

**Soluautomaattien ja agenttipohjaisten mallien käyttö  
kaupunkimallinnuksessa**

Diplomityö  
Rakennetun ympäristön laitos  
Insinööritieteiden korkeakoulu  
Aalto-yliopisto

Espoossa 29 toukokuuta 2017

Tekniikan kandidaatti Antti Federley

Valvoja: Professori Kirsi Virrantaus  
Ohjaaja: Professori Kirsi Virrantaus



---

**Tekijä** Antti Federley

---

**Työn nimi** Soluautomaattien ja agenttipohjaisten mallien käyttö  
kaupunkimallinnuksessa

---

**Koulutusohjelma** Geomatiikka

---

**Pää-/sivuaine** Geoinformatiikka**Koodi** M3002

---

**Työn valvoja** Professori Kirsi Virrantaus

---

**Työn ohjaaja(t)** Professori Kirsi Virrantaus

---

**Päivämäärä** 29.05.2017**Sivumäärä** 57**Kieli** Suomi

---

### Tiivistelmä

Tämän diplomityön tavoitteena on tutkia kaupunkimalleissa käytettyä menetelmää geosimulointi sekä kahta siinä käytettyä tekniikkaa: soluautomaatteja ja agenttipohjaisia systeemeitä. Geosimulointi on tapa mallintaa ympäristöä ottaen huomioon siinä toimivien objektien vuorovaikutus ja siitä syntyvät kompleksiset ilmiöt. Työssä pyritään selvittämään miten kyseistä kahta tekniikkaa käytetään geosimuloinnissa ja mitä haasteita niihin liittyy. Kun haasteet on tunnistettu, tutustutaan kahteen ajankohtaiseen geosimuloinnin malliin ja siihen, miten kumpikin malli toteutti simuloinnin käyttäen agenttipohjaisia systeemeitä ja soluautomaatteja.

Työ alkaa kirjallisuustutkimuksella, jossa tutustutaan soluautomaatteihin ja agenttipohjaisiin systeemeihin. Työssä keskitytään kyseisten tekniikoiden kohdalla erityisesti kaupunkimallinnuksen sovelluksiin. Kummastakin tekniikasta käsitellään lyhyesti tekniikoiden peruseriaatteet ja niiden erityisominaisuudet. Agenttipohjaiset systeemit soveltuvat hyvin mallintamaan yksilöiden toimintaa ympäristössä. Soluautomaateilla kyetään taasen mallintamaan yksinkertaisilla säännöillä ympäristöä ja siihen kohdistuvia muutoksia, jolloin kumpikin tekniikka täydentää toisiaan. Lopuksi tässä osuudessa työtä tutustutaan geosimulointiin sekä yhteen sen moderniin perusmalliin.

Viimeisessä osuudessa työtä suoritetaan geosimulointiin liittyvien haasteiden tunnistaminen ja kahden geosimulointimallin vertailu asetettuihin haasteisiin nähden. Tavoitteena oli tunnistaa millaisia haasteita sekä soluautomaatteihin että agenttipohjaisiin systeemeihin liittyy mallintamisessa. Vertailun tuloksena selvisi, että kumpikin malli toteutti geosimulointimallin eri tarkoitusta varten, jolloin mallit käyttivät tekniikoita eri tavoin. Esimerkiksi yleiskäytännöllisempi malli ei ottanut kantaa mallin validointiin ja verifiointiin samalla tavalla kuin yksityiskohtainen, tiettyä käyttötarkoitusta varten luotu malli. Geosimulointia voidaan käyttää useissa eri sovelluksissa ja kohteissa, joten haasteiden huomioiminen riippuu mallin käyttötapauksesta.

---

**Avainsanat** geosimulointi, paikkatieto, soluautomaatit, agenttipohjaiset systeemit

---



---

**Author** Antti Federley

---

**Title of thesis** Cellular automata and agent-based models in urban modelling

---

**Degree programme** Degree Programme in Geomatics

---

**Major/minor** Geoinformatics

**Code** M3002

---

**Thesis supervisor** Professor Kirsi Virrantaus

---

**Thesis advisor(s)** Professor Kirsi Virrantaus

---

**Date** 29.05.2017

**Number of pages** 57

**Language** Finnish

---

## Abstract

The purpose of this thesis was to investigate a method used in urban modelling called geosimulation and two techniques associated with it: cellular automata and agent-based modelling. Geosimulation is a method to model environment and to take into account interaction of objects within it causing complex phenomena. Thesis aims to find how these two techniques are used in geosimulation and what kind of challenges affiliate with these techniques. Once the challenges are identified, a study is conducted on two current geosimulation models and how these models used cellular automata and agent-based modelling.

The work begins with a literature review that covers cellular automata and agent-based systems. The emphasis on these techniques was urban modelling and its applications. Basics of the both techniques are reviewed along with more advanced characteristics. Agent-based systems are well-suited for modelling individuals and their actions within an environment. Cellular automata on the other hand is a suitable modelling technique to model environment and its changes with simple transition rules. The result is that both methods complement each other. Finally at this part of the work geosimulation and one modern basic model of geosimulation is reviewed.

Final part of this thesis covers identifying challenges associated with geosimulation and comparison of two geosimulation models using these challenges. The objective is to identify what kind of challenges is associated with cellular automata and agent-based systems with urban modelling. Based on the comparison of the models, both geosimulation models had different function and because of that models used methods differently. For example one of the models was more general purpose therefore the model didn't use validation and verification methods same way as the model made for a designated purpose, which was a more detailed model. Geosimulation can be used for multiple different applications and situations therefore taking challenges into consideration depends on the use case.

