On the Predictability of Future Impact in Science

Raj Kumar Pan,[1] Orion Penner,[2] Alexander M. Petersen,[2] Kimmo Kaski,[1] and Santo Fortunato[1]

[1] *BECS, Aalto University School of Science, P.O. Box 12200, FI-00076, Finland*
[2] *IMT Institute for Advanced Studies Lucca, 55100 Lucca, Italy*

Science has evolved a merit driven career advancement process in which an individual is promoted through the various career stages on the strength of his or her past achievements and perceived potential for future achievement. Committees charged with the task of evaluating the past accomplishments and projecting the future success of applicants are at the core of these advancement decisions, whether they be fellowship, grants, tenure track hires, tenure, *etc.* In this context, evaluation is rarely a straightforward matter, as recent case studies indicate that grant committee selection decisions do not necessarily correlate with either the peer-review process or cumulative achievement measures. Faced with applicant pools ranging in size from dozens, for tenure track hires, to thousands for national fellowship and tenure competitions, it is a great challenge to distill the contents of each *curriculum vitae* to an assessment of an individual's past, present and future impact and arrive to an appropriate ranking of candidates. Further, it is important to recognize that *future* impact is at the heart of this matter because the ultimate questions are: Which candidate will be most successful in *this* position? With *this* fellowship? Do the most with *this* grant? Emphasis is typically placed on past success but, for the most part, it is only relevant in so far as it correlates with the future success.

For individual researchers the most widely known measure of impact is Hirsch's h-index [1]. While it has been shown that a correlation exists between a researcher's current and future h-index, h-index is clearly a measure of a researcher's past accomplishments [2]. In recent work Acuna et al. propose a model for a researcher's future h-index and thereby establish a clear and concrete framework for connecting a researcher's current CV to his or her future impact in research [3]. On the conceptual level this aligns much better with the goal of most career advancement decisions, as they are largely focused on what a researcher *will* produce rather than what he or she *has* produced. However, on a technical level cumulative achievement models, such as the Acuna model, suffer from methodological flaws.

Here we present a cross-sectional analysis of 762 longitudinal careers drawn from three disciplines: physics, biology, and mathematics. By applying future impact models to these careers we identify a number of subtle, but critical,

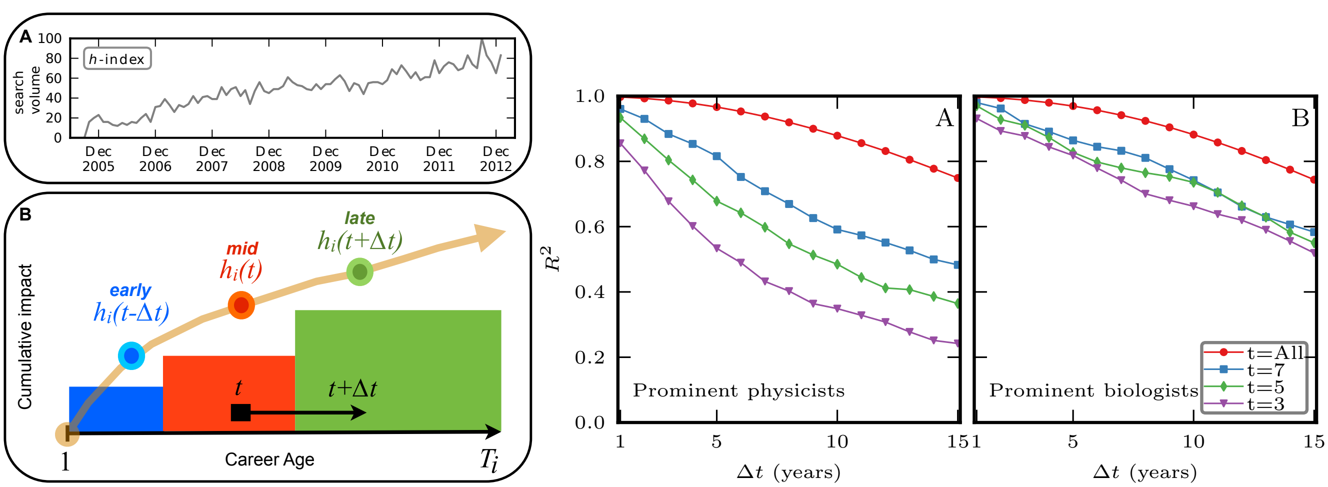

Figure 1. *(Left, A)* Monthly *Google* search volume for the term "h-index", normalized to % peak value. Since the initial publication proposing the h-index on Nov. 15, 2005 [1], there has been roughly a 4-fold increase in h-index search volume over the 7-year period Dec. 2005 - Dec. 2012, capturing the persistent increasing interest and use of h over time [**?**]. *(Left, B)* Schematic illustration of the career stages that define academic careers. The h-index is a cumulative non-decreasing quantity intended to measure both the productivity and impact of a scientist i up to year t. However, models for predicting $h(t+\Delta t)$ must account for two important factors: (i) $h(t)$ is non-decreasing so that "predictability" measures for $h(t+\Delta t)$ can be artificially inflated, and (ii) variations in the "risk" profile and the "production function" of scientists across career stages must be accounted for in predictive models. *(Right)* The "predictive power" of the regression model of the h-index for different disciplines and for different career age cohorts (years since first publication $t = 3, 5, 7$). The curve for $t =$All shows the model where all career ages were lumped together.

