

# Cellulär automat simulerar utbredning av Taraxacum

Examensarbete 15hp

12-06-06

Författare  
Jimmy Bergman  
a06jimbe@student.his.se

Handledare  
Noél Holmgren  
noel.holmgren@his.se

Högskolan i Skövde, Box 408  
541 28, Skövde

## Innehållsförteckning

1. Introduktion	3
2. Metod och material	5
2.1 Modellen	5
2.2 Mått på jämförelse mellan simulering och observation	6
2.3 Quadrat analysis method	7
2.4 Om observation och simulering	8
3. Resultat	9
3.1 Fördelning av observerad data	9
3.2 Simuleringarnas robusthet	9
3.3 Jämförelse mellan observation och simulering	10
4. Diskussion och slutsats	13
5. Referenser	16

# 1. Introduktion

Framsteg inom fysik ger hopp om att modeller som bygger på enkla parametrar kan redogöra för det allmänna uppträdandet av komplexitet i system. Bak et al. (1987) har visat att uppkomsten av komplexitet från enkla lokala interaktioner kan vara spontan och därför kan utgöra källan till komplexitet i naturen. I en studie av cellulära automater (CA) från 1983 fastslår Wolfram et al. att ”en analys av de allmänna dragen i deras beteende kan tänkas ge allmängiltiga resultat rörande många komplexa systems beteende”. CA har använts i predator-bytesmodeller såsom WATOR, en familj ekologiska simuleringar av hajar och fiskar. För växtekologi har modellen använts för att beskriva två eller fler växtarter som producerar frö med slumpmässig spridning och groningen med resultat som verkar förklara ojämnheten i verkliga utbredningar (Ermentrout et al. 2003). Bolliger et al. (2003) visade att en stokastisk CA kan reproducera ett historiskt skogslandskap sett till dess organisation (aggregering) och komplexitet.

I detta arbete undersöks utbredningen av de fleråriga örterna maskrosor (*Taraxacum sect. Ruderalia*). De flesta arter av maskrosor producerar frö asexuellt som kan färdas långväga med vinden. Groning av frö sker på eller nära jordytan. Maskrosens rot kan nå ett djup av 3,5-5 m. Skott växer ut från den övre delen av roten till en krona som kan regenerera sig även om växten skärs av vid eller under ytan. Maskrosor kan överleva i många år och är väl anpassade till mark som regelbundet hävdas genom bete eller klippning.

Sannolikheten att ett frö gror ökar med ökad ljusexponering. På störd mark såsom kultur- och ruderatmark sker kolonisering snabbt. Ett blad i rosetten är 5-25 cm eller längre, vilket innebär rosettradier om 10-50 cm eller större. Rosetten hindrar solljus från att nå ner till jordytan, vilket minskar sannolikheten för groningen av frön. En individ kan ha många bladrosetter på ytan. Ett frö som landar i närheten av en etablerad individ kommer, även om ljusförhållandena är tillräckliga, att

underjord behöva konkurrera med den etablerade individens rotsystem och skott. Så inomarts-konkurrens förekommer både ovan och under jord och påverkar om ett frö gror och om den nya individen överlever på lång sikt.

Den cellulära automatens snabbhet och grafiska klarhet förser oss med ett sätt att testa en mängd modeller och parametrar med nästan ögonblicklig feedback. Den utgör en metod för att formalisera en teori om en påstådd mekanism till något beräkningsbart. Då mycket biologi i bästa fall är kvalitativ är CA ett bra verktyg då den inte kräver att specifika värden på parametrar preciseras för att ge ett resultat (Ermentrout et al. 2003). Maskrosor frodas på hävdad mark. I frånvaro av högre vegetation är maskrosor en konkurrenskraftig art som kanske mest har varandra att konkurrera med om näring och ljus.

Denna studie undersöker vilken grad av konkurrens som ger upphov till ett observerat utbredningsmönster. Den observerade utbredningen kontrolleras även mot slumpmässig fördelning. Studien hypotiserar att maskrosornas spatiala utbredning sett till aggregering och komplexitet kan beskrivas utifrån enkla regler för spridning och konkurrens.

