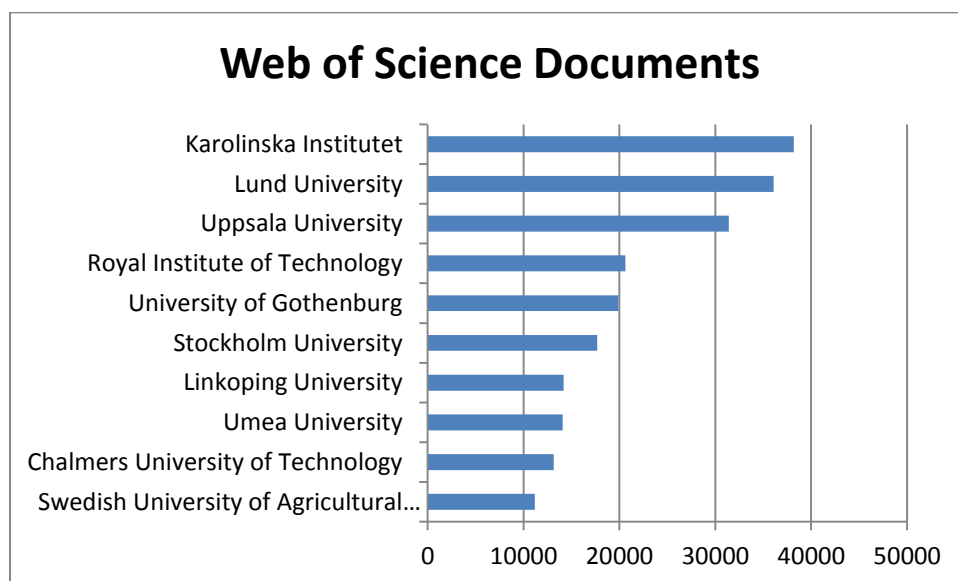


Jämförelse mellan tio svenska lärosäten över tre indikatorer samt den nationella indikatorn

Bo Jarneving, digitala tjänster

2015-10-24

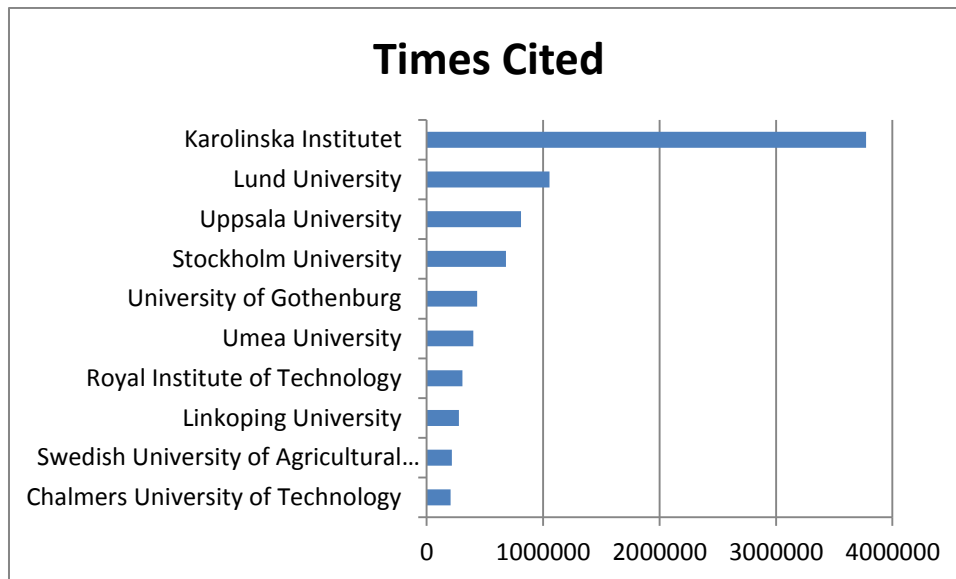
För observationsperioden 2004-2013 har ett dataunderlag för tio svenska universitet/högskolor hämtats från databasen *InCites* (Thomson Reuters) och analyserats. Dataunderlaget baseras på artiklar, reviewartiklar samt konferensbidrag indexerade i Thomson Reuters *Web of Science, Core Collection* uppdaterad 2015-09-11. Den första indikatorn mäter antalet indexerade publikationer (figur 1). Vi ser att Karolinska institutet har flest publikationer i *Web of Science* (WoS), följt av Lunds universitet, Uppsala universitet, KTH och därefter Göteborgs universitet på femte plats. Göteborgs universitet har noterbart fler publikationer än Stockholms universitet vilket troligen beror på att Stockholms universitet saknar en medicinsk fakultet och att Sahlgrenska Universitetssjukhusets publikationer medräknats i denna version av *InCites*. Lägg därtill det faktum att medicin är ett område som har hög täckningsgrad i WoS. I termer av synlighet på den internationella tidskriftsmarknaden ser vi sammanfattningsvis att Karolinska institutet tillsammans med våra äldsta universitet (Uppsala, Lund) tydligt dominerar.



