Evaluation of Strategic Research Programs: The Case of Danish Environmental Research 1993-2002

Peter Ingwersen and Birger Larsen

pi@db.dk, blar@db.dk

Department of Information Studies, Royal School of Library and Information Science,
Birketinget 6, DK 2300 Copenhagen S (Denmark)

Abstract

The article reports on the mid-term and final scientometric evaluations of the Danish Strategic Environmental Research Program (named SMP). SMP consisted of nine virtual research centres during the period 1993-97. Citations are measured 1993-2002. Central indicators are: Centre Impact Factor (CIF), the number of citations received by each centre's SCI-articles; centre Journal Impact Factor (JIF), which is a diachronic IF per journal volume publishing a centre article. Citation and publication data are obtained from the Thomson-Dialog online version of Science Citation Index (SCI). Other indicators applied include the weighted Danish and World domain impact from the National Science Indicators in subject areas selected by the centres.

Top-ranked journal volumes used in SMP in terms of JIF scores were correlated with the corresponding articles' citation values. At the mid-term assessment the Pearson coefficient showed a strong correlation, which disappeared at the final evaluation – except for a few high-impact centres. The publication behaviour varied substantially between centres. Domain indicators weighted according to SMP centre profiles demonstrate significant differences to un-weighted scores. Although SMP as program did not make a strategic difference, measured as a CIF-score 10 % higher than all other indicator values and in particular the Danish domain impact, four centres clearly did. USA was the most knowledge importing country. In a wider perspective a two-step evaluation of, in particular, cost-heavy strategic programs have important implications for the continuation, volume and direction of national research funding and activities.

Introduction

This article reports on the final scientometric evaluation carried out 2003 of the Danish Strategic Environmental Research Program (named SMP) that consisted of nine virtual research centres 1993-1997. The motivation behind this report is threefold. First, a mid-term evaluation, made 1999 and covering the publications 1993-95 as well as citations 1993-1998, resulted in assumptions concerning the centres' publication strategies. When researchers published in top-impact journals, their articles also received many citations (Ingwersen, Larsen & Wormell, 2000; Ingwersen & Larsen, 2005). This might be turned into a fruitful publication strategy for future research in the area. The correlation between diachronic journal impact and the impact of the articles published in the corresponding volumes was then assessed in the final evaluation. Since two-step assessments of entire research programs are rarely done, such mid-term strategic assumptions are seldom tested empirically. Secondly, we wanted to observe whether SMP made a (strategic) difference to the rest of the Danish environmental field and the corresponding world research during the same period. Third: In a wider perspective a two-step evaluation of so-called strategic research programs are important to carry out, since they have implications for the continuation, volume and direction of funding and research activities in the particular field. Such programs may heavily influence the research output from the rest of the research community(ies) concerned over a larger period of time.

SMP attracted approximately 95 million € (700 million DKK) of public funds over the 5 year period (Fisker, 2004). 600 Danish and international researchers participated on an interdisciplinary basis from a range of institutions, connected by Internet communications. Originally SMP consisted of 16 objectives distributed over 13 centres, including humanistic ones. For this reason, only nine centres could be analyzed fairly for citations. They are, with objectives in ( ):

1. Air Pollution Processes & Models (Atmosphere & air pollution)
2. Terrestrial Ecosystem Research (Atmosphere & air pollution)
3. The Groundwater Group (Groundwater; pesticides in groundwater)
4. Agricultural Biodiversity (Soil surface)
5. Root Zone Processes (Soil surface)

Published as:
Ingwersen, P. and Larsen, B. (2007): Evaluation of strategic research programs: the case of Danish environmental research 1993-2002. Research Evaluation, 16(1), 47-57.

6. Freshwater Environmental Research (Freshwater and marine areas)
7. Strategic Environmental Research in Marine Areas (Freshwater and marine areas)
8. Danish Centre for Eco-toxicological Research (Environmentally hazardous substances in the aquatic and terrestrial ecosystems)
9. Centre for Biochemical and Occupational Epidemiology (Human health)

A few other citation analyses of interdisciplinary environmental research have been done, for instance recently on forestry research (Steele & Stier, 2000). Evaluations, including mid-term assessments, are not common. The SMP program is also interesting owing to its mixture of hard science fields with medical and more social science-related disciplines.

The article is organized as follows. The data collection and analysis methods, including the applied indicators, are briefly described. This is followed by the overall results from the mid-term and final evaluations across the nine centres. Indicator results are compared to (un)weighted Danish and world indicator measures, respectively. The original correlation coefficients (Pearson) from the mid-term assessments are compared to the final ones, and the implications of the central results for the strategic program, and evaluation methods in wider perspective, are discussed in the ensuing section.

Data Collection and Analysis Methods

Data was collected from two sources: the online version of Science Citation Index (SCI) hosted by Thomson-Dialog and National Science Indicators (NSI), constructed by Institute for Scientific Information (ISI), 2001. Each centre provided a list of research publications. For the mid-term evaluation the lists covered the period 1993-95. Similarly, the centres provided a supplementary list for the final evaluation covering 1996-97. Hence, at the time of the mid-term evaluation in 1998/99 the entire volume of research output was actually known – but was not fully explored prior to the final evaluation owing to small citation windows. The entries of the lists were all searched online in SCI to establish whether the applied journal was indexed by SCI or not. The non-SCI journals tend to be broader practice-related international journals or magazines in Danish. If indexed, then the entry was verified and journal names were checked against the SCI journal name index, in order to conform the journal data across all centres and time periods. 434 internationally published journal articles constitute the total data population at the final evaluation, with 344 indexed by SCI (79 %). At the mid-term assessment the number of publications was 201 and 151 (75 %), respectively.