## 2. Metod och material

### 2.1 Modellen

En CA består av ett regelbundet fält av celler där varje cell kan tilldelas ett variabelvärde (se fig. 1).

I denna studie kan en cell anta två värden, antingen är cellen tom eller så är den upptagen.

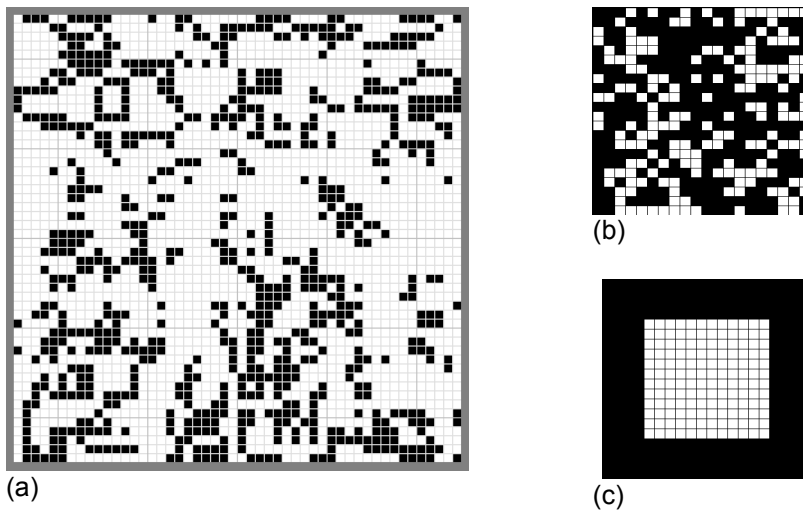
Systemets tillstånd i ett givet ögonblick specificeras sedan genom att ange vilka celler som är tomma och vilka som är upptagna. Systemet utvecklas deterministiskt i tiden enligt regler baserade på tillståndet hos grannceller. Då en cell i detta fall endast kan anta två värden, så kan en upptagen cell antingen överleva eller dö ut. En tom cell kan antingen koloniseras eller förbli tom (kolonisering sker antingen genom rotskott eller groningen av frö).

I denna studie används två regler. Regel A säger att celler med 1-2 levande (upptagna) grannar överlever till nästa generation och att frön gro i celler med 1-2 levande grannar. Denna regel efterliknar en situation med hög känslighet för inomartskonkurrens. Regel B låter celler med 1-4 levande grannceller överleva och frö gro i celler med 1-4 levande grannceller, och förutsätter därmed lägre inomartskonkurrens.

Automaten består av ett toroidialt 50\*50 rutnät, en cell som befinner sig på den ena kanten är granne med en cell på kanten mitt emot. En cells grannskap omfattar 6 celler. För att illustrera, med regel A gro ett frö i en cell om 1/6 eller 2/6 grannceller är upptagna.

Två olika initialvillkor används; ett ordnat och ett slumpmässigt (se Figur 1b och c). Det ordnade villkoret har alla celler ordnade i en kvadrat och en klustersannolikhet (sannolikhet att en slumpmässigt vald cell ingår i ett kluster, här definierat som att en cell har 3 levande grannceller) på 97%. Det slumpmässiga villkoret genereras med Random Bitmap Generator utifrån atmosfäriskt brus och har en klustersannolikhet på 16% (Haahr 2012). Om initialvillkoren visar sig vara

oväsentliga för utvecklingen, visar detta att modellen ger robusta resultat.



Figur 1. (a) visar ett exempel på hur det kan se ut när en CA körs och (b) och (c) visar slumpmässigt respektive ordnat initialvillkor.