Figur 1. Fördelning av artiklar över universitet, perioden 2004-2013.

Om vi sedan skiftar fokus till citeringar och ser dessa som en funktion av *användning av tidigare forskning* är mönstret till väsentliga delar detsamma: Karolinska institutet dominerar än mer och detta avspeglar, som tidigare nämnts, att medicinska forskningsområden är välindexerade i WoS. Karolinska institutet följs sedan av Lunds och Uppsala universitet som tidigare. Däremot ser vi att Stockholms universitet mottar avsevärt fler citeringar än Göteborgs universitet och att KTH faller från fjärde plats till sjunde (figur 2). För en mer uttömmande tolkning av dessa data måste man ta hänsyn

till både indexeringsgrad av olika forskningsområden och respektive lärosätes forskningsprofil. Detta är en mycket omfattande uppgift och ryms inte i denna not.



Figur 2. Fördelningen av citeringar över universitet, perioden 2004-2013.

Både antal publikationer och antal citeringar säger något om det inflytande en institution har på den globala forskningen. Dessa indikatorer avspeglar förstås till stor del storleken på universiteten i termer av antal forskare och andra resurser. För att bättre kunna bestämma forskningsprestationers uppskattade *kvalitet* för vi samman uppgifter om volym citeringar och volym publikationer och relaterar dessa till normer. I detta sammanhang talar vi om *fältnormer* vilka avspeglar den genomsnittliga citeringsmedelvärdet för publikationer inom ett visst forskningsområde. Sådana *relativa indikatorer* kan beräknas på lite olika sätt. Vanligen utgår man från en referensmängd med publikationer som avgränsas av forskningsområde, dokumenttyp och publikationsår för att beräkna ett förväntad värde, *e*. I *Incites* definieras det fältnormerade medelvärdet för en publikation som *Category Normalized Citation Impact (CNCI)* vilket ges som:

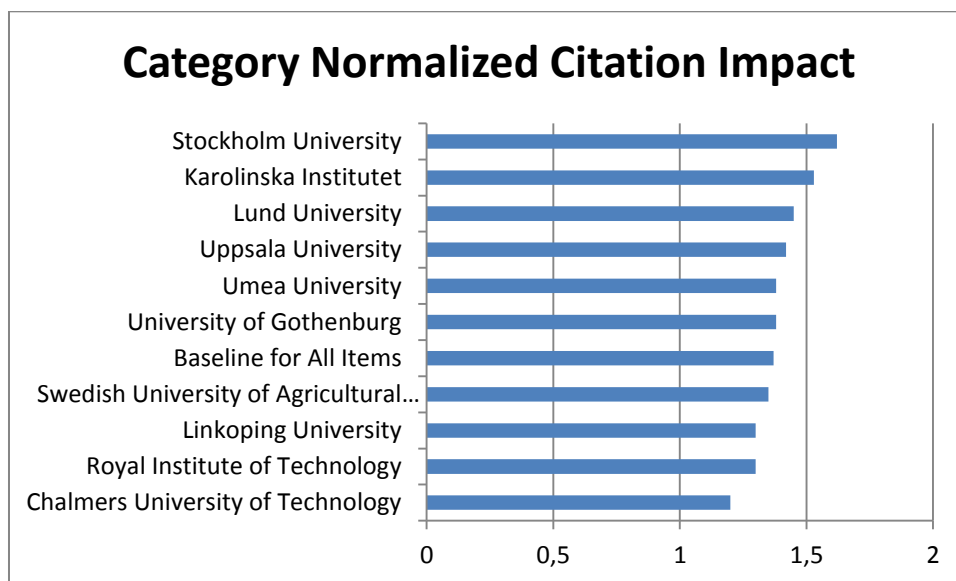
-det observerade antalet citeringar *c* för en publikation *p* delat med det förväntade värdet *e*, där *e* är citeringsmedelvärdet i *Web of Science* för publikationer som tillhör samma ämnesområde, har samma publikationsår och är tilldelad samma dokumenttyp som *p*.