---

**Keywords** geosimulation, geographic information, cellular automata, agent-based systems

---

## Alkusanat

*Haluan kiittää professori Kirsi Virrantausta työni aiheeni valinnassa. Aihealue kosketti osittain jo aiemmin tekemiäni töitä, mutta toi mukanaan myös uutta opittavaa. Lisäksi kommentit työn edetessä ja tärkeiden aineistojen keräämisessä avustaminen olivat erittäin tärkeitä työn loppuun saattamiselle. Myös rohkaisu diplomityön tekemiseen ilman ulkoista rahoitusta tai yhteistyötahoa oli arvokasta.*

*Toisekseen haluan kiittää professori Virrantausta ja myös muita Aalto yliopiston opetushenkilökuntaa inspiroivasta opetuksesta vuosien varrella. Kiitos kuuluu myös niille opiskelutovereille, joiden kanssa pääsin osallistumaan lukuisille kursseille ja tekemään yhteistyötä.*

*Haluan kiittää myös työnantajaani Meritaito Oy:ta tuesta ja ymmärryksestä, että haluan saattaa diplomityöni ja opintoni loppuun työni ohella.*

*Lopuksi tahdon vielä kiittää niitä henkilöitä, jotka tukivat ja avustivat diplomityöni tekemistä. Kiitoksia isälleni, äidilleni ja siskolleni diplomityöni ja opintojeni tukemisesta kaikkina aikoina. Kiitos kuuluu myös ystävilleni, jotka kuuntelivat ja neuvoivat kun sitä tarvitsin. Tahdon vielä lopuksi kiittää TkT Jussi Nikanderia, joka ehti lukea työni lävitse ja tarjota tarpeellista kritiikkiä.*

Helsinki 29.05.2017

Antti Federley



## Sisällysluettelo

Sisällysluettelo .....	1
Termien ja lyhenteiden selitykset.....	3
1 Johdanto .....	4
1.1 Tutkimuksen tausta.....	4
1.1.1 Kaupunkimallit pohjana .....	6
1.2 Tutkimusongelma ja työn rajaus .....	8
1.3 Työssä käytetyt aineistot ja rakenne.....	9
2 Spatiaalisuus ja mallintaminen.....	10
2.1 Yleistä.....	10
2.1.1 Spatiaalisuus.....	10
2.1.2 Paikkatiedon erityisominaisuudet .....	10
2.2 Mallintaminen .....	10
2.3 Simulointi .....	11
3 Soluautomaatti .....	13
3.1 Yleistä.....	13
3.1.1 Taustaa .....	13
3.1.2 Soluautomaatin yleiset ominaisuudet.....	14
3.1.3 Formaali kuvaus soluautomaateille.....	15
3.1.4 Luokittelu .....	16
3.2 Soluautomaattien pääominaisuudet .....	16
3.2.1 Naapurusto .....	16
3.2.2 Tila .....	18
3.2.3 Säännöt.....	19
4 Agentit.....	22
4.1 Yleistä.....	22
4.1.1 Tyypilliset piirteet .....	22
4.1.2 Agenttien luokittelu.....	25
4.2 Agenttien erityispiirteet .....	26
5 Geosimulointi.....	28
5.1 Yleistä.....	28
5.1.1 Yleiset ominaisuudet.....	28
5.1.2 Geosimulointimallien pääpiirteet .....	28
5.1.3 Geosimuloinnin haasteet .....	29
5.2 Agentit geosimuloinnissa .....	30
5.3 Geosimuloinnin moderni malli.....	31
5.3.1 GAS.....	31
5.3.2 GAS tilat .....	32
5.3.3 Geo-referoidut säännöt.....	32
5.3.4 GAS:n naapurustot .....	33
5.4 Soluautomaattien ja agenttien heikkoudet.....	33
6 Geosimuloinnin uusien tutkimuksien arviointi .....	35
6.1 Tutkimuksien ja mallien arvioinnin kriteerit.....	35
6.1.1 Haasteet tarkemmin.....	36

6.1.2	Soluautomaattien haasteita.....	37
6.2	Tutkittavien töiden esittely.....	40
6.2.1	Työ 1: Agent Based Simulation for Understanding Urban Dynamics, Kiranjith, Jonu John Thomas Giridhar K. 2016.....	40
6.3	Työ 2: Geosimulation of urban growth and demographic decline in the Ruhr: a case study for 2025 using the artificial intelligence of cells and agents. Rienow, Andreas. Stenger, Dirk. 2014.....	42
6.4	Tutkimuksien arviointi .....	45
6.4.1	Agenttipohjaisten metodien arviointi.....	45
6.4.2	Soluautomaattien heikkouksien arviointi.....	48
6.5	Loppupäätelmät vertailusta .....	48
7	Päätäntö.....	51

## Termien ja lyhenteiden selitykset

**ABM:** Agent-based Modelling – agenttipohjainen mallinnus.

**CA:** Cellular Automata – soluautomaatti.

**GAS:** Geographic Automata System – maantieteellinen automaattijärjestelmä.

**GIS:** Geoinformation System – paikkatietojärjestelmä.

**SVM:** Support Vector Machine – tukivektorikone.

**UGM:** Urban Growth Model – kaupunkien kasvumalli.

**UML:** Unified Modeling Language – graafinen mallinnuskieli.

# 1 Johdanto

## 1.1 Tutkimuksen tausta

Ympäristön mallinnus pyrkii tuottamaan sekä tietoutta että viestintää liittyen ilmiöön, jotta suunnittelijoilla ja päätöksentekijöillä olisi enemmän ymmärrystä ilmiöstä (Altartouri 2015). Mallintamisen tavoite on lähtökohtaisesti valita ne tärkeät ominaisuudet ja attribuutit ympäristöstä, jotka yksinkertaistavat todellisuutta riittävästi (Yassemi et al. 2007). Ympäristön mallintamisen kompleksisuus on aiheuttanut sen, että useilla eri tieteenaloilla on alettu kiinnittämään huomiota ympäristön mallintamisen tutkimukseen. Aloista esimerkiksi tietojenkäsittely, tilastotieteet sekä sosiaalitieteet ovat tutkineet ympäristön mallintamista ja kyseisiltä aloilta voidaan lainata muille tieteenaloille menetelmiä ja mallinnustapoja. Täten päätöksentekijöille voidaan tarjota aikaisempaa laajempaa ympäristöön liittyvää tietoa. (Altartouri 2015)

