Research Evaluation, 16(1), March 2007, pages 47–57
DOI: 10.3152/095820207X196777; http://www.ingentaconnect.com/content/beech/reev

Evaluation of strategic research programs: the case of Danish environmental research 1993–2002

Peter Ingwersen and Birger Larsen

The article reports on the mid-term and final scientometric evaluations of the Danish Strategic Environmental Research Program (SMP), which consisted of 13 virtual research centers from 1993 to 1997, nine of which are studied bibliometrically here. Citations are measured from 1993 to 2002. Central indicators are: center impact factor (CIF), the number of citations received by each center's Science Citation Index (SCI) articles; and center journal impact factor (JIF), which is a diachronous IF per journal volume publishing a center article. Citation and publication data are obtained from the Thomson-Dialog online version of SCI. Other indicators applied include the weighted Danish and world domain impact from the national science indicators in subject areas selected by the centers. Top-ranked journal volumes used in SMP in terms of JIF scores were correlated with the corresponding articles' citation values and vice versa. At the mid-term assessment the Pearson coefficient showed a strong correlation, which weakened at the final evaluation, except for a few high-impact centers.

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 centers from 1993 to 1997. The motivation behind this report is threefold. First, a mid-term evaluation carried out in 1999 and covering the publications 1993–1995 as well as citations 1993–1998, with citation windows ranging from six to four years, resulted in assumptions concerning the centers' publication strategies. When researchers published in top-impact journals, their articles also received many citations (Ingwersen *et al*, 2000; Ingwersen and Larsen, 2005). This might be turned into a fruitful publication strategy for future research in the area. The correlation between diachronous 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. Second, 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 (DKK700 million) of public funds over the five-year period (Fisker, 2004). Six hundred 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 centers, including

Peter Ingwersen – Email: pi@db.dk – and Birger Larsen – Email: blar@db.dk – are at the Department of Information Studies, Royal School of Library and Information Science, Birketinget 6, DK 2300 Copenhagen S, Denmark.

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 to the ISSI2005 Conference, Stockholm, where a shorter version of the present paper was presented.

 0958-2029/07/010047-11 US$08.00

humanistic ones. For this reason, only nine centers could be analyzed fairly for citations. They are, with objectives in parentheses ():

1. Air Pollution Processes and Models (atmosphere and air pollution)
2. Terrestrial Ecosystem Research (atmosphere and air pollution)
3. The Groundwater Group (groundwater; pesticides in groundwater)
4. Agricultural Biodiversity (soil surface)
5. Root Zone Processes (soil surface)
6. Freshwater Environmental Research (freshwater and marine areas)
7. Strategic Environmental Research in Marine Areas (freshwater and marine areas)
8. Danish Center for Eco-toxicological Research (environmentally hazardous substances in the aquatic and terrestrial ecosystems)
9. Center 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 and 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 centers. 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 center provided a list of research publications. For the mid-term evaluation the lists covered the period 1993–1995. Similarly, the centers provided a supplementary list for the final evaluation covering 1996–1997. The entries of the lists were all searched online in SCI to establish whether the journals in question were 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 centers and time periods. Thus 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, for the two evaluations) was retrieved online for each article, whether being originally indexed in SCI or not, *and* for each corresponding journal publication year applied per article. In the mid-term evaluation the citation windows ranged from six years (1993–1998) to four years (1995–1998). In the final evaluation the maximum citation window was ten years (1993–2002). In that way, we are able to calculate the diachronous journal impact factor (JIF) (Egghe and Rousseau, 1990) online (Christensen and Ingwersen, 1996) *for each time* a journal was used by a center, that is, for each article published. 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). This is because the diachronous JIF is designed to have the same citation windows as the article — therefore it can be compared directly to the real impact of the corresponding research articles. This method is similar to the one applied by Van Raan (1999) on offline data, but in the present analysis publicly available online data is used.