När en publikation är tilldelad mer än ett ämnesområde beräknas medelvärdet för alla kvoter (*c/e*) som:

$$CNCI = \frac{1}{n} \sum_{i=1}^n \frac{c_p}{e_i}$$

där *n* är antalet ämnesområden (journal subject categories). För aggregerade publikationer beräknas den genomsnittliga CNCI. När vi använder denna indikator får vi fram intressant information. Stockholms universitet intar ledningen med *CNCI* = 1,62 följt av Karolinska institutet med 1,53. Göteborgs universitet (1,38) ligger på delad femteplats med Umeå universitet. På nionde och tionde

plats ligger KTH (1,3) och Chalmers (1,2) (figur 3). I detta fall har vi använt oss så kallad hel räkning av publikationer och citeringar, vilket innebär att varje lärosäte får exakt det antal publikationer och antal citeringar som publikationerna genererat, oberoende av hur många andra institutioner och författare som samarbetat om dessa publikationer. Ett annat sätt att räkna innebär att publikationer och citeringar delas upp över författare och/eller institutioner, så kallad fraktionerad räkning. Eftersom det än så länge inte är möjligt att beräkna varje enskild författares (och motsvarande institutions) faktiska bidrag till en publikation är bägge räknesätten lika riktiga och kan komplettera varandra. Fraktionerad räkning avspeglar då inverkan av forskningssamarbete medan hel räkning avspeglar den faktiska antalet publicerade och citerade forskningresultat.



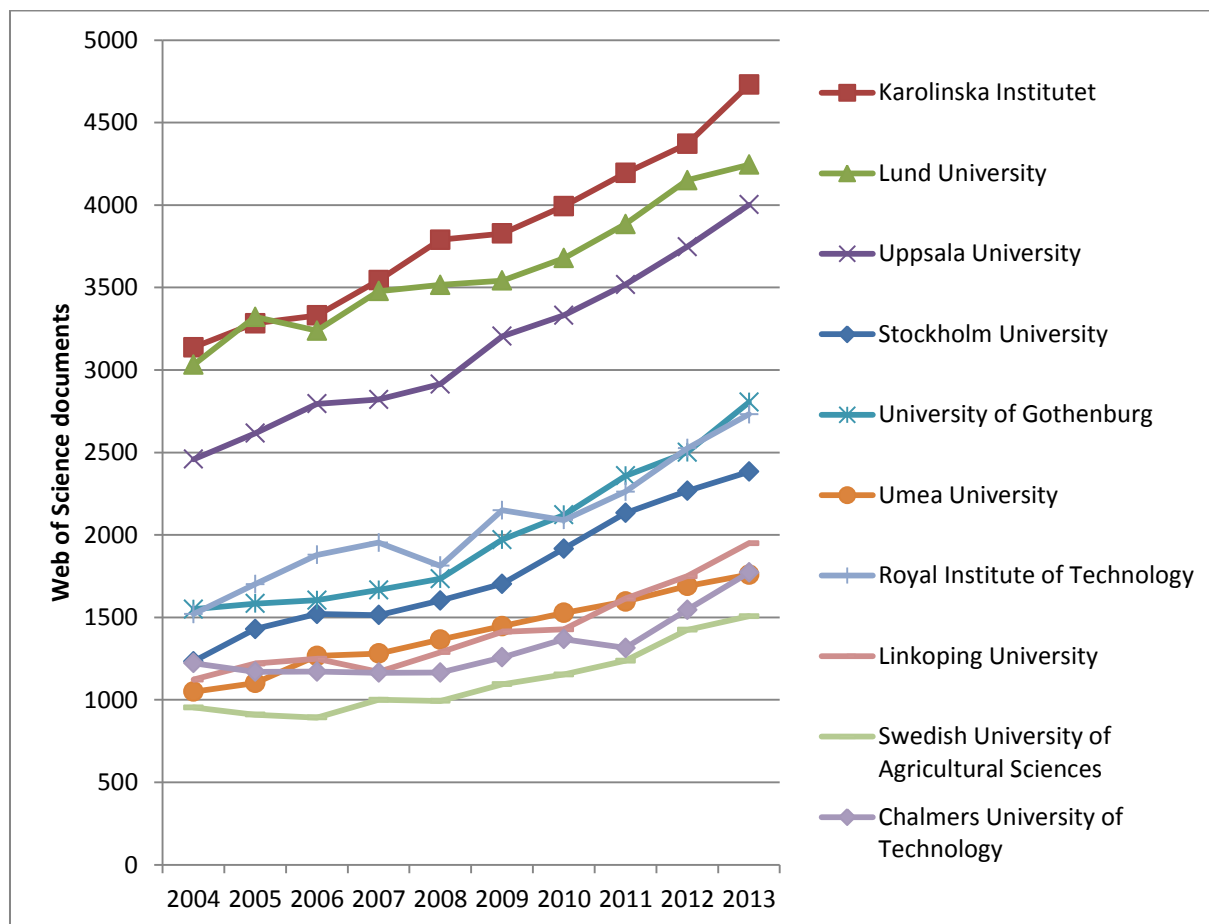
Figur 3. Fältnormaliserad citeringsfrekvens för tio universitet, perioden 2004-2013.