Ympäristön ja sen muodostaman systeemin sekä sen sisältämien kohteiden väliset riippuvuudet ovat olleet yksi tieteellisen tutkimuksen tärkeä osa-alue. Kohteiden välisiä riippuvuuksia sijainnin suhteen kutsutaan spatiaalisuudeksi. Tästä on syntynyt keskeinen käsite maantieteellisessä mallinnuksessa. (Benenson & Torrens 2004) Perinteisissä ympäristömalleissa ei ole otettu spatiaalisuutta yleisesti huomioon, mutta geoinformaatiojärjestelmien (GIS) avulla monimuotoisen datan käsittely ja analysointi on muodostunut tehokkaammaksi, jolloin myös spatiaalisuus on huomioitu malleissa. (Yassemi et al. 2007)

Lisänä ominaisuuksien välisiin suhteisiin ja sijaintiin, on nykyaikaisessa mallinnuksessa aika keskeisenä tekijänä. Menneisyydestä oppiminen, nykyisyyden määrittely ja tulevan ennustaminen ovat kaikki erittäin tärkeitä tutkimuksen kohteita päätöksentekijöille. Haasteita syntyy mallin monimutkaisuudesta johtuen maantieteellisten muuttujien kohdalla ympäristön heterogeenisyydestä. Tällöin moniulotteisen datan käsittely ja visualisointi muodostuu merkittäväksi tekijäksi, kun mallia halutaan hyödyntää.

Ympäristön mallintaminen ja paikkatietoala on kokenut lukuisia edistyksiä. Datan keräämisen osalta satelliitti- ja laser-pohjaiset kaukokartoituksen tekniikat kykenevät tuottamaan korkearesoluutioista ja tarkkaa dataa ympäristöstä jopa pitkien matkojen päästä. Tästä on hyötyä, kun tahdotaan kartoittaa kohdetta, jota ei voida lähestyä turvallisesti, kuten metsäpalot tai luonnonkatastrofit. (Karafyllidis et al. 1996) Suurempi määrä dataa vaatii myös kehittyneemmät menetelmät datan käsittelyyn ja jalostamiseen. Geoinformaatioon liittyvien tutkimuksien myötä on kyetty yhdistämään useita tutkimusmenetelmiä osaksi samaa systeemiä.

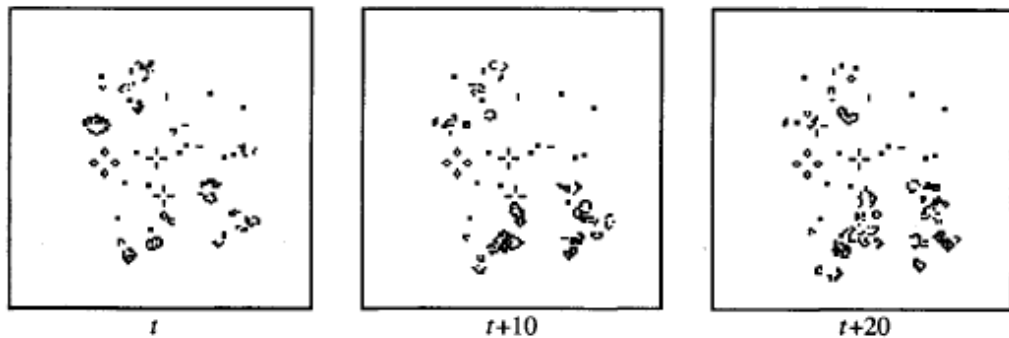
Geosimulointi on kattotermi tavalle mallintaa ympäristöä käyttäen hyväksi geoinformaatiotieteiden ja tietojenkäsittelytieteiden nykyaikaisia menetelmiä. Geosimulointi on yleisesti yhdistetty kaupunkimallinnukseen, mutta sille on myös muita sovelluksia. Verrattuna muihin ympäristön simulointimalleihin, geosimulointi pyrkii ymmärtämään spatiaalisten kohteiden ja objektien muodostamaa systeemiä ja sen toimintatapoja käyttäytymisen pohjalta. Objektikeskeisyys on geosimulointi-mallin keskeinen piirre. Toinen merkittävä piirre on objektien skaala spatiaalisessa merkityksessä. Objekteilla on myös dynaamista liikettä. (Benenson & Torrens 2004)

Johtuen geosimulaatio-mallien korkearesoluutioisesta datasta sekä raskaista prosesseista, on simulointimenetelminä käytettävä tekniikoita, jotka mahdollistavat oikeaa elämää kuvaavien realististen mallien luomisen. Erityinen vaatimus geosimuloinnilla on

toimintansa kannalta objektien käyttäytymisen mallinnus. Tämän johdosta tekniikoina käytetään tyypillisesti automaatteja (automata) mallintamaan objektien realistista käyttäytymistä. (Benenson & Torrens 2004)

