

Application of user statistics and additional data for collection management of Wageningen UR digital library



Marco G.P. Van Veller, Ger Spikman

Wageningen UR Library, the Netherlands

marco.vanveller@wur.nl

ger.spikman@wur.nl, <http://nl.linkedin.com/in/gerspikman>

Abstract

Good collection management is necessary to guarantee the quality, quantity and costs of the online journal collection of Wageningen University and Research Centre (Wageningen UR) Digital Library. *Via* quantitative research on both subscription prices and usage statistics (*i.e.* download data from vendors of online journals and clickthrough data statistics from the library's SFX link resolver) the library obtains data that are helpful to decision making regarding cost-effectiveness of online journals. By calculation of the price per article download for each online journal and holding this against the costs of document delivery from outside the library, a selection of online journals with a low cost-effectiveness is possible. These online journals can further be examined for cancellation. *Via* optimization to cost-effective online journals only, Wageningen UR Library strives to give access to scientific literature that is important ('need to have') to the users and researchers of Wageningen UR while keeping the costs of the subscriptions into the financial budget.

Keywords

Digital library, Collection management, User statistics, Journal subscriptions, Prices

Résumé

Application des statistiques utilisateur et autres données pour la gestion de la collection de la bibliothèque numérique Wageningen

Une bonne gestion des collections est nécessaire pour garantir la qualité, le volume et le coût des collections de revues électroniques de la bibliothèque de l'université de Wageningen. En associant, dans une méthodologie quantitative, les coûts des abonnements et les statistiques d'utilisation (données de téléchargement en ligne fournies par les éditeurs et celles du résolveur de liens SFX de la bibliothèque), la bibliothèque dispose de données et d'indicateurs sur une analyse de type coût-efficacité des revues électroniques. En calculant le prix du téléchargement par article de chaque revue en ligne et en le comparant au coût de sa fourniture ponctuelle (FDD, PPV), il devient possible d'identifier les titres ayant un faible rapport coût-efficacité, et par la même occasion de s'y désabonner. Par cette méthode, la BU de l'Université de Wageningen s'efforce de donner accès à une large collection de revues, indispensables aux usagers, tout en tenant compte des budgets alloués.

Mots-clés

Bibliothèque électronique, collections, statistiques, abonnements, revues, coûts-efficacité, bibliothèque

Marco VAN VELLER est bibliothécaire spécialisé en ressources électroniques à la bibliothèque universitaire de Wageningen, aux Pays-Bas.

Ger SPIKMAN est bibliothécaire à l'Université de Wageningen, aux Pays-Bas. Il est actuellement directeur du département des acquisitions et de la gestion des ressources documentaires.

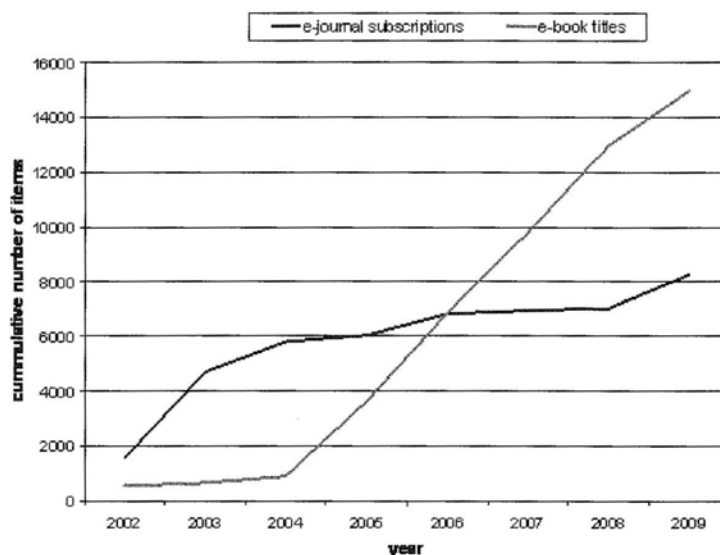
Wageningen University and Research Centre (or Wageningen UR) is an education and research organization in the Netherlands that includes a university, polytechnic college and several research institutes. The main topics for the educational programs and research activities at this organization all deal with “the quality of life” and are concentrated around the following three themes: food and food production, living environment and health, lifestyle and livelihood.