The number of citations received up to a given year (1998 and 2002, respectively) was retrieved online for each article, whether being originally indexed in SCI or not, *and* from each corresponding journal publication year. In that way, we are able to calculate the diachronic journal impact factor (JIF) (Egghe & Rousseau, 1990) online (Christensen & Ingwersen, 1996) for each time a journal was used by a centre. This type of JIF is a fair and *realistic* impact factor, in contrast to the much-criticized synchronous JIF produced annually by ISI (Seglen, 1997), because the diachronic JIF can be compared directly to the real impact of the corresponding research articles. This method is similar to the one applied by van Rann (1999) on offline data, but in the present analysis publicly available online data is used.

When summed up for each centre, and for SMP as such, the number of citations received by the SCI-articles constitutes one primary indicator: the *Centre Impact Factor* (*CIF*). Similarly, the corresponding sum of JIFs per centre and SMP as a program establishes another primary indicator – the *centre JIF*. An overall CIF indicator (*CIF**) corresponds to the total of citations found in SCI to *all* centre publications.

The Danish and the World citation impacts per centre, based on the scientific fields for each centre, constitute secondary indicators. They are calculated by the application of NSI, and are comparable to the CIF and centre JIF covering the same time windows. Each centre had previously pointed out the relevant NSI subject fields from Current Contents that corresponded to their research area. For SMP as a program all the NSI categories applied to all the nine centres were summed up. This means that the Danish and world citation impacts are calculated in terms of *weighted* subject profiles (van Raan, 1999; Ingwersen, Noyons and Larsen, 2001). They mirror the distribution of NSI categories over SMP as a program, as defined by its centres. One may say that they act as a kind of '*shadow*' SMP program. For instance, the category 'Environment/ecology' appears 8 times and 'Biochemistry & biophysics' three times, etc., in the final SMP profile. Since data was not available covering the entire period, NSI data covering 1991-2000 was used to simulate the actual period, 1993-2002. The assumption is that trends in Danish and world impact and volume are similar within such a short time shift.

The Pearson correlation coefficient was applied to the top-25 (mid-term) and top-40 SCI-*journal volumes* (final evaluation), according to their individual JIFs. This means that the diachronic JIF for each volume was paired to the number of citations received by the corresponding article published in that volume. We applied Pearson although the distribution of citations over journals and articles is not a normal distribution – but skewed. In four selected high-impact centres we checked the Pearson results against the average correlation coefficient for the research program.

Finally, also as secondary indicators we observe the patterns of *knowledge export* by means of listing by frequency a) the citing countries and b) the subject categories in the journals that cite a centre. Identical sets of indicators are applied to the two evaluations for comparative reasons. This was done online by means of the Thomson-Dialog Rank command facility (Wormell, 1998).

Major Results of the Analyses

First, we report the results concerned with publication activities and citation impact comparisons. This is followed by the major findings concerned with knowledge export over the entire time window of SMP. Finally, we demonstrate the article-top-journal impact correlation results, both with respect to the mid-term and the final evaluation of the program, as well as for high-impact centres at the final stage.

Publication Activity

The SMP research publication activity in terms of the 344 SCI indexed articles is displayed in Table 1. The number of articles doubles from 1993-1994 and re-doubles from 1994-1996. It triples from 1993-1995. In comparison the entire Danish production in Environmental-ecological research rises from 149 to 256 articles, Table 1. This seems to indicate that the environmental research community remaining outside SMP does *not keep up* a steady line of productivity towards the end of the observed period. SMP seems increasingly dominant on the Danish scene. In 1993/94 SMP constitutes approx. 19 % of all Danish environmental output, but covers approx. 34 % in 1995/96, increasing to approx. 39 % in 1997.

Table 1. SCI-indexed articles published by all centres combined in the SMP program and Danish environmental research (Sources: the centres & NSI, 2001)

Publication Year	No. of SCI-articles	No. of Danish SCI-articles	Proportion of Danish SCI-articles
1993	24	149	16,1%
1994	48	222	21,6%
1995	79	231	34,2%
1996	93	269	34,6%
1997	100	256	39,1%
Total	344	1127	30,5%

The proportion of SCI-articles vs. all articles from SMP as a program was quite stable over the period, but with some variation between the centres. For instance, the Groundwater centre published only 20 SCI-articles out of a total of 35 (57 %) whilst the Freshwater centre produced 36 SCI-articles out of 42 (86 %). The difference between the two kinds of articles essentially mirrors quite different approaches to research publication behaviour among the SMP centres. One of the reasons for the final result of the CIF vs. centre JIF impact scores per centre (see Figure 1 below), probably derives from this dissimilarity in publication behaviour.

Citation Impact Analyses

The final evaluation per centre and for SMP as a program in terms of the primary indicators is displayed in Table 2. It concerns alone the SCI-articles and journals. In addition, the table

demonstrates the Danish and World impact during the same period, as well as the total centre impact (CIF*), for reasons of comparison.
One should note that the average SMP 'Domain DK' and 'Domain World' indicators – in their *un-weighted versions* – would have been 17.49 and 14.93 citations per SCI-publication respectively. The Danish domain impact would hence increase with almost one citation, and surpass the mean SMP *CIF*, while the world impact would decrease slightly[1]. This comparison would have been unfair towards SMP since the program profile's many 'lighter' research areas would disfavor SMP. With a weighted profile imposed on the world research to be compared, the latter world profile acts as a 'shadow' program containing the same proportion of 'light' and 'heavy' research areas as SMP.
Further, 329 of the 344 SCI-articles were cited at least once, corresponding to almost 96 %, including self-citations during the entire period. This figure is higher than the Danish and World shares of cited articles, which is 94 % and 88 % respectively (Source NSI, 2001). At the mid-term evaluation this proportion for SMP as program was slightly lower, i.e., 89 %.