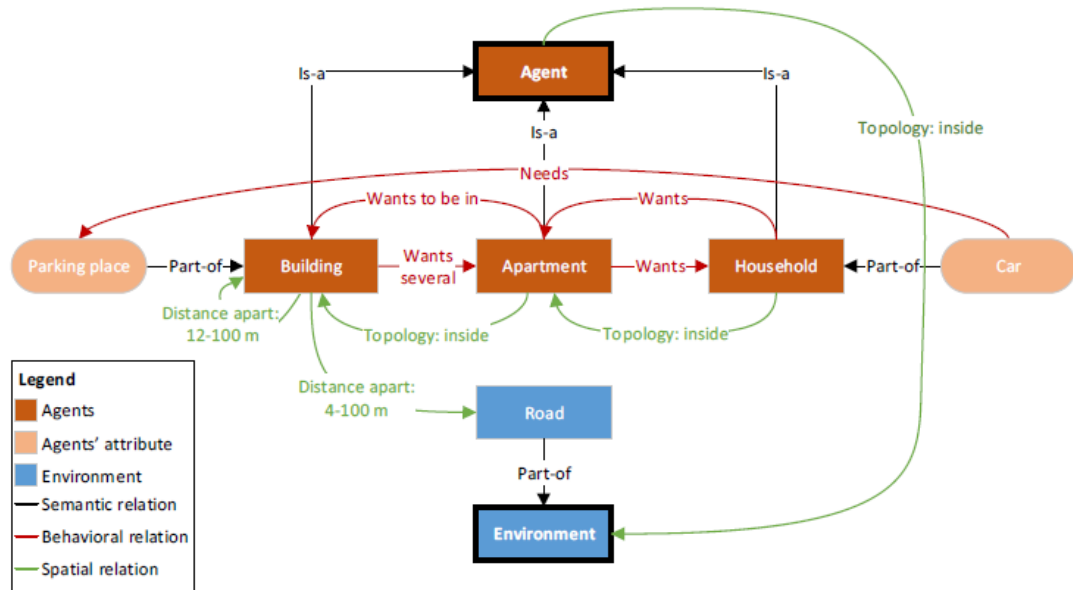
Automaatti on prosessoiva, itseään monistava olio, jonka avulla kyetään yksinkertaisilla säännöillä ja funktioilla toteuttamaan monimutkaisia käyttäytymistapoja. Automaatteja käytetään ympäristön mallintamisessa geoinformaatiotieteissä. Automaateista kaksi on erityisen tärkeitä geosimuloinnille: soluautomaatit (cellular automata), ja agenttipohjaiset systeemit tai moni-agentti systeemit (multi-agent systems).

Soluautomaatit ovat perusmuodossaan sopivia kuvastamaan erilaisia maantieteellisiä alueita ja ympäristöjä. Soluautomaatti koostuu joukosta soluja jaettuna tasaiseksi ruudukoksi ja täten on lähellä usein paikkatietoalalla käytettyä rasterimuotoa. Tärkeä vahvuus soluautomaateille on niiden luontainen spatiaalisuus, koska yhden solun tila riippuu sitä ympäröivistä soluista. Kuvassa 1 on esimerkki yksinkertaisesta soluautomaatista Game of Life, joka tuottaa yksinkertaisilla säännöillä monimuotoisia hahmoja (pattern) eri ajanhetkillä. (Benenson & Torrens 2004)



**Kuva 1** Kuva Game Of Life-soluautomaatista. Mustat solut kuvastavat eläviä soluja ja valkoiset solut kuolleita soluja. Yksinkertaisilla säännöillä voidaan tuottaa monimutkaisia kuvioita. (Chopard & Droz 1998)

Agentit toimivat kuin soluautomaatit, mutta niitä ei ole sidottu yhteen paikkaan ja niiden käyttäytymiseen voidaan antaa erilaisia ehtoja kuten olio-ohjelmoinnissa. Lisäksi, toisin kuin soluautomaatit, agenttien naapurustot (neighborhood) eivät ole yleensä samankaltaisia keskenään, vaan kukin agentti voi ottaa epäsäännöllisen, uniikin muodon. Kuvassa 2 on konseptuaalinen malli usean agentin suhteista toisiinsa ja ympäristöön, jossa toimivat. Agenteilla on erilaisia relaatioita toisiinsa, kuten spatiaalisia, semanttisia tai käyttäytymiseen pohjautuvia. Agenteihin liittyy myös attribuutteja, jotka lisäävät agenteille erityisominaisuuksia. Tällöin agentit ympäristön kanssa muodostavat yhtenäisen systeemin, jossa olevat osatekijät ovat vaikutuksessa toisiinsa.



Kuva 2 Kaavio eri agenttien suhteista toisiinsa ja ympäristöön. (Andreas Hall 2016)

Geosimulointi on kehittynyt pitkälti tietotekniikan, matematiikan ja tilastotieteen edistymisien myötä siihen pisteeseen, että sitä voidaan pitää varteenotettavana simulointimuotona kaupunkimalleille. Aikaisempina haasteina oli mallien liiallinen yleistäminen ja yksityiskohtien puute. Laajempialaisella datalla kyetään mallintamaan seikkaperäisiä yksityiskohtia eri kohteiden käyttäytymisessä ja täten selittämään paremmin ilmiöitä. GIS antaa myös työkaluja mallin hyödyntämiseen useissa erilaisissa analyyseissä ja analyyseistä saadun datan takaisinsyöttämiseen malliin. (Altartouri 2015)

Viimeaikaiset tulokset geosimuloinnin kehittämisessä on nostanut sen varteenotettavaksi tekniikaksi kaupunkimallinnuksessa, mutta myös muilla paikkatiedon aloilla on alettu hyödyntämään geosimuloinnin tarjoamia tekniikoita. Ongelma monissa malleissa on ollut niiden käytännön sovellusten puute, sillä mallit ovat olleet puhtaan abstraktisia. (Benenson & Torrens 2004) Varsinkin agenttien kohdalla on huomattu, että lisää tutkimusta ja työtä vaaditaan, jotta voidaan kehittää käytäntöön soveltuvia menetelmiä.

### 1.1.1 Kaupunkimallit pohjana

Kun puhutaan kaupunkimalleista, mainitaan usein myös systeemiteoria. Kyseinen teoria sai alkunsa 1960-luvun aikana, kun useat eri alojen tiedemiehet ja tutkijat alkoivat hahmottamaan yhteneväisyyksiä lukuisten eri alojen systeemien kesken. Idea siitä, että luonnossa ilmentyviä systeemeitä kyettäisiin mallintamaan, nousi tämän jälkeen tärkeäksi teemaksi.

