Time-dependent Lotkaian informetrics incorporating growth of sources and items

L. Egghe *

Univeristiteit Hasselt (UHasselt), Campus Diepenbeek, Agoralaan, B-3590 Diepenbeek, Belgium
Universiteit Antwerpen (UA), Campus Drie Eiken, Universiteitsplein 1, B-2610 Wilrijk, Belgium

ARTICLE INFO
Article history:
Received 13 February 2007
Received in revised form 8 January 2008
Accepted 11 January 2008

Keywords:
Lotka
Lotkaian informetrics
Growth of sources
Growth of items
Time-dependent
h-index
Hirsch

g-index

ABSTRACT
In a previous article, static Lotkaian theory was extended by introducing a growth function for the items. In this article, a second general growth function – this time for the sources – is introduced. Hence this theory now comprises real growth situations, where items and sources grow, starting from zero, and at possibly different paces. The time-dependent size- and rank-frequency functions are determined and, based on this, we calculate the general, time-dependent, expressions for the $h$- and $g$-index. As in the previous article we can prove that both indices increase concavely with a horizontal asymptote, but the proof is more complicated: we need the result that the generalized geometric average of concavely increasing functions is concavely increasing.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction
Let us have the size-frequency function of Lotka type

$$f(j) = \frac{C}{j^\alpha}$$

with $j \in [1, +\infty]$ and $\alpha > 1$. The theory could be extended to general $\alpha$ provided $j$ is limited to a finite closed interval but this complication would make the arguments in this paper more cumbersome while, essentially, nothing new would be added; therefore we limit ourselves to $j \geq 1$, $\alpha > 1$. Hence, for every $j \geq 1$,

$$\int_j^\infty f(j') \, dj' = r$$

denotes the total number of sources with item-density $j$ or larger. Transforming (1) into a probability density $\varphi$ leads to

$$\varphi(j) = \frac{\alpha - 1}{j^\alpha}$$

as is readily seen (since $\alpha > 1$).

In [5], we let the items grow in time $t$, using a general growth function $G(t)$ as follows: define the new item-densities at time $t$ as

$$jG(t) = k.$$
Based on this we could prove in [5] that the time-dependent size-frequency density, derived from \( \varphi \) is the function

\[
\varphi (k, t) = G (t)^{\alpha - 1} \varphi (k)
\]

(5)

for \( k \geq G (t) \). Here \( \varphi (k) \) is considered to be \( \varphi (k, \infty) \), i.e. for \( t = \infty \).

The rank-frequency function \( \gamma \), associated with \( f \), is generally defined as (cf. [4])

\[
r = \gamma^{-1} (j) = \int_{j}^{\infty} f (j') \, dj'
\]

(6)

(cf. (2)), where \( \gamma^{-1} \) denotes the inverse of the function \( \gamma \) (note that \( \gamma^{-1} \) strictly decreases so that \( \gamma = (\gamma^{-1})^{-1} \) is defined). Note also that \( r \geq 0 \).

The time-dependent rank-frequency function was proved in [5] to be

\[
\gamma (r, t) = G (t) \gamma (r).
\]

(7)

The Hirsch-index (or \( h \)-index, shortly) is defined on a ranked list of sources (ranked in decreasing order of the number of items they produce or have), cf. [1–3,6–14]. It is the highest rank \( h \) such that all sources on rank \( 1, \ldots, h \) have \( h \) or more items. In the above time dependent framework it was proved that (see [5]— see also [8]) the time-dependent \( h \)-index is

\[
h (t) = G (t)^{\frac{\alpha - 1}{\alpha}} \cdot h,
\]

(8)

where \( h \) is the \( h \)-index for \( t = \infty \).

The \( g \)-index (introduced in [6,7,9]) is a solution for the drawback of the \( h \)-index that it does not increase even when the sources with the most items increase their number of items. We define the \( g \)-index to be the highest rank \( g \) such that the sources on rank \( 1, \ldots, g \), together, have \( g^2 \) or more items (note that \( h \) satisfies this, by definition, hence \( g \geq h \)). In the above framework, in [5], we proved that the time-dependent \( g \)-index is

\[
g (t) = G (t)^{\frac{\alpha - 1}{\alpha}} \cdot g,
\]

(9)

where \( g \) is the \( g \)-index for \( t = \infty \).