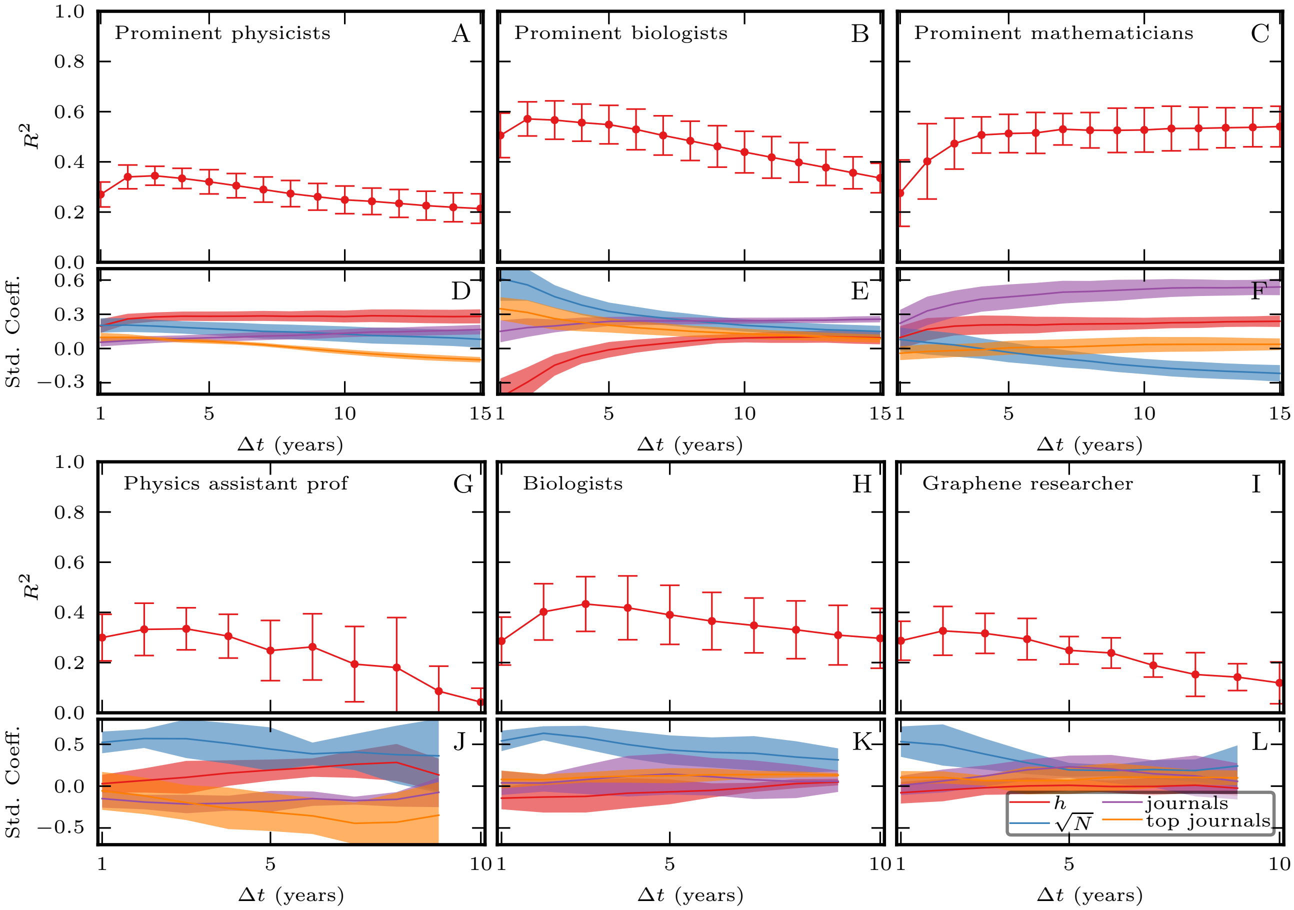

Figure 2. The "predictive power" of h-index increments ($\Delta h(t, \Delta t)$) for different discipline. (A,B,C) Variation of the mean R^2 as a function of time period Δt over which the increment is calculated for established physicists, biologists and mathematicians. The mean is calculated by averaging over different career age cohorts $t = 2, \ldots, 15$. (D,E,F) Variation of the mean standard coefficient as a function of Δt. The shaded region indicates the 95% confidence error bars. Similar plots are also shown for relatively young researchers in (G,H,I) for assistant professors in physics, biologists and graphene researchers. As the careers of young scientists are short, in this case the mean is calculated by averaging over different career age cohorts $t = 2, \ldots, 8$.

flaws in current models. Specifically, cumulative non-decreasing measures like the h-index contain intrinsic spurious autocorrelation, resulting in a significant overestimation of their "predictive power". Applying the model to a scientist's annual h-index change (a non-cumulative measure), the models exhibit far less predictive power. Moreover, the predictive power of these models vary greatly with the career age of scientists, producing least accurate estimates for already risk-burdened early career researchers. These results place in doubt the suitability of linear regression models of future h-index for real application in recruitment decisions and indicate that more effort is needed to develop and benchmark career predictability algorithms [4, 5].

[1] J. E. Hirsch, Proceedings of the National Academy of Sciences **102**, 16569 (2005).
[2] J. E. Hirsch, Proceedings of the National Academy of Sciences **104**, 19193 (2007).
[3] D. E. Acuna, S. Allesina, and K. P. Kording, Nature **489**, 201 (2012).
[4] O. Penner, A. M. Petersen, R. K. Pan, and S. Fortunato, Physics Today **66**, 8 (2013).
[5] O. Penner, R. K. Pan, A. M. Petersen, K. Kaski, and S. Fortunato, ArXiv e-prints (2013), arXiv:1306.0114 [physics.soc-ph].