Systeemiteorian kehityksessä tärkeä suunnanmuutos tapahtui kompleksisten systeemien löytämisen myötä. 60-luvun lopussa alettiin huomaamaan, että kehitetty systeemiteoria ei kyennyt ottamaan huomioon yksinkertaisten, tasapainoiksi luonnehdittujen systeemien kokemia muutoksia. Siihen saakka oli uskottu, että systeemi ottaa vastaan ulkoisista lähteistä tulevia muutoksia, mukautuu niihin ja palaa tasapainoiseen tilaansa ajan myötä. Nämä yksinkertaiset systeemit oli koettu autonomisiksi ja lineaarisiksi, mutta ulkoisten muutosten myötä ne pystyivät kokemaan evoluution, entropian tai kvalitatiivisen muutoksen mm. niiden rakenteessa. Muutosten myötä huomattiin, että

systemit voivatkin olla avoimia ja epälineaarisia, kuten esimerkiksi kaupunki on. (Benenson & Torrens 2004)

Kompleksisuus käsitteenä liittyy emergenssiin eli ilmiön syntymiseen. Yksinkertaiset säännöt ja lait liitettynä joukkoon alkioita voivat toteuttaa monimuotoisia ilmiöitä. Vaikka syntyneillä systeemeillä ei ole näennäisesti struktuuria ja ne voivat vaikuttaa jopa kaoottisilta, on kompleksisissa systeemeissä silti toistuvaa tapahtumaa sekä jäsenneltyjä joukkoja. Kompleksisissa systeemeissä on tärkeänä piirteenä myös keskusjohdoton toteutus, eli yksi tietty taho muita korkeammalla hierarkiassa ei päättä, mitä pienemmät komponentit tekevät. (Torrens 2000)

Tieteen tutkijoita erityisesti maantieteen ja sosiaalisten tieteiden alalla oli kauan haitannut ongelma dikotomiasta liittyen yksittäisiin kohteisiin sekä ryhmittymiin. Esimerkiksi yksittäisen talouden toimet paikallisella tasolla ja sen vaikutukset globaalilla tasolla tuottavat ristiriidan, kun tutkitaan prosesseja sekä ilmiöitä. Tämän lisäksi riittämättömän datan ekstrapolointi voi osoittautua ongelmalliseksi, kun ei tunneta kohteiden välisiä interaktioita tarpeeksi tarkasti. (Torrens 2000) Kompleksisuus aiheuttaa siis sen, että koko kaupunkia mallintaessa voidaan huomata ilmiöitä, jotka eivät taasen sellaisinaan päde paikallisella tasolla pieniin talouksiin. (Benenson & Torrens 2004)

Kaupunkisuunnittelussa kyseinen eri tasojen vaikuttavuus kehityksessä ilmenee kun tarkastellaan rakentamista. Makrotasolla kaupunkia kehitetään suuren joukon toimesta, joka vastaa rakentamisen kokonaisvaltaisesta suunnittelusta. Kyseinen joukko voi ottaa vaikutteita globaalilla tasolla jopa muista kaupungeista ja suunnitteluun liittyy yleensä virallinen strategia. Paikallisemmalla tasolla rakentamisen hoitaa joukko yksilöitä, jotka eivät välttämättä kykene havainnoimaan globaalin tason strategiaa rakentamisen suunnittelussa. Paikallisella tasolla pyritään myös usein maksimoimaan käytetyn rakennustilan tehokkuus. Tällöin näiden kahden tason välillä on ero siinä, mitkä tahot vaikuttavat niiden päätöksiin. Kommunikaatiolla, makrotason päätöksillä sekä ohjeistuksilla voidaan vaikuttaa paikallistason päätöksiin. (Ericsson & Lloyd-Jones 1997) Kaupunkimallinnuksessa yksi ongelma tutkijoiden kesken onkin ollut se, että kaupunkia yritetään mallintamaan yhtenä kokonaisuutena globaalisti, eikä paikallisia tekijöitä ole välttämättä otettu huomioon.

Kaupunkien kehittymisen tutkimuksessa on huomattu, että luontainen ja vähän ohjattu kasvu tuottaa näennäisesti kaoottisia, mutta kokonaisuutena toimivia kaupunkialueita. Tarkastelemalla vain paikkaan sidottuja sosiologisia rakenteita kyetään ymmärtämään vain osa siitä mitkä tekijät vaikuttavat kaupunkialueiden muodostumiseen. Puhutaankin kaupunkirakentamisen ja –suunnittelun muodosta. (Ericsson & Lloyd-Jones 1997) Christopher Alexander teki vuosina 1964 ja 1979 tutkimuksia liittyen urbaanisen kasvun muodoista ja kehitti urbaanin suunnittelun peruseräiteitä. Useat tutkijat jakavat kaupungit suunniteltuihin sekä suunnittelemattomiin alueisiin. Suunnittelemattomista alueista käytetään myös nimitystä orgaaniset alueet niiden luontaisen, epävirallisen vaikutuksen alaisuudessa tapahtuvan syntymisen vuoksi. Kehittyvien maiden nopean kasvun ja kaupunkien kohdistuvan laajenemisen vuoksi kaupunkien suunnittelun tutkimus on erittäin tärkeä aihealue. Perimmäinen ongelma kyseisissä kohteissa on, että suurin osa kasvusta sijoittuu suunnittelemattomille alueille. Kun suunnittelijat kykenevät ennustamaan kasvun muodon ja suuruuden, voidaan tehokkaammin suunnitella kaupungin infrastruktuuria ja palveluita. (Ericsson & Lloyd-Jones 1997)