## 2.2 Mått på jämförelse mellan simulering och observation

En fraktal kan fritt definieras som ett självliknande mönster. En struktur är självliknande om dess delar är likt helheten i miniatyr. Denna egenskap kan analyseras med boxcounting, vilket används i studien för att bestämma de fraktala egenskaperna i utbredningsmönstret. Boxcounting är en analysmetod som bygger på att täcka en fraktal med kvadrater för att se hur många kvadrater som krävs för att täcka ytan och sedan upprepa detta med kvadrater av en annan rutstorlek. Detta ger en logaritmisk funktion för kvadratstorlek (x-axeln) och antalet kvadrater som krävs för att täcka fraktalen (y-axeln). Denna funktions lutning kallas box-counting dimension eller Minkowski–Bouligand-dimensionen (Moisy 2006). Denna dimension kan beskrivas som ett mått mönstrets ytfyllande kapacitet, hur mönstret skalas annorlunda mot ytan det ingår i. En fraktal kurva kan betraktas som för detaljerad för att vara endimensionell, men för enkel för att vara tvådimensionell, dess fraktala dimension ligger därför mellan 1 och 2. En ytfyllande kurva är en kurva som trots att den ej har någon tjocklek är tillräckligt förvriden att den helt fyller en yta.

Klustersannolikhet används för att mäta organisation i den simulerade datan och definieras som sannolikheten för att en slumpmässig levande cell ingår i ett kluster och uttrycks i procent. En cell med tre eller fler levande grannceller betraktas ingå i ett kluster. Metoden anammas från Bolliger et al. (2003).

Algoritmisk komplexitet används för att mäta komplexitet i den observerade och den simulerade utbredningen och definieras som storleken på den enklaste algoritm som kan beskriva datan. Metoden utnyttjar efter Spratt et al. (2002) faktumet att GIF-filer nästan är optimalt komprimerade. Komplexiteten i grafiken över utbredningsmönstren bestäms genom storleken på GIF-filen av utbredningsmönstret efter att headern tagits bort.

### **2.3 Quadrat analysis method**

Slumpmässig fördelning är en mindre vanlig fördelning i naturen och förekommer där individer av en art uppträder i en homogen miljö och mer eller mindre oberoende av varandra. För att undersöka om fördelningen i den observerade datan är slumpmässig används Quadrat analysis method (Rogerson 2001). Med Quadrat analysis läggs ett rutnät över någon slags spatial data. Sedan beräknas antalet datapunkter i varje ruta. I slumpmässigt fördelad data kommer det genomsnittliga antalet punkter per ruta vara ungefär lika med variansen hos antalet punkter per ruta. Om det finns en stor mängd variabilitet i antalet punkter från ruta till ruta, indikerar detta att datan har en tendens att aggregera. Testet använder en  $\chi^2$ -statistika och varians-medelvärde-förhållandet, VMR (Variance Mean Ratio),

$$\chi^2 = \frac{s^2 m - 1}{\bar{x}}$$

$$VMR = \frac{s^2}{\bar{x}}$$

där  $m$  är antal rutor,  $\bar{x}$  och  $s^2$  medelvärde och varians för antalet punkter per ruta. Om  $VMR > 1$  är fördelningen aggregerad, det finns en stor mängd variation i antalet datapunkter per cell, vissa celler har avsevärt många fler punkter än förväntat och andra har avsevärt många färre än väntat.

## **2.4 Om observation och simulering**

Artens utbredning har baserats på blommande individer på en gräsmatta på 20\*20 m. Begreppen överlevnad och kolonisering syftar på den perenna örten. De observerade blommorna ger ingen exakt position för individuella plantor; inte alla individer blommar samtidigt och en individ kan ha flera blommor. Blommorna ger dock information om populationens fördelning i rummet.

Simuleringarna har gjorts med den öppna programvaran Golly (Trevorrow & Rokicki 2011)