For almost 9500 BSc and MSc students and more than 7000 staff positions, Wageningen UR Library gives access to both digital and hardcopy versions of scientific literature, especially in the fields of Agrotechnology, Food Sciences, Plant and Animal Sciences, Soil Science, Geo-information, Landscape and Spatial Planning, Water and Climate, Ecosystem Studies and Social Sciences. The services of the library are not restricted to staff and students. All locations of Wageningen UR Library are also open to the general public. Wageningen UR Library has one main location at the university campus and a few satellite libraries at research institutes and experimental stations that are distributed over the Netherlands. At its main location in Wageningen, the library holds approximately 38000 publications in its (publically accessible) stacks and approximately 600000 publications in its (publically closed) depot.

Besides physical libraries with collections of (*e.g.*) books, journals and reports, Wageningen UR Library also holds a digital library with access to online resources that are possessed by the library or to which the library subscribes. Online resources that can be accessed *via* the digital library can be (*a.o.*) research reports, MSc or PhD theses, websites, databases (mostly bibliographies), electronic books or online journals. Especially, the last five years the number of electronic books that are available *via* the digital library grew substantially. For subscriptions to online journals, the growth started earlier at the end of the 20th and beginning of the 21st century. Currently, Wageningen UR Library holds around 15000 electronic books and subscribes to more than 8000 online journals (see figure 1).

Whereas expenses for electronic books are (in most cases) made only once, for periodicals (like online journals) expenses are made every year for subscriptions. In order to spend the library's budget on journals most efficiently, good collection management on online journals is essential. This paper discusses how Wageningen UR Library monitors prices to journal subscriptions and usage in order to select expensive journals with low usage that are candidates for cancellation. The monitoring information is used for decisions in collection management and optimization of the collection towards cost-effective online journals.

FIGURE 1 – CUMULATIVE NUMBER OF E-BOOKS AND ONLINE JOURNAL SUBSCRIPTIONS* AT WAGENINGEN UR LIBRARY FOR THE LAST EIGHT YEARS



**For the online journals, subscriptions are only taken into account when the starting year of the subscription to the journal by Wageningen UR Library is known.*

◆ **1. Online journal subscriptions at Wageningen UR Library**

Currently, Wageningen UR Library has 9003 online journal subscriptions. Generally, for these online journals, two categories of subscriptions can be distinguished: journal subscriptions in package deals or separate journal subscriptions.

Wageningen UR Library subscribes to a total of 8147 journals in 37 packages with various vendors. For the majority of these journals, subscription is arranged in 'big deals': 7892 journals in 20 packages. The contents and price of these 'big deal' packages are negotiated between the vendors and a Dutch consortium of 13 university libraries (of which Wageningen UR Library is one partner) and the National Library of the Netherlands. Besides the journals in the 'big deal' packages an additional number of 255 journals come in 17 other packages for which Wageningen UR Library negotiates directly with

the vendors. During the negotiations with a vendor, both the number of journals in a package (which varies between 2 and more than 1 500 for the current journal packages to which Wageningen UR Library subscribes) and price for the package is settled for a certain number of years. The result is that for most vendors the contents of the packages is fixed and cannot be modified on a yearly basis; for additional journals separate subscriptions have to be started. Only for a few packages modifications in journal contents is possible.

Besides subscriptions to journals in packages, Wageningen UR Library holds 856 subscriptions to separate journals. For each of these journal subscriptions, the price is negotiated with its vendor (which frequently is a publisher or scientific society) separately. For the separate journals the library holds subscriptions with 391 different vendors. For 72% of these vendors Wageningen UR Library subscribes to only a single journal that the vendor gives access to. Subscriptions to separate journals can often be started or cancelled on a yearly basis which makes them suitable for monitoring and deselection of expensive journals for cost-effective management of online journal subscriptions. Although 'big deals' will only be subscribed to if the contents of the package is of real importance to the library, it is obvious that the expenses made for them put the budget for the rest of the online journal collection under pressure.