Det kan vara svårt att rättvist *rangordna* universiteten efter forskningsprestation utifrån indikatorer baserade på publicerad forskning eftersom olika indikatorer avspeglar olika delar av inflytande (impact) på den globala forskningen. Det finns en stark, positiv korrelation mellan samtliga tre indikatorer och det starkaste sambandet finns mellan *antal publikationer* och *antal citeringar* ($r = 0,96$; $r_s = 0,88$).

Trender

I detta avsnitt studerar vi utvecklingen över tid med avseende på de tre indikatorerna. Över tid ser vi en tydlig linjär tillväxt av indexerade WoS-publikationer för de flesta lärosätena (figur 4).

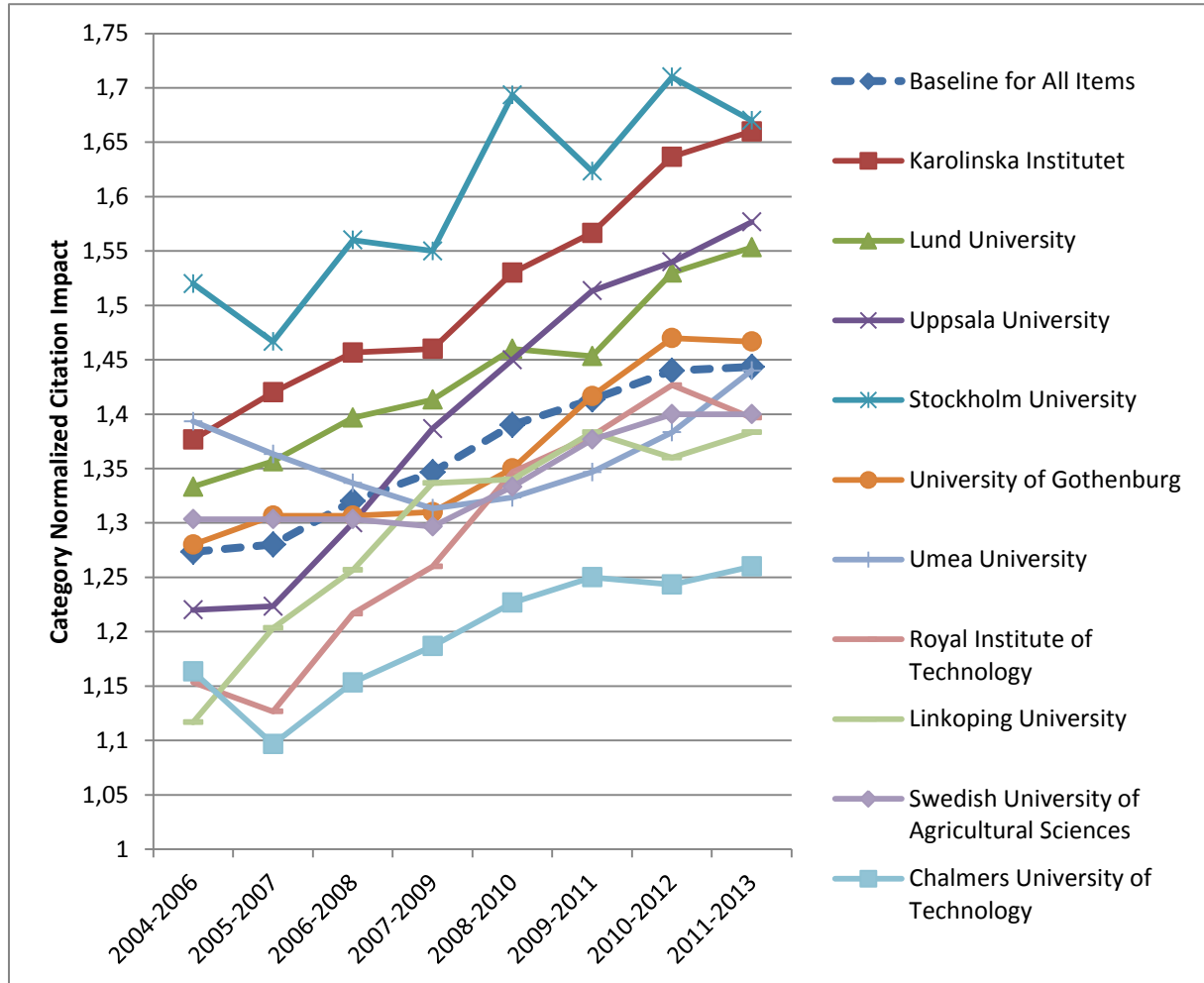
Utvecklingen över tid stämmer förstås med fördelningen av Web of Science Documents i figur 1 och är likartad för de olika lärosätena. Totalt ser vi en genomsnittlig årsproduktion på 2165 WoS-publikationer medan den högsta genomsnittliga årsproduktionen (3820 WoS-publikationer) tillfaller Karolinska institutet och den lägsta (1117 WoS-publikationer) SLU.