Table 2. Final evaluation of the nine SMP centres. Primary and secondary indicators, publications 1993-1997, cited 1993-2002. Sources: SCI, Thomson-Dialog online version & NSI (ISI).

Centre	SCI Publ.	Citations SCI	*CIF*	Diachronic JIF	Domain DK	Domain World	All Publ.	All Citations	CIF*
Air Pollution	29	446	***15.4***	*12.7*	14.7	11.1	34	481	14.1
Terrestrial ecology	39	529	***13.6***	*11.0*	12.4	11.5	51	553	10.8
Groundwater	20	288	***14.4***	*13.9*	12.7	10.7	35	318	9.1
Agro biodiversity	20	209	***10.6***	*7.0*	11.7	9.7	31	267	8.6
Root zones	54	796	***14.7***	*10.4*	10.7	7.8	67	989	**14.8**
Freshwater	36	1009	***28.0***	*11.0*	20.0	19.9	42	1062	25.3
Marine ecology	33	709	***21.5***	*18.4*	20.0	19.9	38	739	19.4
Eco-toxicology	79	891	***11.3***	*13.2*	15.2	14.2	98	1034	10.6
Biochem. Epidem	34	864	***25.4***	*16.6*	23.1	22.5	38	881	23.2
Mean total - *SMP*	***344***	***5741***	***16.7***	***12.8***	***16.6***	***15.2***	***434***	***6324***	***14.6***

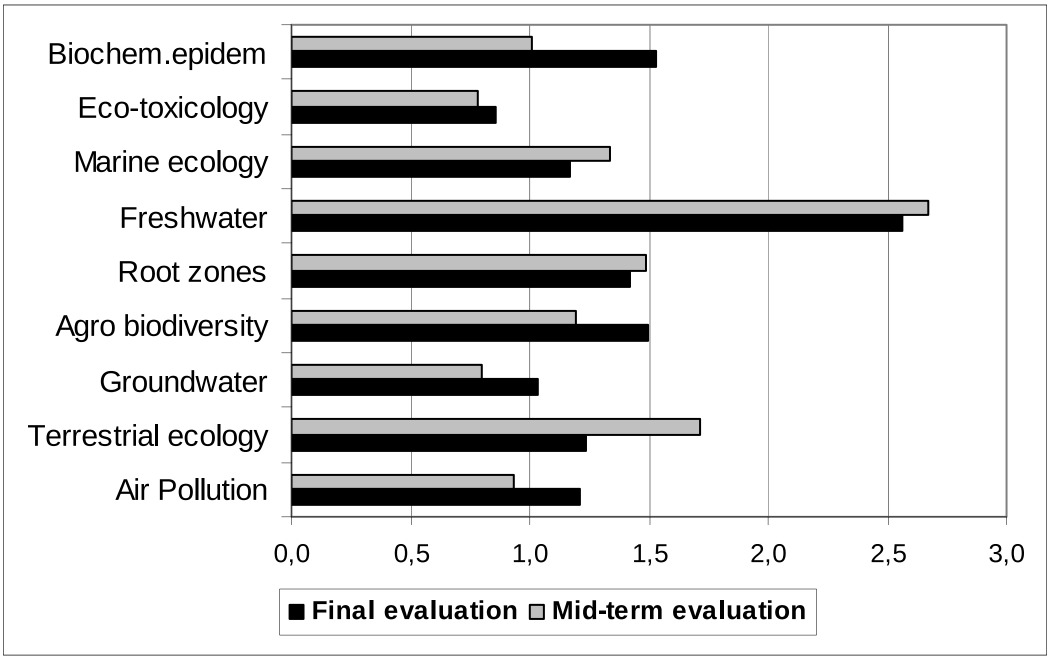

Figure 1. Diagram of the nine SMP centres' citation impact (*CIF*) vs. the corresponding *centre JIF* 1993-2002. Index 1.0 = centre JIF. Source: SCI (ISI, Thomson-Dialog online version).

[1] This supports van Raan's (1999) proposal of applying weighted comparative indicators.

Knowledge Export from the SMP Centres

One kind of knowledge exports concerns which countries that make most use of, i.e., give most citations to, the individual centres. Quite interesting one may observe that USA five times out of nine is the most *highly citing* country with Denmark as the second most citing country. This analysis includes the national self-citations and may hence demonstrate the results of international collaboration – also by means of citations given by partners.

The following cluster of countries constitutes the locations from which ≥ 10 % of each centre's citations are given, i.e., the countries to which at least 10 % of the knowledge export goes; with number of countries represented across the centres in ( ):

1. USA (9) – most citing in 5 centres;
2. Denmark (9) – most citing in 4 centres;
3. Germany (7)
4. United Kingdom (7);
5. The Netherlands (2);
6. Canada (1); France (1); Japan (1); Sweden (1)

A second kind of knowledge export concerns the research areas from which credits are given in the form of citations. We do not demonstrate samples of centre export for lack of space. But obviously the subject categories ranked by the Thomson-Dialog software can be compared to the subject areas chosen by each centre from NSI (i.e., Current Contents) as representative of their research. Checked in this way, it could be observed that, for instance, the Freshwater and Marine centres probably had selected areas (Biology; Biochemistry) somewhat too broad or/and out of tune of their real research foci and of too ambitious nature, i.e., with too high world (and Danish) citation impact. Another trend observed was the knowledge export into applied (engineering) fields from several centres.