◆ 2. Estimation of current prices for journals

Table 1 shows the share in expenses that are made in 2009 on online journal subscriptions in packages as well as for separate journal subscriptions. From the table it follows that package deals dominate in the subscriptions to online journals for Wageningen UR Library. Together, the packages contain 90% of the journal subscriptions and share also 90% of the total costs made on online journals in 2009. The average paid price per journal subscription for separate journals is much higher than for journals that are part of a

TABLE 1 – EXPENSES* IN 2009 ON AND NUMBER OF SUBSCRIPTIONS IN PACKAGES AND SEPARATE JOURNALS

	Share in total costs for online journals	Number of journal subscriptions	Average paid price per subscription
Packages	90%	8 147	€ 212,-
Separate journals	10%	856	€ 491,-
Total		9 003	€ 239-

**Calculations based upon invoices paid to vendors between January and September 2009.*

package deal. For journals that are part of a package deal the average paid price is calculated by dividing the total package price by the number of journals that are included in the package. This average price per journal is much lower than the price that would have to be paid if the library subscribes to the journal separately.

As discussed above, however, only for journal subscriptions that are not part of a package deal between Wageningen UR and a vendor (*i.e.* the “separate journals”) cancellation of most expensive subscriptions with low usage is possible. It is for this reason that a ranking needs to be made based upon prices and usage for these journal subscriptions. Hereby, it is important to use current prices per journal subscription.

For the registration of journal subscriptions, Wageningen UR Library uses a database, named Cardex, which is part of our library content management system. This database contains per journal subscription information on the journal title, print and online ISSN, estimation of the price (based upon indexed prices from last year(s)), paid invoices per year, vendor, etc. From Cardex a selection can be made with estimated (indexed) prices per journal subscription and paid invoices over the current year. Paid invoices over the current year are representative for current prices per journal subscription. Paid invoices over the current year, however, may contain corrections with payments of earlier years. Moreover, the amount of expenses on paid invoices are not complete when current prices per journal subscription need to be known before the end of the year. It is for these reasons that the paid invoices over the current year need to be corrected.

Figure 2 shows the paid invoices, over 2009, for separate journal subscriptions against the estimated prices for these journal subscriptions in September 2009. The red line in this figure represents the relation between the paid and estimated price.

From a regression analysis on the data in figure 2 the following relationship between paid and estimated prices can be derived:

$$\hat{P}_{i,j,2009} = (0.9812 * P_{i,j}) + 35.151$$

With

$\hat{P}_{i,j,2009}$ = Predicted price for subscription *i* of journal *j* in 2009

$P_{i,j}$ = Estimated price for subscription *i* of journal *j*, based upon an indexation of mean prices over earlier years

The price per journal subscription in 2009 is predicted with this relation from the estimated price that is listed in Cardex for each journal subscription separately. In order to obtain current prices per journal, the predicted prices for all subscriptions need to be summed:

$$P_{j,2009} = \sum_{i=1}^n \hat{P}_{i,j,2009}$$

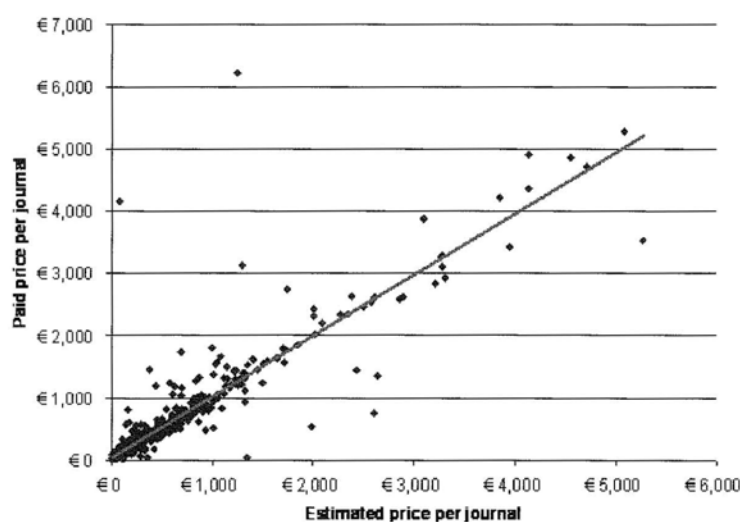
With

$P_{j,2009}$ = Current price for journal j

n = number of subscriptions i to journal j

Sometimes, one (or more) of the subscriptions of a separate journal is also part of a package deal. In this case the price per journal for the package (obtained by dividing the total paid price for the package deal by the number of journals in this package) is included in the summation to obtain the total price for all subscriptions of the journal.