Kaupunkien osa-alueiden tarkastelussa huomataan, että alueiden muoto vaihtelee riippuen siitä, millä tasolla aluetta tarkastellaan. Kuten aiemmin on todettu, vaikutukset ilmenevät eri tasoilla erilaisina, ja sama pätee myös alueiden muodostumisessa. Lähiöalueella voidaan huomata olevan hyvin geometrinen tapa rakentaa asuintaloja ruudukkomaisesti, mutta kun koko aluetta tarkastellaan, niin alueen muoto voi olla hyvin rikkonainen. Tämän vuoksi kaupunkirakentamisen mallinnuksessa ja suunnittelussa on erittäin tärkeää ottaa huomioon se skaala ja alue, johon keskitytään. Tästä johtuen myös erottelu suunniteltuun ja suunnittelemattomiin alueisiin ei ole kaikkien tutkijoiden mielestä mieluisa tapa erotella. Lisäksi kun tutustutaan kaupunkien kehitykseen eri aikaväleillä, voidaan huomata trendejä rakentamisen suunnittelussa. Vaikka voidaan helposti kuvitella eri aikakausina kehitettyjen alueiden olevan erillisiä yksiköitä, on niillä siitä huolimatta relaatioita ja vaikutuksia toisiinsa. Alueet eivät siis ole autonomisia yksiköitä, vaan niillä on vaikutusta toisiinsa kumpaankin suuntaan riippumatta rakentamisen ja kehittämisen aikakaudesta tai niiden muodosta. (Ericsson & Lloyd-Jones 1997)

## **1.2 Tutkimusongelma ja työn raja**

Tässä työssä tutkitaan, miten geosimulointiin liittyvät uudet mallit käyttävät hyväkseen soluautomaatteja ja agenttipohjaisia systeemeitä. Työssä pyritään tarkastelemaan kriittisesti mallien käyttämien menetelmien saavuttamia tuloksia ja niiden soveltuvuutta geosimulointiin. Ymmärryksen lisäämiseksi tehdään kirjallisuustutkimus geosimuloinnista sekä siinä käytettävistä menetelmistä, soluautomaateista sekä agenttipohjaisista systeemeistä.

Geosimuloinnissa on tapahtunut merkittävää edistymistä mallien tarkkuuden ja kattavuuden suhteen. Aikaisemmin malleja pidettiin liian abstrakteina eivätkä ne vastanneet oikeaa todellisuutta tarpeeksi. Kehittyneet mallit ovat lähempänä todellisuutta ja niitä voidaan pitää vartenotettavina. Kun tunnistetaan geosimuloinnin vaatimukset hyvälle mallille, voidaan tehdä arvioita siitä, miten käytetyt tekniikat toteuttavat mallin vaatimuksia. Erityisesti voidaan kiinnittää huomiota siihen, miten tekniikoita käytetään verrattuna muihin kehitettyihin malleihin ja kuinka tekniikoiden kehitystyö on edennyt. Tekniikoiden samankaltaiset käyttötavat sekä vakiintuneiden menetelmien tunnistaminen tutkimuksissa mahdollistaa sen, että voidaan muodostaa kuva siitä, miten tiedeyhteisö käsittää geosimuloinnin.

Geosimuloinnille tärkeä tekniikka on automaattit. Sekä soluautomaattit että agenttipohjaiset systeemit perustuvat samaan ideaan automaateista, mutta niiden käyttötilanteet ovat erilaiset johtuen erilaisista heikkouksista ja vahvuuksista. Geosimulaatiossa hyödynnetään kumpaakin tekniikkaa riippuen mallinnuksen painopisteestä. Kummastakin tekniikasta on ymmärrettävä niiden kehityshistoria sekä nykyiset vakiintuneet perusmuodot. Kummankin tekniikan muotoutuminen nykyaikaan antaa ymmärrystä siitä, mitkä ovat olleet ongelmallisia piirteitä, joita tutkijat ovat pyrkinet kehittämään, sekä mikä on toiminut motivaationa kehitystyön jatkamiselle. Tällä tavoin ymmärtämällä kyetään muodostamaan eheämpi kuva nykyisen kehitystyön tilasta.

Työssä rajataan geosimuloinnin tutkimus koskemaan pääasiassa kaupunkimallintamista. Kaupunkimallintaminen on geosimuloinnin yleisin sovellusalue ja kyseisestä alueesta on saatavilla suurin määrä kirjallisuutta ja tutkimusaineistoa.

Tutkimuskysymykset työssä ovat:

- Miten soluautomaatteja ja agenttipohjaisia systeemeitä käytetään geosimuloinnissa?
- Mitä vaatimuksia hyvälle geosimulointimallille asetetaan?
- Miten nykyaikaiset geosimulointimallit ottavat huomioon hyvän mallin vaatimukset ja haasteet?

### **1.3 Työssä käytetyt aineistot ja rakenne**

Työssä käytetty aineisto koostuu alan kirjallisuudesta, tutkimuksista sekä artikkeleista. Pääasiallinen lähde itse geosimuloinnille on Benensonin ja Torrensin (2004) kirja *Geosimulation*. Teoksessa käydään lävitse geosimuloinnin ja kaupunkimallinnuksen historiaa sekä esitellään malli, jossa yhdistyvät sekä soluautomaatit että agenttipohjaiset systeemit yhdeksi kokonaisuudeksi. Kyseisen kirjan lisäksi tutustutaan nykyaikaisiin tutkimuksiin sekä soluautomaateista että agenttipohjaisista systeemeistä. Tutkimukset selventävät käytettyjä tekniikoita sekä nykyaikaisia sovelluksia.

Johtuen agenttien ja automaattien kehityksen historiasta, on työssä käytetty lähteenä myös muuta kuin paikkatietoalan kirjallisuutta. Erityisesti sosiologian, keinoälyn sekä systeemitekniikan alan kirjallisuutta on käytetty lähteenä, kun on haluttu tutustua kyseisten tekniikoiden kehitykseen sekä perusmalleihin.