The Use of Top-Impact Journals

In the mid-term evaluation we found a quite strong correlation between the diachronic JIF of the applied journal volumes and the citations given to the corresponding articles (Pearson's r = .61 for the 151 journal-article pairs with r^2 = .37; p= .005; CV= .25). For the top-25 journal volumes' diachronic JIF and corresponding articles' citations, r equaled .68 with r^2 = .47 (p= .005; CV= .487) – see Figure 2. One might state that there exists a quite robust correlation between the *expected* average impact of articles in the journal used for publication (the diachronic JIF) and the *actual* citation impact obtained per SMP article over all articles as well as over the top-25 article-journal pairs.

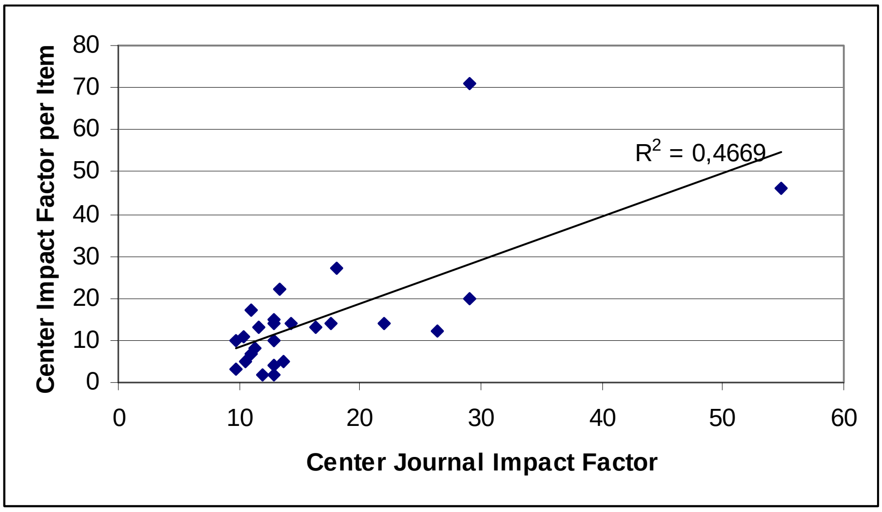

Figure 2. Correlation between the JIF of the top-25 journal volumes and the citation impact of the corresponding articles published by SMP as program 1993-1995, cited 1993-1998.

Effectively this means that almost half of the top-25 SMP articles at the mid-term evaluation obtain high impact when published in high-impact journals. The opposite case, viewing whether the Top-25 high-impact *articles* also appear in high-impact journals, gives a somewhat lower correlation

coefficient (r = .58, with r^2 = .33 (p= .005; CV= .487). The correlation is still above the threshold, but clearly, only a third of the high-impact articles are found in high-impact journals. Two thirds are found elsewhere.
The 151 SMP articles published 1993-95 are also analyzed for their article-journal impact correlation in the extended and final citation window 1993-2002 in order to observe if those articles, earlier on showing the robust correlation, actually did maintain that correlation. The results demonstrate that r = .496; r^2 = .25, with p = .005 and CV = .25. This implies that at the extended citation window of max. 10 years approximately 25 % of the original 151 SMP articles obtain a level of impact corresponding to that received by the journals used for publishing. The correlation coefficient is still substantial but has weakened over the prolonged period.

The top-impact journal - vs. - article correlation changes radically at the final evaluation. Two kinds of correlation analyses were carried out for SMP as such. One observes the top-40 journal volumes from the entire SMP program – see Figure 3. Because the Agro biodiversity centre displayed quite low JIF's the centre did not become represented in the top-40 analysis. Hence, a second analysis took the upper 10 % of the journal volume JIFs per centre. In total 35 JIFs – and all nine centres – are represented.

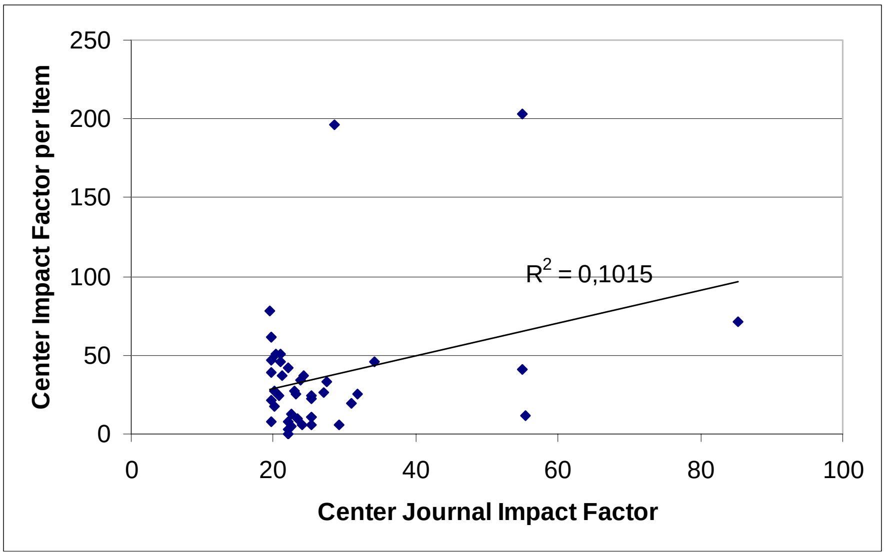

Figure 3. Correlation between the JIF of the top-40 journal volumes and the citation impact of the corresponding articles published by SMP as program 1993-1997, cited 1993-2002.