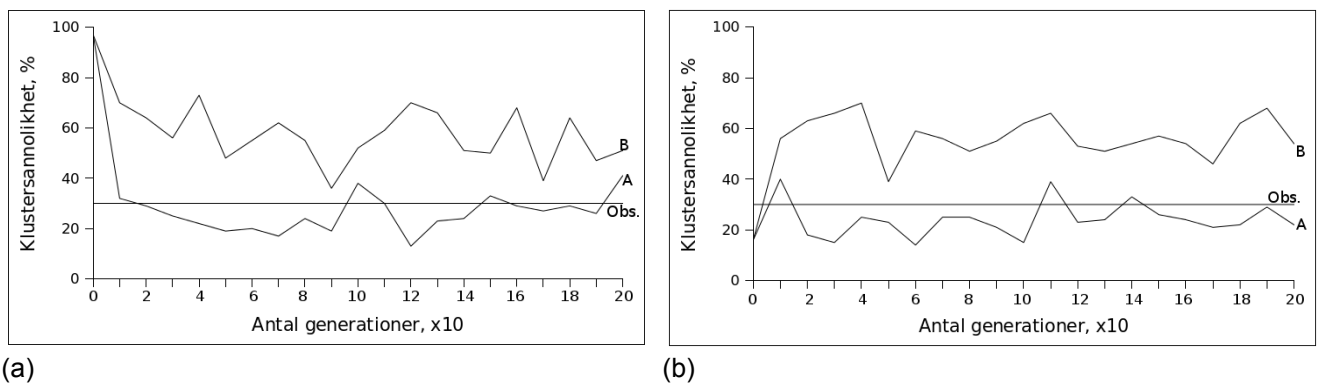
Modellen kördes utifrån de två olika initialförhållandena med två olika regler. Varje simulering har gjorts i 200 generationer där var tionde generation har analyserats.



### 3. Resultat

#### 3.1 Fördelning av observerad data

För att undersöka om fördelningen i den observerade datan är slumpmässig användes Quadrat analysis method. VMR är i detta fall 22,621 vilket innebär att utbredningsmönstret definitivt är klusterartat och ej slumpmässigt fördelat. ( $\chi^2=1877,543$ ,  $df=83$ ,  $p=<0,001$ ).



Figur 2. Klustersannolikhet under 200 generationer för regel A respektive B. Stabil linje visar klustersannolikhet för observation: (a) visar simuleringar med ordnat initialvillkor och (b) visar simuleringar med slumpmässigt initialvillkor.

#### 3.2 Simuleringarnas robusthet

Om modellen är robust borde simuleringar med samma regel ej vara känsliga

för initialvillkoret (Fig. 3a och b). Medelvärdena för klustersannolikhet i simuleringarna med regel

A skiljer sig inte med avseende på initialvillkor ( $t=0,819$ ,  $df=38$ ,  $P=0,417$ ). Differensen i

medelvärde är 1,8 (24,2 och 26). För simuleringarna med regel B skiljer sig inte heller

medelvärdena ( $t=0,104$ ,  $df=38$ ,  $P=0,917$ ). Differensen i medelvärde är 0,3 (56,8 och 57,1).

Om man däremot analyserar hela fördelningen av antal celler per klusterstorlek för en generation

per simulering (generation vald som närmast simuleringens medelvärde för klustersannolikhet) så

skiljer sig dessa åt beroende på initialvillkoren. Med regel A ( $\chi^2=38,779$ ,  $df=8$  och  $P=<0,0001$ ), med

regel B ( $\chi^2=28,432$ ,  $df=8$ ,  $P=0,0004$ ).

### 3.3 Jämförelse mellan observation och simulering

Den generation i simuleringarna som ligger närmast den observerade fraktala dimensionen är regel A med det slumpmässiga initialvillkoret (Tab. 1). Funktionerna har en näst intill identisk lutning (fig. 4a och b). Det slumpmässiga utbredningsmönstret skiljer sig dock minst från observerat värde.

Tabell 1. Fraktal dimension för observation och simuleringar.

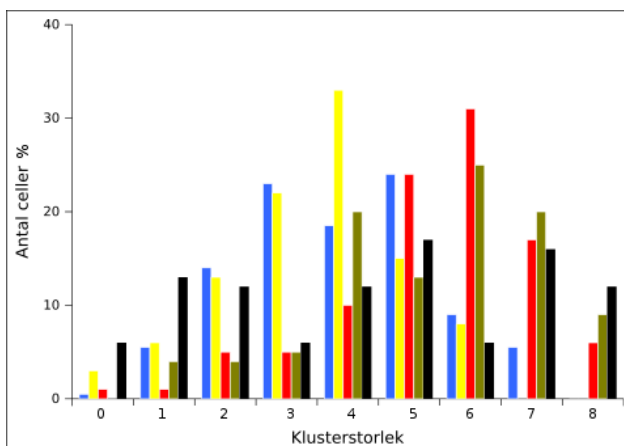
Data	Fraktal dimension
Observation	1,807 +/- 0,137
Simulering, Regel A, ordnat initialvillkor	1,782 +/- 0,156
Simulering, Regel A, slump. initialvillkor	1,794 +/- 0,148
Simulering, Regel B, slump. initialvillkor	1,588 +/- 0,315
Simulering, Regel B, ordnat initialvillkor	1,708 +/- 0,206
Slumpmässigt genererad bitmap	1,811 +/- 0,140