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

The Danish and the world citation impacts per center, based on the scientific fields for each center, constitute secondary indicators. They are calculated by the application of NSI, and are comparable to the CIF and center JIF by covering identical citation windows, that is, ranging from a ten-year window for the 1993 publications to a five-year window for the 1998 articles. Each center 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 centers were summed up. This means that the Danish and world citation impacts are calculated in terms of *weighted* subject profiles (Van Raan, 1999; Ingwersen *et al*, 2001). They mirror the distribution of NSI categories over SMP as a program, as defined by its centers. One may say that they act as a kind of ‘shadow’ SMP program. For instance, the category ‘Environment/ecology’ appears eight times and ‘Biochemistry and biophysics’ three times, and so on, in the final SMP profile. Since data was not available covering the entire period, NSI data covering 1991 to 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 calculated for the citation values obtained by the entire population of article–journal pairs at the mid-term evaluation (151 pairs) and at the final assessment (344 pairs). This means that the diachronous JIF for each volume used was paired to the number of citations received by the corresponding published article. Further, the 151 mid-term pairs were reassessed at the final evaluation separately for comparative reasons. The top-25 (mid-term) and top-40 journal volume–article pairs (final evaluation) were analyzed by means of the Pearson calculation. We applied Pearson although the distribution of citations over journals and articles is not a normal distribution — but skewed. In four selected high-impact centers we checked the Pearson results for all the article–journal impact pairs against the average correlation coefficient for the SMP research program.

Finally, also as secondary indicators we observed the patterns of *knowledge export* by means of listing, by frequency, both the citing countries, and the subject categories of the journals that cite a center. 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 journal–article impact pair correlation results, both with respect to the mid-term and the final evaluation of the program, as well as for high-impact centers 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 to 1994 and redoubles from 1994 to 1996. It triples from 1993 to 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 approximately 19% of all Danish environmental output, but covers approximately 34% in 1995/96, increasing to approximately 39% in 1997.

Table 1. SCI-indexed articles published by all centers combined in the SMP program and Danish environmental research

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%

Sources: the centers and NSI, 2001

The proportion of SCI articles to all articles from SMP as a program was quite stable over the period, but with some variation between the centers. For instance, the groundwater center published only 20 SCI articles out of a total of 35 (57%) while the freshwater center 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 centers. One of the reasons for the final result of the CIF versus center JIF impact scores per center (see Figure 1) probably derives from this dissimilarity in publication behaviour.

Citation impact analyses

The final evaluation per center 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 center impact (CIF*), for reasons of comparison.

One should note that the average SMP ‘Domain DK’ and ‘Domain world’ indicators — in their *unweighted 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 (which supports Van Raan’s [1999] proposal of applying weighted comparative indicators). This comparison would have been unfair towards SMP since the program profile’s many ‘citation-light’ 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, almost 96%, including self-citations during the entire period. This figure is higher than the Danish and world shares of cited articles, which are 94% and 88% respectively (NSI, 2001). At the

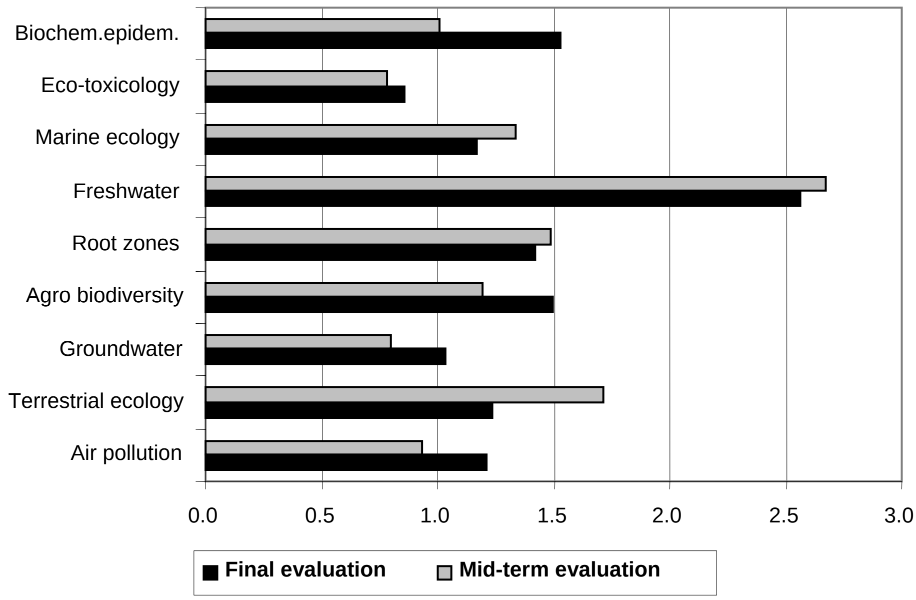

Figure 1. The nine SMP centers' citation impact (CIF) vs. the corresponding center JIF 1993–2002
Note: Index 1.0 = center JIF
Source: SCI (ISI, Thomson-Dialog online version)

mid-term evaluation this proportion for SMP as program was slightly lower, at 89%.