Figur 4. Fördelning av Web of Science Documents över universitet och år, perioden 2004-2013.

Trender baserade på citeringsdata är ofta svåra att visualisera och tolka eftersom enstaka högt citerade publikationer lätt åstadkommer ett oscillerande mönster. Av den anledningen har vi valt att redovisa fältnormaliserade citeringsfrekvenser som glidande medelvärden över treårsintervall (figur 5) med avsikten att presentera ett mindre kaotiskt mönster. I figur 5 är också en baseline baserad på all citeringsdata för dessa lärosäten inlagd i diagrammet så att man kan avgöra vilka lärosäten som samlas över respektive under en gemensam baseline. Även här är trenden tydligt positiv över tid och

det glidande medelvärdet för baseline rör sig från 1,28 till 1,44 över observationsperioden.

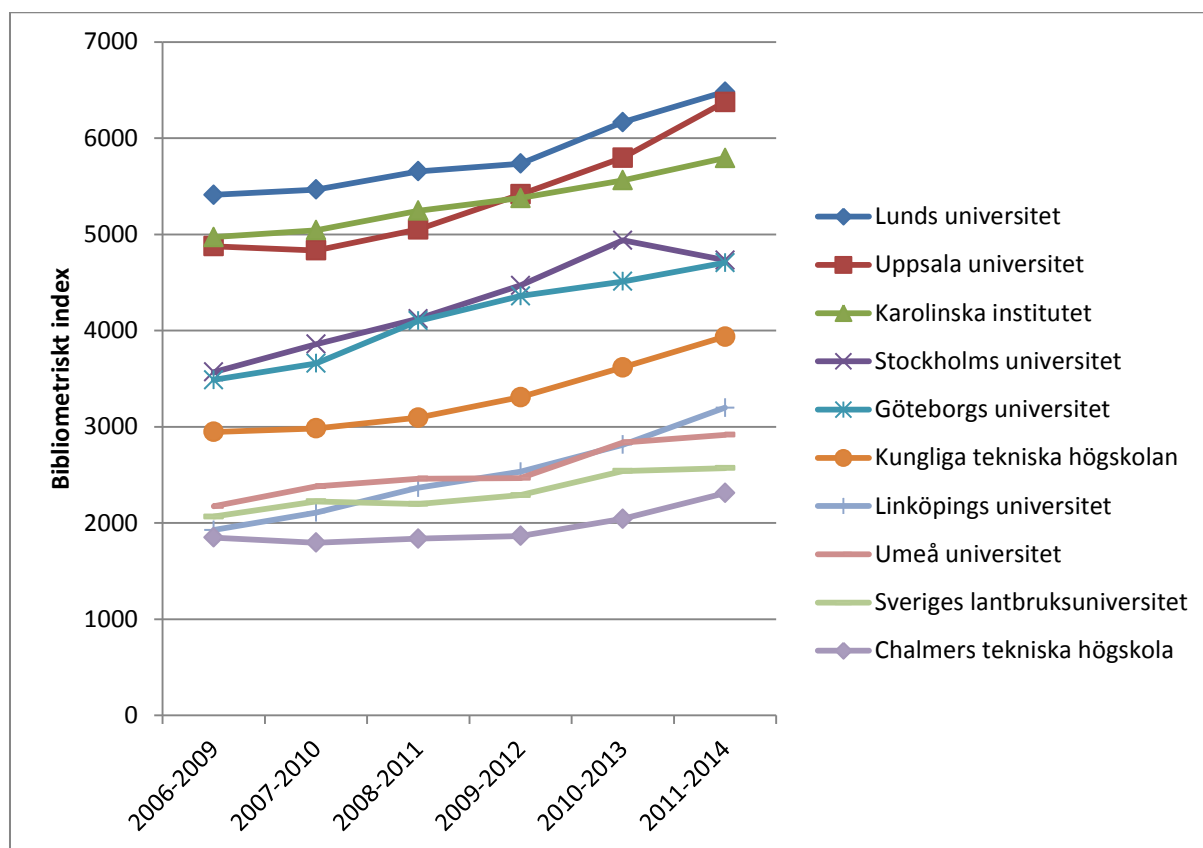


Figur 5. Fältnormaliserad citeringsfrekvens för tio universitet, perioden 2004-2013. Glidande medelvärden och treårsintervall.

Jämförelse med den nationella indikatorn

Vetenskapsrådets nationella indikator genererar ett s.k. *bibliometriskt index* som uppdateras årligen och används för medelsfördelning till lärosäten i Sverige. Detta index baseras på både volym publikationer och citeringar. Dessutom räknar man fraktionerat och normerar för både fält och produktion av vetenskapliga publikationer. Denna indikator är komplex och mindre transparent än de indikatorer som redovisats här. Hitintills har denna indikator använts inför sex budgetår (2011-2016) där varje observationsperiod som ligger till grund för beräkningarna omfattar fyra år. Inför varje nytt budgetår flyttas sedan observationsperioden fram med ett år så att den överlappar med föregående period med tre år. Vi ser att de fördelningar som redovisas i figur 6 samlar Lunds- och Uppsalas universitet samt Karolinska institutet i det övre intervallet, följt av Stockholms universitet, Göteborgs universitet samt KTH. Jämför vi detta mönster med indikatorn Web of Science Documents (figur 4) ser vi en överensstämmelse med samma lärosäten i det övre intervallet, men enstaka rangordningar skiljer sig åt. Indikatorn Times Cited (antal citeringar) överensstämmer relativt väl med