In the next section we will extend this theory by allowing for two general growth functions: one is the function \( G \) as in (4) (from now on called \( G_1 \)) and another function \( G_2 \) acting on source ranks. We then determine the general time-dependent Lotkaian theory by proving formulae for the time-dependent size- and rank-frequency functions (using both growth functions \( G_1 \) and \( G_2 \)).

Formula (5) is extended to

\[
\varphi (k, t) = G_2 (t) G_1 (t)^{\alpha - 1} \varphi (k)
\]

(10)

and formula (7) is extended to

\[
\gamma (s, t) = G_1 (t) \gamma (r)
\]

(11)

with

\[
r G_2 (t) = s
\]

(12)

(the defining relation for \( G_2 \)).

In the third section we show that the general time-dependent formulae for the \( h \)-index and the \( g \)-index are:

\[
h (t) = G_1 (t)^{1 - \frac{1}{\alpha}} G_2 (t)^{\frac{1}{2}} \cdot h
\]

(13)

\[
g (t) = G_1 (t)^{1 - \frac{1}{\alpha}} G_2 (t)^{\frac{1}{2}} \cdot g
\]

(14)

where \( h \) and \( g \) are the \( h \)-index and \( g \)-index for \( t = \infty \). Hence both \( h (t) \) and \( g (t) \) are the values at \( t = \infty \), multiplied by a time factor which is a generalized geometric mean of the functions \( G_1 (t) \) and \( G_2 (t) \). We prove that such a generalized geometric mean of concavely increasing functions \( G_1 \) and \( G_2 \) is concavely increasing, hence the same is true for the functions \( h (t) \) and \( g (t) \), growing to their values \( h \) and \( g \) (formulae proved earlier will be repeated here) for \( t = \infty \).

2. Time-dependent Lotkaian theory, incorporating general growth models for items and sources

2.1. Basic relations

Basic in this theory is the “static” law of Lotka (1), normalized so that we have a size-frequency distribution (see (3))

\[
\varphi (j) = \frac{\alpha - 1}{j^\alpha},
\]

(15)
where \( j \geq 1 \) and \( \alpha > 1 \). This distribution is considered as the “limiting” situation at time \( t = \infty \), where \( \varphi (j) \) denotes the density of sources with item density \( j \) (see [4]).

For the growth in time \( t \geq 0 \) we assume the following simple relations:

(i) Item growth: Let \( G_1 (t) \) be the cumulative growth distribution of the item densities expressed by (cf. (4) with \( G = G_1 \))

\[
jG_1 (t) = k.
\]

Hence, since \( j \geq 1 \), we have that \( k \geq G_1 (t) \).

(ii) Source growth: Let \( G_2 (t) \) be the cumulative growth distribution of the rank densities (of the sources) expressed by (new equation)

\[
rG_2 (t) = s.
\]

Hence \( s \geq 0 \) since \( r \geq 0 \).

We underline that \( G_1 \) and \( G_2 \) are arbitrary cumulative growth distributions and, of course \( G_1 \) and \( G_2 \) can be different. As cumulative growth distributions we can (and will) suppose that \( G_1 \) and \( G_2 \) are strictly increasing concave functions such that \( G_1 (0) = 0 \), \( G_2 (0) = 0 \), \( \lim_{t \to \infty} G_1 (t) = 1 \), \( \lim_{t \to \infty} G_2 (t) = 1 \). A classical example of such a function is

\[
G_i (t) = 1 - a_i^t,
\]

where \( 0 < a_i < 1 \), \( t \geq 0 \), but this is only an example (we do not need (18) for our model further on).