Knowledge export from the SMP centers

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

The following cluster of countries constitute the locations from which more than 10% of each center's citations are given, that is, the countries to which at least 10% of the knowledge export goes; with the number of times the country is represented across the centers in parentheses ():

1. USA (9) – most citing in five centers
2. Denmark (9) – most citing in 4 centers
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 center 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 center from NSI (i.e. Current Contents) as representative of their research. Checked in this

Table 2. Final evaluation of the nine SMP centers

Center	SCI Publ.	Citations SCI	CIF	Diachronous 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	1,009	***28.0***	*11.0*	20.0	19.9	42	1,062	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	1,034	10.6
Biochem. epidem.	34	864	***25.4***	*16.6*	23.1	22.5	38	881	23.2
Mean total – SMP	**344**	**5,741**	***16.7***	***12.8***	**16.6**	**15.2**	**434**	**6,324**	**14.6**

Note: Primary and secondary indicators, publications 1993–1997, cited 1993–2002
Sources: SCI, Thomson-Dialog online version; NSI (ISI)

One kind of knowledge export concerns which countries give most citations to the individual centers; five times out of nine USA is the most highly citing country, with Denmark as the second

way, it could be observed that, for instance, the freshwater and marine centers probably had selected areas (biology; biochemistry) somewhat too broad and/or out of tune with their real research foci and of too ambitious a nature, that is, with too high world (and Danish) citation impact. Another trend observed was the knowledge export into applied (engineering) fields from several centers.

The use of top-impact journals

In the *mid-term evaluation* we found a quite strong correlation for the 151 journal–article pairs between the diachronous JIF of the applied journal volumes and the citations given to the corresponding articles (Pearson's $r = 0.61$ with $r^2 = 0.37$; p = 0.005; Critical Value: CV = 0.25).

For the top-25 journal volumes' diachronous JIF and corresponding articles' citations, r equaled 0.68 with $r^2 = 0.47$ (p = 0.005; CV = 0.487) — see Figure 2. One might state that there exists a quite robust correlation between the *expected* average impact of articles in the journal set used for publication (the diachronous JIF) and the *actual* citation impact obtained per SMP article *over all articles* as well as over the top-25 pairs defined by the 25 high-impact journals.

Effectively the two mid-term correlation analyses imply that, first, in the overall analysis more than a third of the 151 articles published received a real impact corresponding to that of the journal used. Almost 60% did not. Second, for the top-25-impact journals the proportion increased to almost half of the corresponding 25 SMP articles obtaining a 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 ($r = 0.58$, with $r^2 = 0.33$ (p = 0.005; CV = 0.487). The correlation is still above the threshold (the CV) but, clearly, only a third of the high-impact articles are found in high-impact journals. Two thirds are found elsewhere, that is, in lower-impact journals; see also below on the publication behaviour of the centers.

The 151 SMP articles published from 1993 to 1995 were also analyzed for their journal–article pair impact correlation separately for the period 1999–2002, and in the total citation window 1993 to 2002. These analyses were done in order to observe if those initial articles, earlier on showing the robust correlation, actually did maintain that correlation. The results demonstrate that, first, for the late separate four-year citation window 1999–2002 the 151 articles published from 1993 to 1995 showed a correlation $r = 0.43$; $r^2 = 0.18$. Second, as to the total citation window 1993–2002, the results for the same 151 articles demonstrate that $r = 0.496$; and $r^2 = 0.25$ (for both analyses: p = 0.005 and CV = 0.25). This implies that, first, at the shorter later period the initially published 151 SMP articles did *not maintain* the strong impact correlation to the applied journal impact, as demonstrated for the initial period up to the mid-term evaluation, 1993–1998. It weakened substantially from $r^2 = 0.37$ to $r^2 = 0.18$. Second, at the extended citation window of a maximum of ten years only approximately 25% of the initial 151 SMP articles obtained a level of impact corresponding