In the first analysis of the top-40 journals Pearson's r is equal to .32 with r^2= .102 (p = .005; CV= .41). This is an *extremely weak* correlation far below the critical value. Only in the application of one out of ten high-impact journals a correspondingly high article impact was obtained. Even when the four outliers are discharged from the analysis r only increases to .34. The second correlation is similarly weak, r = .38, CV= .42. These scores are far from the mid-term evaluation results demonstrated above. Consequently, in order to observe if at least some of the centres might show a robust correlation coefficient at the final evaluation, we carried out correlation analyses between *all* the journal-article pairs in the four high-impact SMP centres displayed on Figure 1 and Table 4. We also analyzed the centre for Groundwater research for comparative reasons owing to its average impact – Table 3.

Table 3. Pearson correlation coefficients for journal and article pairs of impact for five SMP centres. Critical values correspond to no. of analyzed pairs. 'cit' refers to citation impact obtained by the centre articles (p = .005).

Centre	No. of SCI Publ. (pairs)	All pairs *r*	All pairs r^2	Critical value	Top-20 JIF/cit *r*	Top-20 cit/JIF *r*
Terrestrial ecology	39	0.19	0.04	0.42	-0.23	0.28
Root zones	54	0.44	0.19	0.34	0.54	0.22
Freshwater	36	0.71	0.51	0.42	0.82	0.72
Biochem. Epidem.	34	0.45	0.20	0.42	0.37	0.38
Groundwater	20	0.63	0.39	0.56	0.63	0.63

The analyses also revealed that if one *single outlier* was removed from the correlation analyses of the centres of Terrestrial ecology (journal impact outlier), Root zones and Biochemical epidemics (article impact outliers) the correlations changed:

- Terrestrial ecology, $r = .38$; $r^2 = .14$; the *Outlier*: JIF: 55.5 vs. cit: 12.0
- Root zones, $r = .53$; $r^2 = .27$; the *Outlier*: JIF: 10.03 vs. cit: 70.0
- Biochem. Epidem., $r = .62$; $r^2 = .38$; Top-20 JIF/cit $r = .59$; *Outlier*: JIF: 28.67 vs. cit: 196

In the Freshwater and Groundwater centres no outliers were observed, and the four remaining centres did not show significant outliers.

Discussion

The growth rate from year to year, and over the entire five-year period of research, is impressive and substantially higher than the corresponding Danish production growth – Table 1. Denmark produced in total 1127 SCI-articles including the 344 SMP items during 1993-1997 in environmental research. The growth rate demonstrates that in terms of productivity the SMP program became a clear success.

On the other hand, it can also be argued that SMP increasingly seems to devour resources – like a cuckoo in a nest – and by its increasing domination may inhibit the common development in the rest of the field during the same period (Fisker, 2004). From that stand the strategic research program may not necessarily be seen as giving a national strategic edge to the field in question. Other aspects of the program may indeed provide reinforcement of the field in the long run, e.g., improved research education. But such aspects are not forming part of this evaluation.

The SMP productivity showed also high variation from centre to centre – Table 2. Three centres clearly produced below the average of 38 SCI-articles while four centres were close to average. Similarly, the proportion of *non-SCI articles* varied across centres. The highest ratio was observed for the Groundwater centre (43 % non-SCI articles), and the lowest for the centres of Biochem. Epidem. (11 %), Marine ecology (13 %), and Freshwater (14 %). Nonetheless, this does not mean that centres with a high ratio of non-SCI articles did not receive a high impact, see Tables 2 and 4. The Groundwater centre still kept its above-average *CIF*. The Root zones centre faired even better by means of the non-SCI-articles, Table 2. Only the Agro biodiversity centre (36 % non-SCI articles) demonstrated a weak *CIF* compared to the Danish field impact. In contrast, the Eco-toxicology centre with a lower ratio on non-SCI-articles (19 %) shows a *CIF* far below all other indicators.

Basically, non-SCI-articles serve to mediate research results and applications to a broader audience and are increasingly required by government funding agencies in order to distribute scientific knowledge into society. Notwithstanding, this publication strategy may indeed produce substantial additional citation impact – see, for instance, the Root zone centre, Table 2, for which the CIF* is *higher* than the *CIF*.

Citation Impact Comparisons – the Strategic Difference

During the mid-term evaluation period 1993-95, 151 SCI-articles were produced that yielded 1386 citations 1993-1998. The same 151 articles received 1486 citations 1999-2002 – not a bad yield over this shorter citation window at a later period in such interdisciplinary research areas.

In total the 344 SCI-articles from 1993-97 provided the SMP program with 5741 citations at a mean of *CIF* = 16.7 citations per SCI-article compared to the diachronic centre JIF of 12.8. This difference (23 %) demonstrates that SMP made a *success* of its *publication strategy*.

However, the JIF indicator is *relative*. It displays the *expected* citation impact for the publishing journal volume, i.e., for the publication strategy at a centre as expressed in the journals chosen for publication. It does not state anything about the optimal level of impact value. For that analysis it is very important to look into the balance between the centre JIF and the Domain DK (and World) indicators – Table 2.

The central issue is that the mean centre JIF is far below the Danish weighted citation impact (23 %); all centre JIFs, except for Groundwater, are similarly below the equivalent Danish impact. This demonstrates that the common strategy for publication of research results in SMP was to publish in lower-impact journals, compared to what seems common in the field, also world wide (16 %). In that sense SMP does not demonstrate any strategic difference.