2.2. Size-frequency functions, in function of time \( t \)

Let \( T \) denote the total number of sources at \( t = \infty \). Since (17) implies

\[
\frac{r}{T} G_2 (t) = s \frac{T}{T} \tag{19}
\]

and since

\[
\frac{r}{T} = \int_1^\infty \varphi (j') \, dj'
\]

denotes the fraction of sources (at \( t = \infty \)) with item-density \( j \) or more (cf. (2) divided by \( T \), hence using (3)), we have, by (19) and (20) that

\[
G_2 (t) \int_j^\infty \varphi (j') \, dj'
\]

denotes the fraction of sources (with respect to \( t = \infty \)) with \( k \) or more as item-density, where

\[
jG_1 (t) = k
\]

by (16). Let \( \phi (k, t) \) denote this fraction of sources. Hence

\[
\phi (k, t) = G_2 (t) \int_{\frac{k}{G_1 (t)}}^\infty \varphi (j') \, dj' = G_2 (t) \int_{\frac{k}{G_1 (t)}}^\infty \frac{\alpha - 1}{j^{\alpha}} \, dj' = G_2 (t) \left( \frac{G_1 (t)}{k} \right)^{\alpha - 1}
\]

since \( \alpha > 1 \). Since

\[
\phi (k, t) = \int_k^\infty \varphi (k', t) \, dk'
\]

by definition of the time-dependent Lotka function \( \varphi (k, t) \), we clearly have

\[
\varphi (k, t) = - \varphi' (k, t), \tag{25}
\]

where \( \varphi' (k, t) \) denotes \( \frac{\partial \varphi (k, t)}{\partial k} \). Consequently, by (23)

\[
\varphi (k, t) = G_2 (t) \left( \alpha - 1 \right) \frac{G_1 (t)^{\alpha - 1}}{k^\alpha} \tag{26}
\]

\[
\varphi (k, t) = G_2 (t) G_1 (t)^{\alpha - 1} \varphi (k) \tag{27}
\]

where \( k \geq G_1 (t) \) and by (15). Note that, here, we extend (15) to \( k \geq G_1 (t) \) (by (26)).
Note that \( \varphi (k, t) \) is the density of source fractions (with respect to \( t = \infty \)) (cf. (20) and (21)) with item-density \( k \). Hence, denoting by \( T \) the total number of sources at \( t = \infty \),

\[
\begin{align*}
f (k, t) &:= T \varphi (k, t) \\
f (k, t) &= G_2 (t) G_1 (t)^{\alpha - 1} T \varphi (k) \\
f (k, t) &= G_2 (t) G_1 (t)^{\alpha - 1} f (k)
\end{align*}
\]

(28)

(29)

(30)

denotes the actual density of sources (with respect to \( t = \infty \)) with item density \( k \) at time \( t \) and where \( f (k) = T \varphi (k) \) denotes the same at \( t = \infty \). Note that also here (15) is extended to \( k \geq G_1 (t) \).

2.3. Total number of sources and items at time \( t \)

The above function \( f (k, t) \) already allows one us calculate

\( T (t) \) = the total number of sources at time \( t \)

\( A (t) \) = the total number of items at time \( t \).

We have, by the definition of \( f (k, t) \) (cf. (29)) and since \( k \geq G_1 (t) \):

\[
\begin{align*}
T (t) &= \int_{G_1 (t)}^{\infty} f (k, t) \, dk \\
T (t) &= G_2 (t) G_1 (t)^{\alpha - 1} T \int_{G_1 (t)}^{\infty} \varphi (k) \, dk \\
T (t) &= G_2 (t) G_1 (t)^{\alpha - 1} T G_1 (t)^{1 - \alpha} \\
T (t) &= G_2 (t) T
\end{align*}
\]

(31)

a logical result. In the same way, for \( \alpha > 2 \), we have (cf. (29))

\[
\begin{align*}
A (t) &= \int_{G_1 (t)}^{\infty} kf (k, t) \, dk \\
&= TG_2 (t) G_1 (t)^{\alpha - 1} \int_{G_1 (t)}^{\infty} k \varphi (k) \, dk \\
&= TG_2 (t) G_1 (t)^{\alpha - 1} \frac{\alpha - 1}{\alpha - 2} G_1 (t)^{2 - \alpha}.
\end{align*}
\]

(32)

But, as is readily seen, on \( t = \infty \) (cf. [4, Chapter 2]), since \( \alpha > 2 \)

\[
\mu = \frac{A}{T} = \frac{\int_{G_1 (t)}^{\infty} k \varphi (k) \, dk}{\int_{G_1 (t)}^{\infty} \varphi (k) \, dk} = \frac{\alpha - 1}{\alpha - 2},
\]

(33)

(\( T = \) the total number of sources at \( t = \infty \), \( A = \) the total number of items at \( t = \infty \)).