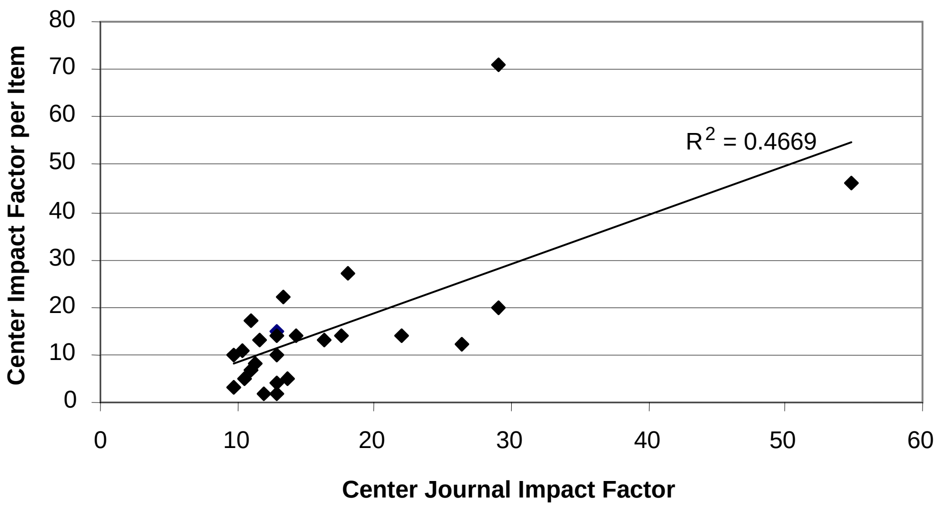

Figure 2. Correlation between the JIFs of the top-25 journal volumes and the citation impact of the corresponding articles published by SMP as program 1993–1995, cited 1993–1998

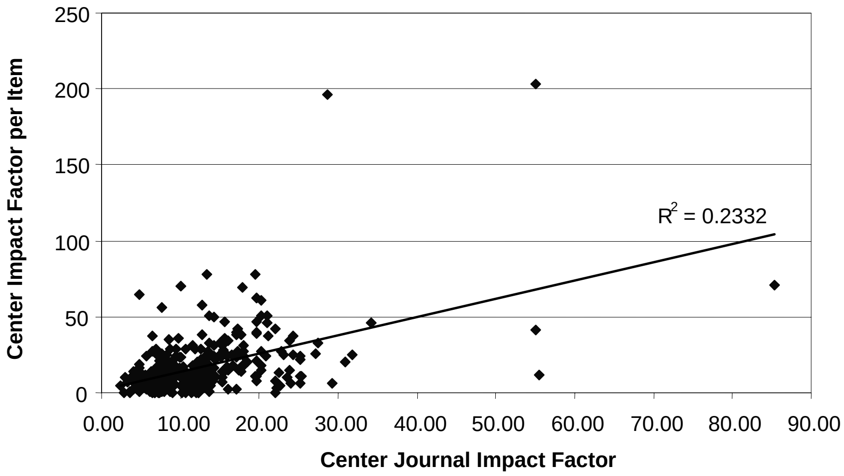

Figure 3. Correlation between the JIFs of all 344 journal volumes and the citation impact of the corresponding 344 articles published by SMP as program 1993–1997, cited 1993–2002

to that received by the journals used for publishing. Both correlation coefficients are still far above the CV value but have weakened substantially over the prolonged period.

These latter findings correspond to the result of the citation correlation analysis done for the entire program's 344 articles and corresponding journal volume pairs, covering the publication years 1993 to 1998 and the citation window 1993 to 2002 at the *final evaluation*. Here the maximum citation period is again ten years and the minimum window five years (1998–2002). Pearson's r = 0.482; r^2 = 0.23, with p = 0.005 — see Figure 3. The coefficient is statistically significant. We observe a number of outliers in Figure 3; see also below for selected centers. Approximately 75–80% of all the SMP articles obtain a level of citation impact *not corresponding* to that received by the journals used for publishing. The observed trend for the initial 151 SMP articles is continued for the total population of articles.