A second issue is that the average SMP *CIF* is almost identical to the weighted Danish domain impact (16.6) – Table 2 – and only 1.5 citations above the weighted world impact. As briefly mentioned above: had the Danish domain impact been analyzed un-weighted, like in most traditional and unfair bibliometric research evaluations, it would in fact have beaten the *CIF* score by .8 citations. However, the world impact would have decreased and been further beaten by the SMP *CIF*.

The reason here is owing to the heavy single weights of research fields like Biochemistry & biophysics and Microbiology in the total SMP spectrum – if all fields are regarded equal in citation impact! Since they are *not* viewed as equal in our weighted method, the fair correspondence is demonstrated in Table 2. This implies also that the Danish (shadow) research profile, as defined by SMP, turns out far better (> 10 %) than the corresponding shadow world spectrum.

In order to demonstrate a clear strategic difference, a SMP centre must – in our opinion – show a rather higher *CIF* than any of the other indicators – for instance at least a 10 percent positive difference. This constitutes a conservative measure. As observed above, the Danish weighted impact factor in itself is more than 10 % higher (1.5 citations) than the world impact; in addition, it obviously contains the SMP *CIF*.

The *Danish research* as such thus demonstrates a *strategic difference* in the Environmental sciences – probably supported substantially by the SMP initiative. On the other hand, the mean SMP *CIF* does *not* demonstrate any clear strategic edge compared to the Danish research. However, four centres on the surface do show a robust average difference in relation to all other indicators, Table 4.

Table 4. SMP centres with equal or more than 10 % difference between *CIF* and the other indicators (source: SCI, Thomson-Dialog online version; NSI, ISI, 2001).

Centre	CIF	Centre JIF	Domain DK	Domain World
Terrestrial ecology	13.6	11.0	12.4	11.5
Root zones	14.7	10.4	10.7	7.8
Freshwater	28.0	11.0	20.0	19.9
Biochem. Epidem.	25.4	16.6	23.1	22.5

The interesting observation is that the first and second centres both show quite homogeneous impact factors whereas the last two centres demonstrate very high *CIF*, Domain DK and Domain World indicators – but rather low *centre JIF*. The former two centres have applied the journals ‘common’ to or slightly lower than the research area worldwide as a publishing strategy. Indeed the Terrestrial ecology centre had one outlier, see above at Table 3, which was a high-impact *journal* in which the SMP article received very few citations. That outlier alone stands for a Centre JIF of *1.4* over all the 39 SCI journal volumes used. Without that outlier the Centre JIF would thus have been 9.6 and the centre CIF improved by 0.3 impact points to 13.9, i.e., showing an even greater strategic difference versus the field.

The latter two centres have applied far *lower-impact journals* than ordinarily used in the areas in general as well as in Denmark. In line with the top-impact journal discussion below, these observations clearly demonstrate that the positive average SMP *CIF* score does *not* really derive from publishing in top-impact journals *and*, at the same time, obtaining much more citations than expected. Rather, the two centres (*Freshwater* and the *Biochem. Epidem*. centre) published in lower-impact-than-average journals, but with success – and thus received a much improved centre impact (*CIF*). In three of the four centres the outliers served to improve the CIFs for each centre by *1.3* (Root zones) and *5.8* (Biochemical epidemics) impact points over all centre articles, i.e., that the latter centre did not at all demonstrate a cutting edge – on the contrary.

This phenomenon is typical for small research entities (Aksnes, 2003) in which one single research article may make the entire difference.

Two other centres, *Air pollution* and *Marine ecology*, are close to have a 10 % advantage over the other indicators, but not quite; and the *Groundwater* centre is in line with its *centre JIF*. The two remaining centres are less advantageous in comparison to the competing indicators. They are directly below the Danish citation impact (*Agro diversity* and *Eco-toxicology*), the latter centre also scoring below the *centre JIF* and the world impact – see also Figure 1. So, essentially three centres (a third of SMP) demonstrate a *strategic difference*, as measured by at least a 10 % difference over all other indicators by the centre CIFs. This difference may later be beneficial to the Danish environmental research, e.g., due to the training of upcoming younger researchers during the five-year program. Already the ensuing years, 1998-2000, demonstrated a definitive increase from previous years (Table 1) in Danish environmental publications: 315, 299, 321 articles per year.

Finally, Table 2 demonstrates that the overall centre impact for *all* articles, *CIF**, is quite high (14.6), i.e., only 2 citations per article below the mean SMP *CIF*. Whereas the *CIF** for the *Groundwater* centre is quite lower in value than the *CIF*, most other centres possess a *CIF** that is close to the corresponding *CIF* in value. One centre, the *Root zone*, displays a *CIF** (14.8) that is *higher* than its *CIF* (14.7). The reason for this rare phenomenon is that the 13 non-SCI articles all were cited heavily by SCI-journals and thus have appeared to be highly useful to the scientific community. Besides, the CIF*-CIF difference demonstrates the average impact of research information mediated to practitioners of the field in question and the society as such.

Comparing the Mid-term and Final Evaluations

Figure 1 demonstrates that at the mid-term assessment three centres obtained index scores *below the baseline*, 1.0, signifying that their *CIF* did not reach the corresponding *centre JIF*. At the mid-term the mean SMP *CIF* = 9.2, with a diachronic *JIF* = 7.8. At the final evaluation two such centres managed to pass the JIF index baseline (the *Groundwater* and *Air pollution* centres). The high-impact centre, *Freshwater*, managed to hold its index value but several centres dropped in values at the final evaluation: the *Marine*, *Root zone* and *Terrestrial ecology* centres.