Hence, (33) in (32) yields

\[
A (t) = AG_1 (t) G_2 (t),
\]

also a logical result since items at \( t \) are determined by the items in the already existing sources; the former one is determined

by \( G_1 \) and the latter one by \( G_2 \).

Since rank-frequency functions are equivalent with size-frequency functions (by the general relation (6), but now \( t \)-dependent) we are now in a position to calculate the time-dependent rank-frequency function \( \gamma (s, t) \).

2.4. The rank-frequency function, in function of time \( t \)

Since, at \( t \), the rank-densities are denoted by \( s \) via (17) and by (6), interpreted at time \( t \), we have (the inverse refers to the variable \( k \))

\[
\begin{align*}
s &= s (t) = \gamma^{-1} (k, t) = \int_{k}^{\infty} f (k', t) \, dk' \\
s &= TG_2 (t) G_1 (t)^{\alpha - 1} \int_{k}^{\infty} \varphi (k') \, dk',
\end{align*}
\]

(34)

by (29). Hence, using (15)

\[
s = s (t) = \frac{C}{\alpha - 1} \frac{G_2 (t) G_1 (t)^{\alpha - 1}}{k^{\alpha - 1}}
\]
denoting \( T (\alpha - 1) = C \) (hence \( f (k) = \frac{C}{k^\alpha} \)). This yields
\[
k = \gamma (s, t) = \frac{C^{\frac{1}{\alpha - 1}}}{((\alpha - 1) s)^{\frac{1}{\alpha - 1}}} G_2 (t)^{\frac{1}{\alpha - 1}} G_1 (t).
\]
Invoking (17) yields
\[
k = \gamma (s, t) = \frac{C^{\frac{1}{\alpha - 1}}}{((\alpha - 1) r)^{\frac{1}{\alpha - 1}}} G_1 (t)
\]
\[
\gamma (s, t) = \gamma (r) G_1 (t)
\]
with \( r \) and \( s \) related as in (17). Note that (35) follows from the fact that \( f (k) = \frac{C}{k^\alpha} \) and by applying the same calculation as performed here (but for \( t = \infty \), hence \( G_1 (\infty) = G_2 (\infty) = 1 \)). In fact the result can also be found in [4, Exercise II.2.2.6 (p. 134)] and its proof is also given in [10, Appendix].

In the next section, we will apply these results in the calculation of the general forms of the time-dependent \( h \)- and \( g \)-indices.

3. Formulae for the time-dependent \( h \)- and \( g \)-indices and properties of these functions of time

3.1. The time-dependent \( h \)-index

The \( h \)-index in this model is defined to be \( h (t) \) such that
\[
\int_{h(t)}^\infty f (k, t) \, dk = h (t) .
\]
Hence by (29) and (15) we have
\[
TG_2 (t) G_1 (t)^{\alpha - 1} \int_{h(t)}^\infty \frac{\alpha - 1}{k^\alpha} \, dk = h (t) .
\]
So
\[
h = h (t) = (TG_2 (t))^{\frac{1}{\alpha}} G_1 (t)^{\frac{\alpha - 1}{\alpha}} .
\]
Invoking that, for \( t = \infty \)
\[
h = h (\infty) = T^{\frac{1}{\alpha}}
\]
(see [12,10]). So (37) and (38) yield
\[
h (t) = G_1 (t)^{1 - \frac{1}{\alpha}} G_2 (t)^{\frac{1}{\alpha}} h .
\]
Note that the time-dependent \( h \)-index is proportional to the generalized geometric mean of the growth functions \( G_1 (t) \) and \( G_2 (t) \). In the following we will study properties of this generalized geometric mean.