Also the top-impact journal–article impact correlation changed radically at the final evaluation. Two kinds of correlation analyses were carried out for SMP as such. One observed the top-40 journal volumes from the entire SMP program. Because the agro biodiversity center displayed quite low JIFs the center was not represented in the top-40 analysis. Hence, a second analysis took the upper 10% of the journal volume JIFs per center. In total 35 JIFs, and all nine centers, are represented.

In the first analysis of the top-40 journals Pearson's r equals 0.32 with r^2 = 0.102 (p = 0.005; CV = 0.41). This is an *extremely weak* correlation far below the critical value (CV). The second correlation coefficient covering all centers is similarly weak, r = 0.38, CV = 0.42. These scores are far from the mid-term evaluation results demonstrated above. Hence, what seemed a rather promising publication strategy at the mid-term assessment (to publish in high-impact journals and obtain a high citation impact in real terms) did not turn out so well at the end of the project period.

Consequently, in order to observe if at least some of the centers 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 centers displayed in Figure 1 and Table 3. We also analyzed the center for groundwater research for comparative reasons owing to its average impact (Table 4).

The analyses also revealed that if one *single outlier* was removed from the correlation analyses of the centers of terrestrial ecology (journal impact outlier), root zones and biochemical epidemiology (article impact outliers) the correlations changed, in some cases becoming statistically valid:

- Terrestrial ecology — r = 0.38; r^2 = 0.14; the *outlier*: JIF: 55.5 vs. cit: 12.0
- Root zones — r = **0.53**; r^2 = 0.27; the *outlier*: JIF: 10.03 vs. cit: 70.0
- Biochem. epidem. — r = **0.62**; r^2 = 0.38; Top-20 JIF/cit r = **0.59**; *outlier*: JIF: 28.67 vs. cit: 196

The four remaining centers did not show significant outliers.

In Table 4 we observe that only the high-impact freshwater center demonstrated a strong correlation coefficient (0.71) and a substantial r^2 value (0.51)

Table 3. SMP centers with equal or more than 10% difference between CIF and the other indicators

Center	CIF	Center 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

Sources: SCI, Thomson-Dialog online version; NSI, ISI (2001)

Table 4. Pearson correlation coefficients for all journal and article pairs of impact for five SMP centers

Center	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**

Notes: Critical values correspond to no. of analyzed pairs.
'cit' refers to citation impact obtained by the center articles (p = 0.005).
Figures in **bold** are statistically significant.

and also showed high coefficients for its top-20 journal–article impact pairs. Clearly it succeeded in publishing in high-impact journals *and* obtained a high impact. However, its high-impact journals' actual impact seems relatively lower than the average used in the field (see Table 3). The remaining three high-impact centers have had different strategies or may not have succeeded in the former publication strategy. In contrast, the groundwater center demonstrates a rather successful strategy: 40% of its articles obtain similar citation impact to the journal volumes in which the articles were published. This is also true for its top-20 correlation analyses.

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 1,127 SCI articles including the 344 SMP items during 1993 to 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 the 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, for example, improved research education. But such aspects do not form part of this evaluation.

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

The SMP productivity showed also high variation from center to center (Table 2). Three centers clearly produced below the average of 38 SCI articles while four centers were close to average. Similarly, the proportion of *non-SCI articles* varied across centers. The highest ratio was observed for the groundwater center (43% non-SCI articles), and the lowest for the centers of biochemical epidemiology (11%), marine ecology (13%), and freshwater (14%). Nonetheless, this does not mean that centers with a high ratio of non-SCI articles did not receive a high impact (see Tables 2 and 4). The groundwater center still kept its above-average CIF. The root zones center faired even better by means of the non-SCI articles (Table 2). Only the agro biodiversity center (36% non-SCI articles) demonstrated a weak CIF compared to the Danish field impact. In contrast, the eco-toxicology center 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 in society. Notwithstanding, this publication strategy may indeed produce substantial additional citation impact — see, for instance, the root zone center (Table 2), for which the CIF* is slightly *higher* than the CIF.