FIGURE 2 – PAID PRICES AGAINST ESTIMATED (INDEXED) PRICES FOR ONLINE JOURNAL SUBSCRIPTIONS. ONLY PRICES FOR SEPARATE JOURNAL SUBSCRIPTIONS THAT DO NOT COME IN A PACKAGE ARE CONSIDERED



◆ **3. Journal usage data**

Access to online journals to which Wageningen UR Library subscribes is possible for licensed users *via* the digital library (see introduction). In general, the users can obtain articles, published in online journals, from the vendors website *via* four different pathways:

- *via* the catalogue of Wageningen UR Library,
- *via* a (bibliographical) database,
- *via* internet (e.g. Google Scholar), or
- directly *via* the journal website.

When consulting the library catalogue, a (bibliographical) database or using Google Scholar for access to articles, the user gets access to the online journal by clicking on a Get it! button that appears for every record in the catalogue or journal to which Wageningen UR Library subscribes.

After clicking on the button, a SFX-menu is shown in which the user can fill in the volume, issue and start page of the article that is needed.

After further clicking on a Go button in the SFX-menu the user is directed to the vendors (or publishers) website where the requested article can be downloaded if the user is licensed (*i.e.* student or staff member of Wageningen UR).

The clicks in the SFX-menu on the Go button for obtaining an article are counted per journal as “clickthroughs” and give an indication of the use of online journals *via* the Wageningen UR Library.

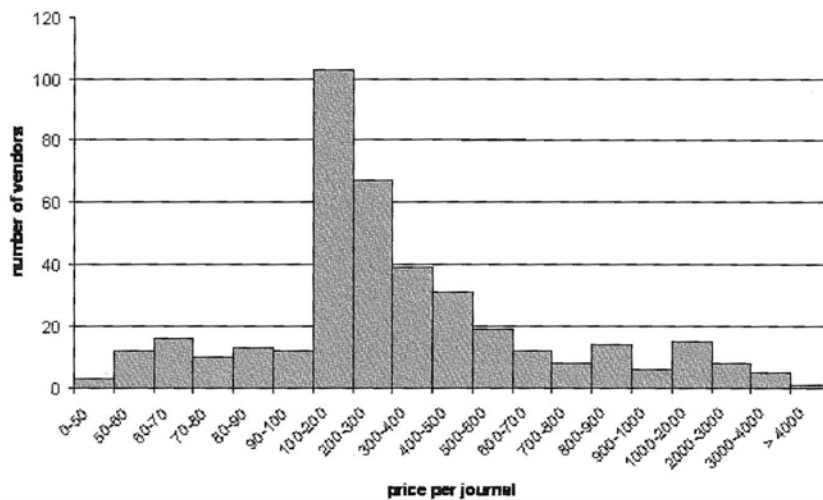
Articles in online journals, however, can also be obtained within the network of Wageningen UR by directly going to the website of the journal. The latter is done when users know where to find the online journal on the internet or when they received alerts on new articles in the journal with direct links to the journal website. There may be also other cases of direct access to articles in online journals. In all these cases of access to articles by directly going the website of the journal the SFX-menu is not shown and no clickthroughs are counted.

Independent of the used path for access to the online journals (*via* the catalogue, a (bibliographical) database or directly the journal website), articles need to be downloaded from the vendors (or publishers) website. Often, these article downloads are counted per journal for all licensed users from Wageningen UR and can be obtained in reports from vendors that make them available. These counts of article downloads give the most precise and complete indication of online journal use.

◆ 4. Selection of expensive journals with low usage

Figure 3 gives the distribution of the average price per journal for vendors of separate journals to which Wageningen UR Library subscribes. For the selection of journals that are not cost-effective (*i.e.* expensive journals with low usage) one can start with the journal subscriptions from vendors with the highest average price per journal and collect usage data for these journals to order to assess their cost-effectiveness.

FIGURE 3 – DISTRIBUTION OF PRICE PER JOURNAL PER VENDOR FOR 391 VENDORS WITH IN TOTAL 856 JOURNALS TO WHICH WAGENINGEN UR LIBRARY SUBSCRIBES SEPARATELY



The 50 vendors with the highest average price per journal give access to 125 journals together. For half of these journals counts of article downloads can be obtained from the vendors. For the other half of the journals a further estimation on usage needs to be made since no download data are available.