3.2. The time-dependent \( g \)-index

There are two, equivalent, ways to define the \( g \)-index (at time \( t \)), denoted \( g (t) \): if
\[
\int_k^\infty k f (k', t) \, dk' = g (t)^2
\]
then
\[
g (t) = \gamma^{-1} (k, t) .
\]
Equivalently, \( g (t) \) is defined as
\[
\int_0^{g(t)} \gamma (r, t) \, dr = g (t)^2 .
\]
The latter approach has been followed in [5] (but only using the growth function \( G_1 \)—there denoted \( G \)). Here we will follow the former approach. Formula (40) yields, using (15) and (19) (and the notation \( C = T (\alpha - 1) \), already used above):
\[
G_2 (t) G_1 (t)^{\alpha - 1} \int_k^\infty \frac{C}{k^\alpha} \, dk' = g (t)^2
\]
\[
G_2 (t) G_1 (t)^{\alpha - 1} \frac{C}{\alpha - 2} k^{2 - \alpha} = g (t)^2 ,
\]
(43)
supposing \( \alpha > 2 \). But (34) yields (using (15) again)

\[
\gamma^{-1}(k, t) = G_2(t) G_1(t)^{\alpha - 1} \int_{k}^{\infty} \frac{C}{k^{\alpha}} dk' \\
g(t) = \gamma^{-1}(k, t) = G_2(t) G_1(t)^{\alpha - 1} \frac{C}{\alpha - 1} k^{1 - \alpha}
\]

by (41). So, solving (44) for \( k \), yields

\[
k = \left( \frac{\alpha - 1}{C} \frac{g(t)}{G_2(t)} \right)^{\frac{1}{\alpha - 2}} G_1(t)
\]

(45) in (43) yields, after some calculation and using that \( T = \frac{C}{\alpha - 1} \):

\[
g(t) = G_1(t)^{1 - \frac{1}{\alpha}} G_2(t)^{\frac{1}{2}} \left( \frac{\alpha - 1}{\alpha - 2} \right)^{\frac{\alpha - 1}{\alpha}} T^{\frac{1}{2}}.
\]

Invoking that, for \( t = \infty \),

\[
g = g(\infty) = \left( \frac{\alpha - 1}{\alpha - 2} \right)^{\frac{\alpha - 1}{\alpha}} T^{\frac{1}{2}}
\]

(see [7]). So (47) in (46) yields

\[
g(t) = G_1(t)^{1 - \frac{1}{\alpha}} G_2(t)^{\frac{1}{2}} g.
\]

Hence \( g(t) \) has the same proportionality factor as \( h(t) \) (see (39)). We will now study this proportionality factor which determines the evolution of \( h(t) \) and \( g(t) \) in time.

3.3. Properties of \( G_1(t)^{1 - \frac{1}{\alpha}} G_2(t)^{\frac{1}{2}} \)

Since \( G_1(t) \) and \( G_2(t) \) are cumulative growth distributions we have that

\[
G_1(0) = G_2(0) = 0 \\
\lim_{t \to \infty} G_1(t) = \lim_{t \to \infty} G_2(t) = 1
\]

and that \( G_1 \) and \( G_2 \) are strictly increasing in \( t \) and this increase is concave (example: (18) but we do not need this special case here).

So it is then trivial that \( h(t) \) strictly increases from 0 to \( h(\infty) \) and that \( g(t) \) strictly increases from 0 to \( g(\infty) \).

The lemma in the Appendix shows that the generalized geometric mean of strictly increasing concave functions is strictly increasing and concave so that also \( h(t) \) and \( g(t) \) are strictly increasing and concave. Of course, since the \( g \)-index is larger than or equal to the \( h \)-index we also have, for every \( t \):

\[
g(t) \geq h(t).
\]

4. Conclusions and open problems

In this paper we studied the aspect of time-dependence in the evolution of a Lotkaian system. Source and item densities are growing in time using two general cumulative growth distributions. Using this general framework yields formulae for the time-dependent size- and rank-frequency functions in the function of the final situation at \( t = \infty \), where we suppose we have a Lotkaian system. We show that at every time \( t \), we have a Lotkaian system with the same Lotka exponent, but where the growth is dependent on – in addition to the two growth distributions – this Lotka exponent.

Similar results are proved for the time-dependent \( h \)- and \( g \)-indices. Here we prove that both indices are proportional to their values at \( t = \infty \), and where the time-dependent proportionality factor is a generalized geometric mean of the two growth distributions. Also here the Lotka exponent is involved. We prove in a lemma in the Appendix that a general geometric mean of strictly increasing concave functions is strictly increasing and concave. Hence the same is true for the time-dependent \( h \)- and \( g \)-indices.