med bibliometriskt index i figur 6.¹ Slutligen kan vi konstatera att indikatorn Category Normalized Citation Impact (fältnormerad citeringsfrekvens) i figurerna 3 och 5 stratifierar lärosätena på ett likartat sätt men att relativt stora skillnader finns vad gäller enskilda rangpositioner.



Figur 6. Bibliometriskt index över sex budgetår. Källa: Vetenskapsrådet.

Dessa impressionistiska jämförelser säger oss att de olika indikatorerna i är relaterade med varandra i större eller mindre utsträckning. Skillnader mellan indikatorerna är trots allt påtagliga och avspeglar att olika aspekter lyfts fram (till exempel förmågan att publicera sig i centrala tidskrifter gentemot förmågan att rikta sig mot forskningsfronten och bli högt citerad). Man kan förstås använda dessa diagram för att anställa parvisa jämförelser där ett sådant intresse finns, men vi behöver ett annat mätinstrument för att mer generellt fastställa sambandet mellan den sammansatta nationella indikatorn och de mer grundläggande indikatorerna. För detta ändamål kan vi mäta korrelationen mellan indikatorerna. Innan denna kan mätas måste vi ordna data så att observationsperioderna överensstämmer. För samtliga indikatorer väljs då intervallet 2006-2013 och vi tillämpar samma periodisering som Vetenskapsrådet med fyra års överlappande perioder. I tabell 1 är korrelationskoefficienter för sambandet mellan bibliometriskt index (vetenskapsrådets nationella indikator) och de tre tidigare presenterade indikatorerna ordnade över fem perioder. Om vi läser tabellen i horisontalplanet kan vi följa korrelationen över dessa perioder. Vi ser ett tämligen konsistent mönster för sambandet mellan bibliometriskt index och de övriga indikatorerna: korrelationen är starkast mellan bibliometriskt index och Times Cited (antal citeringar) och svagast