Työ alkaa kirjallisuustutkimuksella, joka kattaa luvut 2-5. Luku 2 aloitetaan käsittelemällä lyhyesti käsitteet spatiaalisuus ja mallintaminen. Tämän jälkeen luvussa 3 tutustutaan soluautomaatteihin. Luvussa käsitellään peruskäsitteet sekä edistyneemmät soluautomaatteihin liittyvät tekniikat. Luvussa 4 käsitellään taasen agenttipohjaisia systeemeitä. Luvussa kerrotaan agenttien käytöstä yleisesti sekä miten niitä voidaan hyödyntää geosimuloinnissa. Tämän jälkeen luvussa 5 esitellään geosimulointi. Luvussa käydään lävitse geosimuloinnin perusasiat sekä erityisiä ominaisuuksia. Lisäksi erityiset geosimuloinnin haasteet käydään lävitse liittyen soluautomaatteihin ja agentteihin. Tätä seuraavissa luvuissa käsittelemään kahta ajankohtaista geosimulointiin liittyvää tutkimusta. Töiden lähtökohdat esitellään ja tämän jälkeen kattavammin käytetyt tekniikat. Arviointia varten koostetaan tekniikoiden tunnettuihin haasteisiin liittyvät kriteerit. Kummankin työn esitellyn jälkeen tehdään kriittinen tarkastelu siitä, miten eri tekniikoita on käytetty. Lopuksi suoritetaan pohdintaa saatujen tuloksien mukaan. Päätännössä kerrataan työssä esitellyt tutkimusongelmat ja miten saadut tulokset työssä vastaavat ongelmiin.

## 2 Spatiaalisuus ja mallintaminen

### 2.1 Yleistä

#### 2.1.1 Spatiaalisuus

Spatiaalisuus on tilassa olevien kohteiden ja ilmiöiden suhteellinen sijainti toisiinsa. Kohteille voidaan määritellä kahdenlaisia spatiaalisia suhteita: topologisia ja metrisiä. Metrinen suhde käsittää kohteiden välisen fyysisen, mitattavissa olevan etäisyyden. Topologinen suhde kohteiden välillä taas kuvastaa niiden topologiaa, kuten vierekkäisyyttä, yhteneväisyyttä tai päällekkäisyyttä. (Gorodetz kaya 2013)

#### 2.1.2 Paikkatiedon erityisominaisuudet

Paikkatiedosta tekee erityistä muutamia ominaisuuksia. Ensimmäinen lähes kaikelle oikeasta maailmasta kerätylle tiedolle kyetään määrittelemään sijainti. Tällöin paikkatieto on hyvin laaja-alaista. Tämä myös tarkoittaa sitä, että kerätty tieto ja sen käyttötilanteet voivat olla hyvin käytännönläheisiä. Toiseksi paikkatietoa sanotaan myös spatiaaliseksi tiedoksi, koska tieto on sidottu tiettyyn tilaan eikä se välttämättä tarkoita maantieteellistä sijaintia. Terminologisessa mielessä on käytetty myös geospatiaalista tietoa kuvaamaan vain maapallon ympäristöön ja tilaan liittyvää spatiaalista tietoa, mutta yleensä puhutaan vain paikkatiedosta yleisellä tasolla. (Koistinen 2011)

Longley et al. (2011) määritteli paikkatiedolle ominaisuuksia, jotka tekevät siitä erityistä. Tällaisia ominaisuuksia ovat paikkatiedon moniulotteisuus, paikkatiedon datamäärien suuruus, paikkatietoa joudutaan analysoimaan erityisillä analysointimetoodeilla, koska perinteiset tilastolliset analysointimetodit eivät sovellu autokorrelaation vuoksi ja paikkatiedon esittäminen sekä visualisointi.

Kyseisiä ominaisuuksia ei tässä työssä käydä lävitse tarkemmin, mutta on hyödyllistä pitää mielessä, että kyseiset erikoisominaisuudet asettavat haasteita myös geosimuloinnin kohdalla. Kyseiset ominaisuudet osoittavat myös, kuinka paikkatieto erottuu omaksi tieteenalaksi, ja tarjoaa motivaatiota tutkimuksen tekoon.

### 2.2 Mallintaminen

Paikkatiedon mallintaminen on pelkistetyin tai yleistetyin esityksen luontia jostakin reaali maailman tilasta tai alueesta. Julkisen hallinnon suositus JHS 162 Paikkatietojen mallintaminen tiedonsiirtoa varten (2012) kattaa käsitteet ja periaatteet, jotka liittyvät mallintamiseen. Kun malli kattaa kyseisen alan käsitteet ja niiden väliset suhteet, kutsutaan sitä käsittemalliksi. Paikkatiedon tapauksessa käsittemalli kattaa paikkatiedon kohteet ja niiden väliset riippuvuussuhteet mallinnettavalla alueella. (JHS 162 2012)

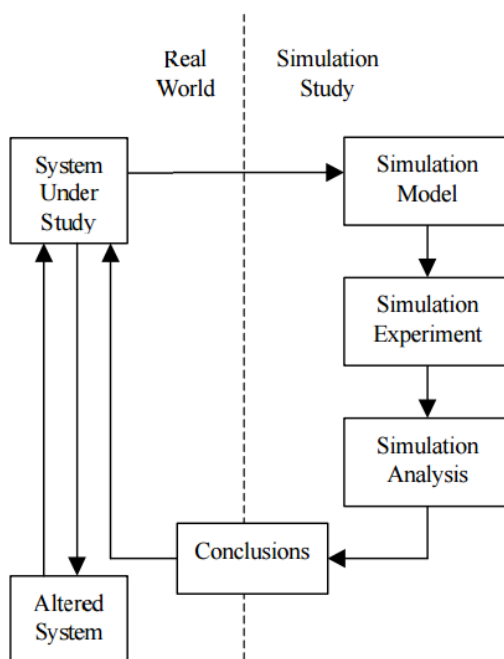
Malli on yksinkertaistettu tai abstrakti kuvaelma todellisesta maailmasta. Mallilla voidaan esittää, selventää, ennustaa tai suunnitella ja se on yksi analysointiprosessin tekniikka. Mallissa olevien elementtien ja objektien väliset riippuvuudet ja niiden käyttäytyminen riippuvat säännöistä, joita oikeasta maailmasta pyritään mallintamaan. Karkeasti erotellen mallit pyrkivät simuloimaan joko yksittäisten kohteiden tai suurempien yksilöistä koostetun joukon yleistysten kautta oikean maailman ilmiöitä. Myös käytössä oleva resoluutio ja tarkkuus rajoittavat ilmiöiden ja käyttäytymisten mallinnusta. (Koistinen 2011)