The Use of Top-impact Journals

When comparing the two SMP evaluations, the most interesting results deal with the use of top-impact journals – Figures 2-3. At the mid-term study the 25 SCI-journal volumes with top-JIF scores were correlated with the citation scores obtained by the articles published in the same volumes. At the final investigation the top-40 SCI journal volumes were used. Further, the initial 151 articles published in SCI journals did not continue their robust impact correlation with their corresponding journal volumes. The coefficient dropped from .61 to .496 when the correlation window became extended from min. 4 to max. 10 years.

Figure 2 demonstrates that there are few outliers compared to the diagram, Figure 3. In the latter diagram some outliers hold very high article impact scores, as shown in relation to Table 3 above. Further, there exists a concentration of pairs situated between JIF-values of 20 and 35 citations and article impact values ranging from zero to 50 citations. The reason for the weak to bad correlation coefficients obtained at the final evaluation lies in the fact that in too many cases articles published in (the few) top-impact journal volumes received much less citations than the corresponding diachronic JIF. Besides, many high-impact articles were published in lower-impact journal volumes. By removing three outliers from the calculation, Figure 3, the improvement is marginal and still very weak (r = .34, p = .005; CV= .41).

Note however, that exactly the same article impact outliers in question seem responsible for the favorable but relative CIF/JIF difference obtained by, in particular, the smaller Biochemical Epidemics centre, displayed on Table 4.

Only the Freshwater centre displayed at the same time a robust correlation over the total time span for all as well as top-20 article-journal pairs, Table 3, did not possess significant outliers, Fig. 4, and demonstrated also a strategic difference. The Groundwater centre also showed a fairly good correlation score, although much smaller, had no outliers, but did not display a strategic difference.

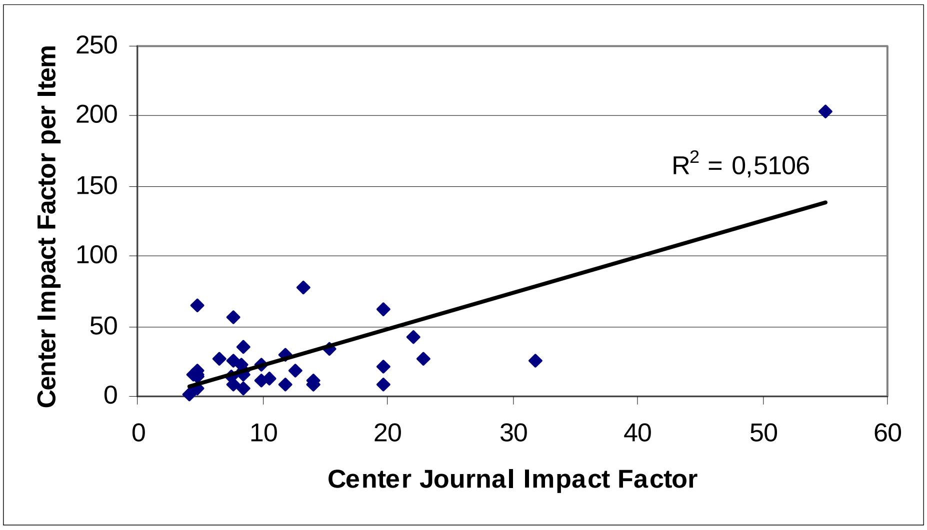

Figure 4. Correlation between the JIF of the 36 journal volumes and the citation impact of the corresponding articles published by the Freshwater Centre 1993-1997, cited 1993-2002.

The Pearson correlation coefficient is a relative measure in the sense that variation from a mean plays the central role – regardless the value of that mean. The extension to which a journal also holds centre articles with equal (or higher) citation impact scores can be observed at the detailed data level underlying the Table 2 values per centre: For all nine centres 175 articles (of a total of 344 SCI-items = 50 %) yield better impact than the corresponding *centre JIF*. This proportion of better-than-JIF cited articles is interesting compared to the common trend (Seglen, 1997, p. 498) that "articles in the most cited half of articles in a journal are cited 10 times as often as the least cited half". Centres with a percentage > 50 % typically belong to the three to four centres showing a *strategic difference* compared to the other indicators, Table 4. Again, this additional indicator specifies that SMP as program did not entirely make a difference – but some centres did.

Methodological Issues

As stated at the start of this discussion the dominant position of the strategic research program seems to influence the productivity of the remaining research community outside SMP in environmental research. At the mid-term evaluation one might have made use of all available data in order to gain additional information on the behaviour of the program. One might here think of the research work and articles underway or actually published after 1995 and into 1997/98, when the mid-term evaluation was carried out. We did not take that data into consideration. However, Table 1 on the productivity could (and should) have been made available.

Similarly, one might have made use of the Immediacy Index values, which is the only easily available citation indicator made by ISI of diachronic nature. Such values per applied journal could have been compared to the corresponding article's immediacy impact. This additional data and assessment information might hence have made the mid-term evaluation more comprehensive. The issue at stake is that mid-term evaluations have important implications for the continuation of research funding, its magnitude, and its direction – not only for a large-scale strategic research program like SMP, but for the remaining local research community. It is rarely so that strategic programs obtain large additional funding – it merely seizes (by a given political priority at a point in time) what commonly is *already in the system*. Then somebody else receives less.

The application of *weighted comparative indicators* of national and world properties – forming a 'shadow' of the unit under investigation – we found, like van Raan (1999), to be of central methodological importance. The present evaluation clearly showed that un-weighted comparisons would be unfair to the unit under analysis. A central role is that they be compared to the research program's average as well as single CIFs *and*, in particular, to the *centre JIFs*. The reason is that the latter performance indicator is *highly relative* and must be seen in context: We have observed how some units are able to produce article citation impact factors quite high above the average impact of the corresponding journals (the centre JIF) applied by the unit for publication. By comparing to the weighted research indicators of the same field(s) it becomes observable whether the journals used for publishing in a unit are 'light-weight' impact journals or more average or high-impact journals in that field. See for instance the cases in Table 4 and Figure 3. The interpretation of the evaluation outcome becomes definitively more reliable.