Citation impact comparisons

During the mid-term evaluation period 1993–1995, 151 SCI articles were produced that yielded 1,386 citations from 1993 to 1998. The same 151 articles received 1,486 citations from 1999 to 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 to 1997 provided the SMP program with 5,741 citations at a mean of CIF = 16.7 citations per SCI article compared to the diachronous center 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, that is, for the publication strategy at a center as expressed in the journals chosen for publication. It does not state anything about the optimal level of absolute impact. For this type of analysis it is very important to look into the balance between the center JIF and the Domain DK (and world) indicators (Table 2).

The central issue is that the mean center JIF is far below the Danish weighted citation impact (23% below); all center 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 worldwide (16% lower). In that sense SMP does not demonstrate any strategic difference. Generally speaking SMP received a much higher impact than deserved by the JIFs for the journals used by SMP, a fact also observed by means of the weak Pearson correlation coefficients as to journal–article pair-wise impact over the total of 344 SMP articles.

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 unweighted, as in most traditional and unfair bibliometric research evaluations, it would in fact have beaten the CIF score by 0.8 citations. However, the world impact would have decreased and been further beaten by the SMP CIF.

The reason here is that the heavy single weights of research fields such as biochemistry and biophysics, and microbiology, in the total SMP spectrum distort the picture if all fields are regarded as equal in the calculation of citation impact! Since they are *not* viewed as equal in our weighted method, a fair correspondence can be computed as 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 center must, in our opinion, show a rather higher CIF than any of the other indicators — for instance at least a 10% 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 overall Danish research. However, we did not explicitly investigate the remaining Danish non-SMP environmental research output. Nevertheless, four centers on the surface do show a robust average difference in relation to all other indicators (Table 3).

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

The latter two centers have applied far *lower-impact journals* than ordinarily are used worldwide 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 centers (freshwater and the biochemical epidemiology center) published in lower-than-average-impact journals, but with success — and thus received a much improved center impact (CIF). In three of the four centers the outliers served to improve the CIFs for each center by 1.3 (root zones), 5.5 (freshwater) and 5.8 (biochemical epidemiology) impact points over all center articles, that is, that the latter center 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 centers, air pollution and marine ecology, are close to have a 10% advantage over the other indicators, but not quite; and the groundwater center is in line with its center JIF. The two remaining centers are less advantageous in comparison to the competing indicators. They are directly below the Danish citation impact (agro diversity and ecotoxicology), the latter center also scoring below the center JIF and the world impact — see also Figure 1.

So, essentially three centers (a third of SMP) demonstrate a *strategic difference*, as measured by at least a 10% difference over all other indicators by the center CIFs. This difference may later be beneficial to the Danish environmental research, for example, 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, respectively.

Finally, Table 2 demonstrates that the overall center impact for *all* articles, CIF*, is quite high (14.6), that is, only two citations per article below the mean SMP CIF. Whereas the CIF* for the groundwater center is a lot lower in value than the CIF, most other centers possess a CIF* that is close to the corresponding CIF in value. One center, the root zone, displays a CIF* (14.8) that is slightly *higher* than its CIF (14.7). The reason for this rare phenomenon is that the 13 non-SCI articles were all 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 centers obtained index scores *below the baseline*, 1.0, signifying that their CIF did not reach the corresponding center JIF. At the mid-term the mean SMP CIF = 9.2, with a diachronous JIF = 7.8. At the final evaluation two such centers managed to surpass the JIF index baseline (the groundwater and air pollution centers). The high-impact center, freshwater, managed to hold its index value but several centers dropped in values at the final evaluation: the marine, root zone and terrestrial ecology centers.

The use of top-impact journals

When comparing the two SMP evaluations, the most interesting results deal with the impact correlations between the SMP articles and their corresponding journal volumes — Figures 2 and 3. The initial 151 articles published in SCI journals did not continue their robust impact correlation with their corresponding journal volumes. The coefficient dropped from 0.61 to 0.496 when the correlation window became extended from minimum four to maximum ten years. At the final assessment covering the entire period 1993–2002, the 344 journal–article volume impact pair correlation dropped further to r = 0.482. At the mid-term study the 25 SCI journal volumes with top-JIF scores correlated strongly with the citation scores obtained by the 25 articles published in the same volumes (r = 0.68). At the final investigation of the top-40 journal–article citation impact matrix this correlation turned out to be statistically insignificant.