Figure 4 shows the number of downloads against the number of clickthroughs for 62 journals from vendors with the highest average price per journal. The red line in this figure represents the relation between the number of downloads and clickthroughs for these journals.

From a regression analysis on the data in figure 4 the following relationship between the number of downloads and number of clickthroughs in 2009 can be derived:

$$\hat{D}_{j,2009} = 7.2273 * C_{j,2009}$$

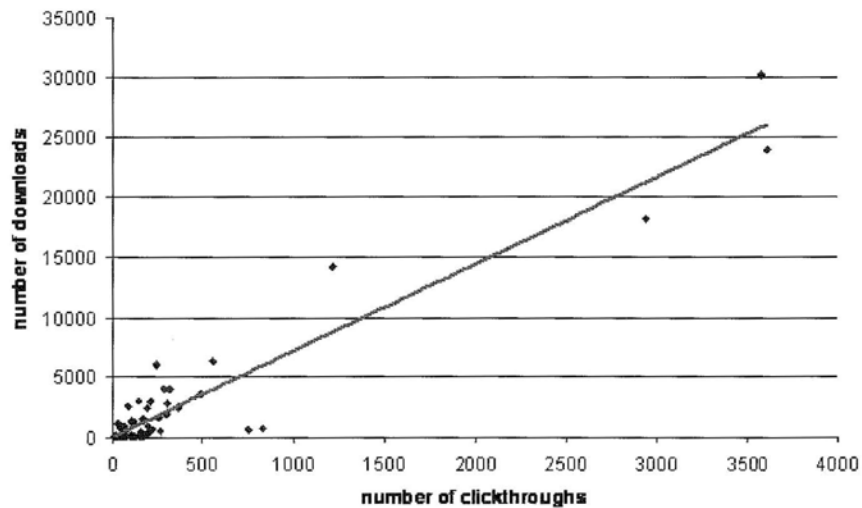
With

$\hat{D}_{j,2009}$ = Predicted number of downloads for journal j

$C_{j,2009}$ = Counted number of clickthroughs in SFX-menu for journal j

For the journals for which the vendors do not provide any article download data the number of downloads in 2009 is predicted *via* this relationship.

FIGURE 4 – NUMBER OF DOWNLOADS AGAINST NUMBER OF CLICKTHROUGHS FOR ONLINE JOURNAL SUBSCRIPTIONS FROM VENDORS WITH THE HIGHEST AVERAGE PRICE PER JOURNAL. ONLY SEPARATE JOURNAL SUBSCRIPTIONS THAT ARE NOT PART OF A PACKAGE DEAL ARE CONSIDERED



With the (predicted) current price and the (counted or predicted) number of article downloads a measure for the cost-effectiveness of an online journal is the price per article download. By calculation of this measure and listing the journals starting with highest price per article download, the least cost-effective online journals can be selected. In case the number of article downloads is zero, cancellation of the online journal subscriptions needs to be considered since there is no indication of any use on these journals. For the online journals with a high price per article download one may also consider cancellation if delivery of articles from other libraries is less expensive than the price per download which is measured.

For the 125 online journals from the 50 most expensive vendors, 10 journals did not have any article downloads in 2009. Further, another 22 online journals have a price higher than € 6.50, which is the price for article delivery from Wageningen UR Library to another library. These 32 journals, that show the lowest cost-effectiveness, will further be examined for cancellation.

◆ **5. Concluding remarks**

The methodology described above can be used to get an impression on the cost-effectiveness of online journals based upon currently paid prices for their subscription and current usage data. By calculation of the price per download for each journal a decision measure is provided for selection of low cost-effective journals. Additional examination of these journals before cancellation of their subscription, however, is very important since there may be additional reasons to have these journals in the collection of Wageningen UR Library.

A variant of the above described methodology is used for judging ('big deal') packages. In this judgment, the total usage of each package is held against the paid price of the package. The price per article download for the journals in the package is calculated to assess cost-effectiveness.

In addition to the price per article download, also the total expenses for subscribing to the 'need to have' journals from a package is calculated and held against the paid price for the package to control whether the package still is cost-effective for the library. Hereby it is important to make sure that long term commitments to vendors of 'big deal' packages will not be a threat to the online journal collection as a whole.