Concluding Remarks

The two-step research evaluation covers as a minimum a citation window of 7 years (1997-2002) with a 10-year widow as maximum (1993-2002). This ensures robustness in the citation analyses. Perhaps owing to the extended citation window at the final evaluation, the variation of citations received per SCI-article increased, and the promising substantial correlation observed at the mid-term evaluation between top-ranked journals and their corresponding centre articles did not continue. Too often a top-cited article from a centre was published in a journal displaying a much lower diachronic *JIF*; but the opposite phenomenon also took place. Perhaps this difference in correlations over time simply may signify that what seems manifest at a short-term impact perspective becomes more changeable or scattered in a long-term citation scenario. Of course, the mere values of the involved correlation vectors are also of importance: smaller impact values tend to correlate better in Pearson than when the impact increases in values, since the variation from the mean figures may increase.

As a program SMP did make a *success* associated with the volume of research publications published over the five-year period. The doubling and tripling of output is significant. But perhaps at the cost of the rest of the field's research development. It becomes hence of interest to follow up the Danish environmental research production from 1998 onwards. Already the years 1998-2000 look promising.

The *Danish environmental research* as such thus demonstrates a strategic difference in the Environmental sciences – probably helped substantially by the SMP initiative dominating the national field. From a wider perspective, in the cases of assessments of strategic research programs, one should always attempt carrying out a two-step evaluation package, which makes use of *all* available data at the time of investigation. Further, one might profit from observing what happens to the research communities outside the program – *its context* so to speak – in terms of publication growth, citation impact, and relationships to the program. This is owing the large influence such programs have for a substantial period of time on resource allocation, researcher affiliation and head hunting, and the volume of funding in particular fields.

The variation between the SMP centres in publication behaviour seemed quite large. Two centres – *Groundwater* and *Eco-toxicology* – published quite often articles in non-SCI journals without receiving enough citations to compensate. This behaviour resulted in *CIF*-scores below or just on the diachronic JIF baseline. The two centres probably contributed most to the fact that SMP, as research program, *did not* make a substantial *strategic difference*.

Notwithstanding 3-4 centres did make a strategic difference compared to the Danish (and world) citation impact in the relevant research areas. The four high-impact centres produced a *CIF* at least 10 % higher than any of the other indicators applied. One of the centres actually showed a *CIF** for all SCI and non-SCI articles that is *higher* than its *CIF* – the *Root zone* centre. Characteristically, the 3-4 high-impact centres also demonstrate a substantial proportion of SCI-papers that receive more citations than the mean *centre JIF* scores for each centre. In total 57 % satisfies this condition, which supports their strategic advantage. However, single citation impact outliers play a significant role in some centres, as concerns their impact behavior. This owes to the size of such centres.

Finally, one may observe that in terms of *knowledge export* from the SMP program, the surprising observation is that USA is the predominant knowledge importer followed, not surprisingly, by Denmark.

Acknowledgement

We wish to thank the Strategic Research Program for Environmental Research for sponsoring the research and the Nordic Research School in Library and Information Science (NORSLIS) for its travel support.

References

Aksnes, D.W. (2003). Characteristics of highly cited papers. *Research Evaluation*, 12(3), 159-170.

Christensen, F.H. & Ingwersen, P. (1996). Online citation analysis: a methodological approach. *Scientometrics*, 37(1), 39-62.

Egghe, L. & Rousseau, R. (1990). *Introduction to Informetrics*. Amsterdam: Elsevier.

Fisker, M. (2004). *Miljøviden i Verdensklasse? Læren af 90'ernes strategiske miljøforskning [World class environmental knowledge? Lessons learned from the strategic environmental research of the 1990es]*. København: Mandag Morgen. 71 p. (In Danish)

Ingwersen, P., Noyons, E. and *Larsen, B.* (2001): Mapping national research profiles in social science disciplines. *Journal of Documentation*, 57(6), 715-740.

Ingwersen, P. and *Larsen, B.* (2005): Evaluation of strategic research programs: The case of Danish environmental research 1993-2002. In: Ingwersen, P. and *Larsen, B.* eds. *Proceedings of ISSI 2005 – the 10th International Conference of the International Society for Scientometrics and Informetrics, Stockholm, Sweden, July 24-28, 2005, Volume 2*. Stockholm: Karolinska University Press, p. 450-459.

Ingwersen, P., Larsen, B. & Wormell, I. (2000). Applying diachronic citation analysis to research program evaluations. In: Cronin, B. & Atkins, H.B. (eds.), *The Web of Knowledge: Festschrift for Eugene Garfield*. (pp. 373-388). Medford, NJ: Information Today, 2000.

Seglen, P.O. Why the impact factor of journals should not be used for evaluating research. *British Medical Journal*, 314, 498-502.

Steele, T.W. & Stier, J.C. (2000). The impact of interdisciplinary research in the environmental sciences: The forestry case study. *Journal of American Society for Information Science*, 51(5), 476-484.

van Raan, A.F.J. (1999). Advanced bibliometric methods for the evaluation of universities. *Scientometrics*, 45(3), 417-423.

Wormell, I. (1998). Informetric analysis of the international impact of scientific journals: How 'international' are the international journals. *Journal of Documentation*, 54(5), 584-605.