Figure 3 demonstrates that there are two outliers with very high article impact scores, as discussed above in connection to the freshwater and biochemistry centers. Further, there is a dense concentration of pairs situated between JIF-values of two and 15 citations, and article impact values ranging from zero to 30 citations. The reasons for the weaker correlation coefficients obtained at the final evaluation lie in the fact that in too many cases articles published in (the less frequent) top-impact journal volumes received many fewer citations than the corresponding diachronous JIF. Besides, many higher-impact articles were published in lower-impact journal volumes.

Only the freshwater center displayed at the same time a robust correlation over the total timespan for all as well as top-20 article–journal pairs, (Table 4), and possessed a significant outlier, which did not reduce its strategic difference if removed. The groundwater center also showed a fairly good correlation score, although much smaller; it had no outliers, but did not display a strategic difference.

The Pearson correlation coefficient is a relative measure in the sense that variation from the mean plays a central role — regardless of the value of that mean. The extent to which a journal also holds center articles with equal (or higher) citation impact scores can be observed at the detailed data level underlying the Table 2 values per center. For all nine centers 175 articles (from a total of 344 SCI items = 50%) yield better impact than the corresponding center JIF — also demonstrated by the points above the trend-line, Figure 3. This proportion of better-than-JIF cited articles is interesting compared to the common trend (Seglen, 1997: 498) that “articles in the most cited half of articles in a journal are cited 10 times as often as the least cited half”. Centers with a percentage greater than 50% typically belong to the three to four centers showing a strategic difference compared to the other indicators, Table 3. Again, this additional Pearson-based indicator specifies that SMP as program did not entirely make a difference — but some centers did.

Methodological issues

One might have made use of the immediacy index values, which is the only easily available citation indicator made by ISI of diachronous 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 been applied at the mid-term evaluation on the current articles–journal volume pairs and made that evaluation more comprehensive. A later final assessment could then falsify the immediacy results. 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 the case 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 unweighted 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 center 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 center JIF) applied by the unit for publication. By comparison to the weighted research indicators of the same field(s) it becomes observable whether the journals used for publishing in a unit are 'lightweight' impact journals or more average or high-impact journals in that field. See for instance the cases in Table 3 and Figure 3. The interpretation of the evaluation outcome becomes definitively more reliable.

Concluding remarks

The two-step research evaluation covered as a minimum a citation window of four years (1995–1998) with a ten-year window 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 center articles did not continue. Too often a top-cited article from a center was published in a journal displaying a much lower diachronous JIF; but the opposite phenomenon also took place. Perhaps this difference in correlations over time simply signifys 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 was 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. Hence it becomes of interest to follow up the Danish environmental research production from 1998 onwards. In 2002–2005 Denmark produced 365 SCI articles per year on average, with a Crown indicator of 1.29, that is, showing a definitive strategic edge (NSI, 2006).

Danish environmental research as such thus demonstrated 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 to carry out a two-step evaluation package, which makes use of all available data at the time of investigation. Further, one might profit from observing in more detail 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 to 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 centers in publication behaviour seemed quite large. Two centers — groundwater and eco-toxicology — quite often published articles in non-SCI journals without receiving enough citations to compensate. This behaviour resulted in CIF scores below or just on the diachronous JIF baseline. The two centers probably contributed most to the fact that SMP, as research program, did not make a substantial strategic difference.

Nevertheless, three centers did make a strategic difference compared to the Danish (and world) citation impact in the relevant research areas. One of the centers actually showed a CIF* for all SCI and non-SCI articles that was slightly higher than its CIF — the Root zone center. Characteristically, the three or four high-impact centers also demonstrated a substantial proportion of SCI-papers that receive more citations than the mean center JIF scores for each center, a condition that supports their strategic advantage.

However, single citation impact outliers play a significant role in at least one high-impact and other lower-impact centers. This is because of the size of such centers.

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

References

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

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

Egghe, L and R Rousseau 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 1990]* (in Danish). København: Mandag Morgen.

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

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

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

National Science Indicators (NSI), 2001/2006. *NSI Database*. Produced by Institute for Scientific Information, Thomson Scientific, USA.

Seglen, P O 1997. Why the impact factor of journals should not be used for evaluating research. *British Medical Journal*, **314**, 498–502.

Steele, T W and J C Stier 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), 417–23.

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.