We formulate as an open problem the definition of other growth processes of sources and items, where the growth functions are acting on the source- and item-densities in a nonlinear way and the derivation from this framework of the time-dependent size- and rank-frequency functions and the time-dependent \( h \)- and \( g \)-indices.
Appendix

Let \( x \) and \( y \) be two positive real numbers. We say that \( z \) is a generalized geometric average of \( x \) and \( y \) if there exists a \( \lambda \in [0,1] \) such that
\[
z = x^{\lambda} y^{1-\lambda}.
\] (A.1)

The classical geometric mean is given when \( \lambda = \frac{1}{2} \); then \( z = \sqrt{xy} \).

In Section 3 we found that
\[
h(t) = G_1(t)^{-\frac{1}{2}} G_2(t)^{\frac{1}{2}} h
\]
(A.2)
\[
g(t) = G_1(t)^{-\frac{1}{2}} G_2(t)^{\frac{1}{2}} g
\]
(A.3)

hence \( h(t) \) and \( g(t) \) are equal to their values at \( t = \infty \) multiplied by a generalized geometric mean of the growth functions \( G_1(t) \) and \( G_2(t) \) (note, indeed, that, since \( \alpha > 1 \) we have that \( \lambda = 1 - \frac{1}{\alpha} \in [0,1] \)).

We have the following general lemma.

**Lemma.** Let \( G_1, G_2 \) be two functions of the variable \( t \) which are strictly increasing and concave. Then their generalized geometric average (for \( \alpha > 1 \))
\[
G(t) = (G_1(t))^{1-\frac{1}{\alpha}} (G_2(t))^{\frac{1}{\alpha}}
\]
(A.4)

is strictly increasing and concave.

**Proof.** That \( G \) is strictly increasing in \( t \) is trivial. For the concavity we need to calculate \( G'' \). We have
\[
G'(t) = \frac{\alpha - 1}{\alpha} \left( \frac{G_1(t)}{G_2(t)} \right)^{-\frac{1}{\alpha}} G_1(t) + \frac{1}{\alpha} \left( \frac{G_1(t)}{G_2(t)} \right)^{\frac{\alpha - 1}{\alpha}} G_2(t)
\]
(A.5)
\[
G''(t) = \frac{\alpha - 1}{\alpha^2} \left( \frac{G_1(t)}{G_2(t)} \right)^{-\frac{1}{\alpha}} \frac{G_1(t) G_1'(t) - G_1(t) G_1'(t) - G_1(t) G_2'(t)}{G_2(t)^2} + \frac{1}{\alpha} \left( \frac{G_1(t)}{G_2(t)} \right)^{\frac{\alpha - 1}{\alpha}} \frac{G_2(t) G_2'(t) - G_2(t) G_2'(t) - G_2(t) G_2'(t)}{G_2(t)^2}
\]

The second and fourth terms are clearly strictly negative since \( G_1 \) and \( G_2 \) are concave. The first and third term, together, equal
\[
\frac{\alpha - 1}{\alpha^2} \left( \frac{G_1(t)}{G_2(t)} \right)^{-\frac{1}{\alpha}} \frac{G_1(t) G_1'(t) - G_1(t) G_1'(t) - G_1(t) G_2'(t)}{G_2(t)^2} + \frac{1}{\alpha} \left( \frac{G_1(t)}{G_2(t)} \right)^{\frac{\alpha - 1}{\alpha}} \frac{G_2(t) G_2'(t) - G_2(t) G_2'(t) - G_2(t) G_2'(t)}{G_2(t)^2}
\]
\[
= \frac{\alpha - 1}{\alpha^2} \left( \frac{G_1(t)}{G_2(t)} \right)^{-\frac{1}{\alpha}} \frac{G_1(t) G_1'(t) - G_1(t) G_1'(t) - G_1(t) G_2'(t)}{G_2(t)^2}
\]
\[
\leq 0, \quad \text{since} \quad \alpha > 1.
\]

So \( G''(t) < 0 \) for all \( t \) and hence \( G(t) \) is concave.

\[\square\]

**References**