Medelvärde för klustersannolikhet (Tab. 2) mellan observation och regel A med ordnat initialvillkor skiljer sig inte ( $t=1,281$ ,  $df=25$ ,  $P=0,211$ ). Regel A och slumpmässigt initialvillkor skiljer sig inte heller från observation ( $t=1,845$ ,  $df=25$ ,  $P=0,076$ ) Observation och regel B med ordnat liksom slumpmässigt initialvillkor skiljer sig dock ( $6,188$ ,  $df=25$ ,  $P=<0,0001$ , respektive  $t=7,938$ ,  $df=25$ ,  $P=<0,0001$ ).

Tabell 2. Klustersannolikhet för observation och simuleringar. Klustersannolikhet för observation har beräknats genom stickprov.

Data	Klustersannolikhet, %
Observation	30,075 +/- 3,064
Simulering, Regel A, ordnat initialvillkor	26 +/- 6,943
Simulering, Regel A, slump. initialvillkor	24,2 +/- 6,955
Simulering, Regel B, slump. initialvillkor	57,1 +/- 7,635
Simulering, Regel B, ordnat initialvillkor	56,8 +/- 10,319
Slumpmässigt genererad bitmap	16

Samtliga simuleringar skiljer sig från observationen vad gäller fördelningen av antal celler per klusterstorlek för en generation per simulering (se Figur 3 och Tab. 3).



Figur 3. Klusterstorlek för simuleringar jämfört med observation. Svarta staplar är observerad data, gula slupm. regel A, blåa ordnat regel A, gröna ordnat regel B, och röda slupm. regel B.

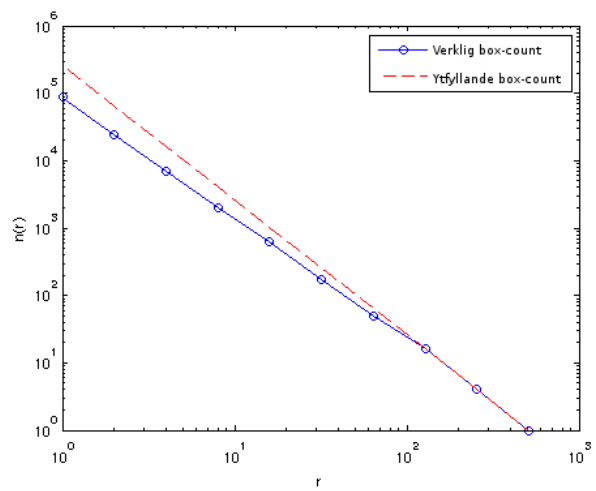
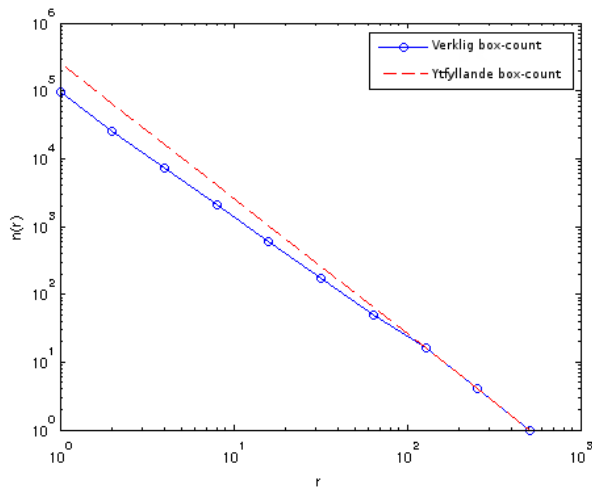
Tabell 3. Klusterstorlek för simuleringar jämfört med observation.

