati ISI Web of Knowledge è stata applicata a 37 Istituti di ricerca operanti in ambito ambientale e appartenenti a 12 diversi Paesi Europei.

Inizialmente la metodica proposta è stata utilizzata per determinare la produttività di 7 Istituti di 3 diversi Paesi, che conducono attività di ricerca nel settore dell'inquinamento atmosferico. In questo caso, il metodo ha permesso di confrontare Istituti diversi di un unico Paese e al contempo di confrontare questi con Istituti stranieri che si occupano della stessa tematica di ricerca.

I dati raccolti per i 37 Istituti sono stati raccolti in una tabella e rappresentati mediante un grafico a bolle, che ha consentito di confrontare le loro performance sulla base dei parametri PPR e CPP, correlandoli alla dimensione dell'Istituto. Questo ha consentito di fare alcune considerazioni in merito alla massa critica di un Istituto.

Infine è stato fatto, è stato fatto un tentativo di applicazione del metodo per la comparazione di Istituti francesi e tedeschi. In questa occasione i parametri bibliometrici utilizzati per i singoli Istituti sono stati elaborati in modo tale da renderli riconducibili al Paese, ottenendo così indici rappresentativi della produttività scientifica nazionale. Gli indici che sono stati elaborati consentono un confronto sia in termini di numero di pubblicazioni (attraverso i parametri PPR e PPS) sia in termini di numero di citazioni (attraverso i parametri CPP e HI'09).

Lo studio di comparazione degli istituti potrebbe essere ulteriormente approfondito mediante la stima della capacità produttiva del singolo ricercatore dell'Istituto sulla base delle risorse economiche a disposizione. Pertanto, potendo accedere alle informazioni inerenti ai fondi a disposizione dell'Istituto, sia di tipo interno che provenienti da progetti esterni, si arriverebbe al confronto fra i ricercatori dei singoli Istituti e, mediando questi valori per Paese, anche in questo

caso il confronto potrebbe essere esteso a livello nazionale e permettere il confronto fra Paesi diversi.

Mentre l'analisi appena fatta consentirebbe di approfondire lo studio relativamente alla quantità della produzione scientifica di un istituto, uno studio bibliometrico puro delle pubblicazioni potrebbe portare a qualche conclusione sulla qualità scientifica delle stesse. In questa prospettiva sarebbe possibile, date le categorie del JCR in cui ricadono le testate oggetto delle pubblicazioni di un determinato ente di ricerca, desumerne l'IF aggregato e medio e vedere come si collocano rispetto a questi numeri le singole riviste, su cui quell'ente pubblica; sarebbe inoltre possibile effettuare dei confronti su questa base tra i diversi paesi e verificare quale sia il contributo offerto dai diversi enti all'IF di categoria. Da tutt'altro punto di vista, e cioè in un'ottica di valutazione degli strumenti da cui si tirano fuori i dati, sarebbe altresì interessante vedere come cambiano i numeri a seconda del database utilizzato; abbiamo prima accennato al fatto che il WoS non è più l'unico strumento a fare valutazioni di questo tipo, ma, accanto a piattaforme che rispondono ad una logica di "archivi aperti", ancora, tuttavia, di uso un po' farraginoso, esistono altri database (ad esempio Scopus), che, ancorché meno blasonati, cominciano ad assumere caratteristiche analoghe a quelli dell'ISI: di qui l'interesse a vedere quale tipo di oscillazione può ricavarsi dall'usare l'uno o l'altro.

8. Bibliografia

1. G. Abramo, C.A. D'Angelo, M. Solazzi (2009), Assessing public-private research collaboration: is it possible to compare university performance?, *Scientometrics*, published online: 14 November 2009 (DOI 10.1007/s11192-009-0104-0).
2. M. Coccia (2009), Research performance and bureaucracy within public research labs, *Scientometrics* 79 (1): 93-107.
3. M. Coccia (2005), A scientometric model for the assessment of scientific research performance within public institutes, *Scientometrics* 65 (3): 307-321.

4. J.S. Dietz, B. Bozeman (2005), Academic careers, patents, and productivity: industry experience as scientific and technical human capital, *Research policy* 34: 349-367.
5. L. Egghe (2006), Theory and practice of the g-index, *Scientometrics* 69 (1): 131-152.
6. J. Guan, J. Wang (2004), Evaluation and interpretation of knowledge production efficiency, *Scientometrics* 59 (1): 131-155.
7. J.E. Hirsch (2005), An index to quantify an individual's scientific research output. *PNAS* 102 (46): 16569–16572
8. S.M. Nejati, Hosseini Jenab (2009), A two-dimensional approach to evaluate the scientific production of countries (case study: the basic sciences), *Scientometrics*, published on line 7 November 2009 (DOI 10.1007/s11192-009-0103-1).