¹ Här gör vi en jämförelse med ett stapeldiagram (figur 2) och ett linjediagram över tid (figur 6) eftersom indikatorn antal citeringar (times cited) är mindre informativ som tidsserie (citeringar avtar snabbt med ett minskande citeringsfönster).

mellan bibliometriskt index och Category Normalized Citation Impact (fältnormaliserad citeringsfrekvens). Indikatorn Web of Science Documents har näst starkast samband med bibliometriskt index med undantag av perioden 2006-2009.

Tabell 1. Korrelationen (*r*) mellan bibliometriskt index och tre indikatorer över fem observationsperioder (2006-2013): antal publikationer (Web of Science Documents), antal citeringar (Times Cited) och fältnormerad citeringsfrekvens (Category Normalized Citation Impact).

	Times Cited 2010-2013	Times Cited 2009-2012	Times Cited 2008-2011	Times Cited 2007-2010	Times Cited 2006-2009	Web of Science Documents 2010-2013	Web of Science Documents 2009-2012	Web of Science Documents 2008-2011	Web of Science Documents 2007-2010	Web of Science Documents 2006-2009	Category Normalized Citation Impact 2010-2013	Category Normalized Citation Impact 2009-2012	Category Normalized Citation Impact 2008-2011	Category Normalized Citation Impact 2007-2010	Category Normalized Citation Impact 2006-2009
Bibliometriskt index 2010-2013	,86					,74					,54				
Bibliometriskt index 2009-2012		,92					,76					,52			
Bibliometriskt index 2008-2011			,93					,78					,76		
Bibliometriskt index 2007-2010				,85					,79					,72	
Bibliometriskt index 2006-2009					,78					,83					,74

Sammanfattning

De tre indikatorer som presenterades inledningsvis baseras helt på Web of Science data och avspeglar impact på utifrån Thomson Reuters databaser med allt vad det innebär av biased täckning och indexering. För medicin och naturvetenskap är Web of Science traditionellt viktig för den vetenskapliga kommunikationen och bibliometriska mätningar av densamma. För humaniora och samhällsvetenskap, och då i synnerhet för sådan forskning som har mer lokal anknytning, är Web of Science citeringsdatabaser inte särskilt användbara. I detta sammanhang får man också komma ihåg att ett lärosätes forsknings- och ämnesprofil bestämmer i vilken utsträckning dess publikationer ger rättvisande avtryck i Web of Science. Jämförelser mellan universitet som syftar mot att fastställa skillnader i forskningsprestation (forskningskvalitet och impact) kan därför med tveksamhet anställas med tillgängliga metoder och dataunderlag.

Med detta sagt kan vi konstatera att samtliga tre indikatorer avspeglar en stratifiering med Karolinska institutet, Uppsala- och Lunds universitet i det övre stratat och Linköpings Universitet, SLU och Chalmers i ett undre strata. KTH, Göteborgs universitet och Umeå universitet hamnar någonstans mitt emellan och Stockholms universitet ligger i topp vad gäller fältnormerad citeringsfrekvens. Rangordningen varierar en del över indikatorerna vilket uttrycker att samtliga indikatorer bör studeras i ett och samma sammanhang. Så kan man till exempel se att Stockholms universitet producerar relativt lite Web of Science publikationer (mindre än till exempel KTH) men att dess publikationer har ett hög genomsnittlig fältnormaliserad citeringsfrekvens. Vi har också noterat hur en hög indexeringsgrad i Web of Science och förmodat hög andel citerade publikationer gett Karolinska institutet en mycket hög citeringsfrekvens och synlighet på den internationella tidskriftsmarknaden.

Jämförelserna mellan den nationella indikatorn och de tre övriga indikatorerna indikerar att de senare inte särskilt väl avspeglar bibliometriskt index, trots att dataunderlaget torde vara detsamma (Web of Science). Av de indikatorer som jämförts med bibliometriskt index är Times Cited (totalt antal citeringar) den indikator som korrelerar starkast.