Simulering	Jämfört med observation
Slupm. initialvillkor, regel B	$X^2=129,938$ , $df=8$ , $P=<0,0001$
Slupm. Initialvillkor, regel A.	$X^2=113,671$ , $df=8$ , $P=<0,0001$
Ordnat initialvillkor, regel A	$X^2=85,618$ , $df=8$ , $P=<0,0001$
Ordnat initialvillkor, regel B	$X^2=85,922$ , $df=8$ , $P=<0,0001$

Den simulerade datan som komprimeras till en storlek närmast den observerade datan (6323 bytes) är simulering med regel B och slumpmässigt initialvillkor (6213 bytes). En slumpmässigt genererad GIF-fil komprimeras till 7427 bytes (Tab. 4).

Tabell 4. GIF-komprimering för observation och simuleringar.

Data	GIF-komprimering
Observation	6 323 bytes
Simulering, Regel A, ordnat initialvillkor	7 402 bytes
Simulering, Regel A, slupm. initialvillkor	7 370 bytes
Simulering, Regel B, slupm. initialvillkor	6 213 bytes
Simulering, Regel B, ordnat initialvillkor	6 760 bytes
Slupmässigt genererad bitmap	7427 bytes



(a)

(b)

Figur 4. Box-counting med kvadratstorlek ( $r$ ) på x-axeln och antal kvadrater som krävs för att täcka mönstret ( $n(r)$ ) på y-axeln: (a) Box-counting av observerad utbredning (1,807) och (b) box-counting av simulering med slumpmässigt initialvillkor och regel A (1,794). Den blå kurvan är den verkliga funktionen medan den röda är funktionen för en ytfillande tvådimensionell bild.

## 4. Diskussion och slutsats

Resultaten visar att en enkel diskret modell till en viss grad kan beskriva maskrosornas utbredning sett till dess aggregering och komplexitet. Utifrån klustersannolikhet är det regel A, en högre inomartskonkurrens, som bäst beskriver den observerade utbredningen sett till dess aggregering. Kolonisering och överlevnad för celler med 1-2 levande grannceller är en mer realistisk utveckling för maskrosornas utbredning än kolonisering på och överlevnad för celler med 1-4 levande grannceller. Mekanismen kan vara inomartskonkurrens, men med endast två olika regler kvantifieras ej effekten.

Utbredningens självliknande egenskap beskrivs bäst av en slumpmässig utbredning (skillnad 0,004). Den beskrivs också bra av regel A (skillnad 0,013). Funktionernas lutning avviker från en ytfyllande box-count (funktionen för en ytfyllande tvådimensionell bild), vilket indikerar ett möjligt fraktalt beteende för skalor där  $r < 100$  (figur 3a och b). Detta beteende är logiskt om man föreställer sig hur den första maskrosen koloniserar en tom yta. Genom rotskottsfrökning skjuts skott ut åt alla riktningar, vilka i sin tur växer en egen rot och producerar sina egna skott, vilka skjuter ännu längre bort. På detta sätt formas teoretiskt en fraktal geometri (likt ett itererande funktionssystem). Dock kan ju fler frö gro och kolonisera denna yta. Då kommer konkurrens in i bilden och utbredningens självliknelse minskar antagligen. Den observerade graden av självliknelse överensstämmer dock bra med regel A och en högre inomartskonkurrens.

Utifrån befintlig data kan det inte uteslutas att simuleringarna inte är känsliga för initialförhållandena vad gäller vald klusterstorlek. En indikation på motsatsen är skillnaden mellan simuleringarna då klusterstorlek per cell beräknas för ett stickprov från varje fördelning. Detta gäller dock endast en generation och skulle behöva undersökas ytterligare.

Inget stickprov från simuleringen liknar till fördelningen av cell per klusterstorlek den observerade utbredningen. Om detta är grund för att förkasta modellen kan diskuteras. Att förvänta sig en perfekt reproduktion av observationens hela fördelning är inte realistiskt. Det finns en ganska stor variation i simuleringarna emellan och antagligen också generationer emellan. Den observerade fördelningen verkar vara bimodal (tvåtoppig). Utifrån fördelning av klusterstorlek finns ett tecken på att bimodal fördelning reproduceras med ordnade initialvillkor, men detta är osäkert.

Att det är regel B som bäst beskriver maskrosornas verkliga utbredning sett till algoritmisk komplexitet beror på att den uppvisar en hög aggregering med resulterande öppna områden emellan. Regel A genererar ett mer utspritt mönster som uppenbarligen är svårare att beskriva. Detta kan tolkas som att individer kan leva närmare in på varandra än vad regel A förutsätter.

Då maskrosor har fröspridning skulle det vara mer verklighetstroget att använda ett större grannskap, såsom 16 eller 32 celler, men modellen medger ej detta. Med ett grannskap om 6 celler kan man tänka sig att den vegetativa förökningen inryms, medan fröspridningen ej inryms på samma vis. Ett sätt att fånga denna spridning kan vara att introducera stokastisk kolonisation.

Observation av blomning är tacksam, med den uppenbara nackdelen att blommorna sällan befinner sig just på platsen för bladrossetten. Detta bidrar till osäkerhet i den observerade datan. Genom upprepade, noggranna observationer och simuleringar kunde resultatens statistiska säkerhet ökas.

Denna studie ser endast till utbredningen av en art. Ett mer rigoröst sätt att testa modellens förmåga att simulera ekologiska spatiala fördelningar skulle vara att se på flera samspelande arter. Inbördes interaktioner mellan arter introducerar en annan nivå av komplexitet och dynamik. Genom att variera de enkla regler som bestämmer den cellulära automatens utveckling kan antagligen både

jämviktssystem liksom dynamiska, icke-jämviktssystem framgångsrikt simuleras. Modellen tillåter en mängd tillstånd och därför skulle mycket komplicerade mönster eller system med många komponenter, såsom artrika miljöer, kunna simuleras med intressanta resultat.

Modelleringen kan vara ett verktyg för att förstå de simulerade biologiska processerna både ur spatialt och temporalt perspektiv. CA är speciellt intressant då den framför en möjlig förståelse av hur samspelet individer emellan ger upphov till ekosystemets övergripande dynamik och utveckling. Eftersom regler om utdöende och kolonisering kan skraddarsys så kan olika biologiska mekanismer inkorporeras i modellen. Detta gör den värdefull inom ekologin.

## 5. Referenser

Bak, P., Tang, C. & K. Wiesenfeld (1987). "Self-organized criticality: an explanation of  $1/f$  noise". *Physical Review Letters* 59 (4): 381–384.

Bolliger, J., Julien C. Sprott & David J. Mladenoff (2003). "Self-organization and complexity in historical landscape patterns". *OIKOS* 100: 541–553.

Ermentrout, G. Bard & Leah Edelstein-Keshet (1993). "Cellular automata approaches to biological modeling". *Journal of Theoretic Biology*. 160, 97-133.

Haahr, Mads (2012). Random Bitmap Generator, baserad på atmosfäriskt brus.  
<http://www.random.org/bitmaps/>

Moisy, Frederic (2006). "1D, 2D and 3D Box-counting". Scientific/educational MATLAB database. Technische Universität München.  
<http://m2matlabdb.ma.tum.de/>

Rogerson, Peter (2001). "Statistical methods for geography." Första upplagan. Sage Publications. 248 s.

Sprott, J. C., J. Bolliger, & D.J. Mladenoff (2002). "Self-organized criticality in forest landscape evolution." *Physics Letters. A* 297: 271 –276 .

Trevorrow, Andrew & Tom Rokicki. Golly (Ver. 2.3). [Mjukvara]. Hämtad från:  
<http://golly.sourceforge.net/>

Wolfram, Stephen (1983). "Statistical mechanics of cellular automata." *Reviews of Modern Physics*. 55. s. 601.