Staattiset ja dynaamiset mallit liittyvät mallinnuksen temporaaliseen ulottuvuuteen eli aikaan. Jos mallin syöte ja ulostulo on samassa ajanhetkessä, sanotaan mallia staattiseksi. Jos taas ulostulo on myöhemässä ajanhetkessä, sanotaan mallia dynaamiseksi. Staattisten ja dynaamisten mallien erona on pääasiassa dynaamisten mallien kyky ennustaa tulevaisuuteen. Tällöin dynaamiset mallit ovat hyödyllisiä erityisesti skenaarioiden suunnittelussa. Dynaamisten mallien kehityksessä kompleksiteoria ja kaupunkimallit ovat olleet merkittävässä asemassa. (Gorodetskaya 2013)

Kohdeluokka on nimettyä kohdetta, abstraktiota tai ilmiötä koskeva mallinnus. Kohdeluokka on hierarkkisesti kattoterminä ja luokan sisäistä yksittäistä kohdetta kutsutaan kohdeluokan ilmentymäksi. Tällöin esimerkiksi kohdeluokka ”Rakennus” kattaa kaikki rakennukseksi luokiteltavat kohteet alueella ja yksittäinen rakennus on kohdeluokan ”Rakennus” ilmentymä. Kohdeluokan sisäisten kohteiden ominaisuuksia kutsutaan kohdetyypeiksi. Kohdetyypillä on nimi sekä ominaisuusarvo. Jokaisen kohdeluokan ilmentymän on sisällettävä kyseiset kohdetyypit, jotta ne voisivat kuulua yhteiseen kohdeluokkaan. Tällöin esimerkiksi aiemmin mainittu ”Rakennus” – kohdeluokka voisi sisältää kohdetyyppitietona rakennuksen iän, talonumeron, omistusoikeuden ja kerrosmäärän. Paikkatiedon yhteydessä on tärkeitä kyetä yksilöimään kohteita, jolloin talonumero voisi toimia yksittäisen rakennuksen uniikkina tunnuksena. Pistemäisten kohteiden kohdalla sijainti voidaan ilmoittaa yksinkertaisina koordinaattiarvoina, mutta monimutkaisemmat kohteet, kuten tiet, tarvitsevat geometriatiedon, jolla ne voidaan määrittellä tarpeellisella tarkkuudella. (Koistinen 2011)

### 2.3 Simulointi

Simulointi on kehitetyn mallin käyttöä oikean maailman systeemistä. Mallille voidaan tehdä kalibrointeja, jotta malli tuottaa erilaisia tuloksia, jotka ovat joko lähempänä oikean maailman vastinetta systeemistä tai kokeellisia tuloksia. Kokeellisilla toimenpiteillä opitaan systeemin toimintaperiaatteita ja miten systeemin sisällä olevat alisysteemit toimivat keskenään. Tiedon kerääminen on tällöin yksi simuloinnin mahdollinen tavoite. (Maria 1997)



Kuva 3 Kaaviokuva systeemin mallintamisesta ja simuloinnista. (Maria 1997)

Geosimulointi on yksi simulointityyppi, joka keskittyy spatiaalisten ilmiöiden simuloimiseen. Kaupungit ja niihin liittyvät ilmiöt ovat yksi geosimuloinnin tyypillisin sovelluskohde. Perinteisten simulointimallien ja geosimulointimallien erona on, että perinteiset mallit rakennetaan yleensä top down -periaatteella, jossa lähtökohta on kokonaisuus, jota puretaan pienemmiksi palasiksi. Kun edetään pienempiin yksityiskohtiin, kyetään ymmärtämään systeemiä tällöin paremmin. Geosimulointi tehdään taasen bottom up -periaatteella, jossa muodostetaan ensin systeemin peruselementit yksityiskohtaisesti. Kyseiset peruselementit muodostavat monimutkaisempia, korkeamman tason systeemeitä ja systeemeissä toimivia ilmiöitä. (Geosimulation 2017)

## 3 Soluautomaatti

### 3.1 Yleistä

#### 3.1.1 Taustaa

Soluautomaatit kehitettiin alun perin 1940-luvulla John Von Neumannin tekemien tutkimuksien myötä. Von Neumannin alkuperäinen ajatus soluautomaateille oli kehittää ihmisen aivoja kuvastava systeemi, jossa sekä data että prosessorit olisivat samalla tasolla. Von Neumann alkoi kehittää systeemiä, joka kykeni monistautumaan itsensä kanssa ilman biologiatieteiden prosesseja. (Choppard & Droz 1998) Motivaationa oli kehittää tapa ratkaista monimutkaisia ongelmia, jotka siihen aikaan vaativat suurta laskentatehoa (Yassemi et al. 2007). Von Neumannin luomassa soluautomaatin perusmallissa solulla voi olla rajallinen joukko tiloja, jotka määräytyvät solujen yhteisillä säännöillä. Jokaisen solun tila muuttuu samaan aikaan funktion ja naapuruston perusteella, jolloin jokainen solu on kompleksisuudeltaan samanlainen kuin muut solut. Tämä oli merkittävää, koska siihen saakka oli uskottu, että koneet kykenisivät tuottamaan vain itseään yksinkertaisempia objekteja. (Chopard & Droz 1998)

Yleisen systeemiteorian tultua suosituksi 60- ja 70-luvuilla, kiinnostus soluautomaatteihin lisääntyi. Soluautomaatit kykenivät yksinkertaisilla säännöillä mallintamaan ja järjestämään komplekseja systeemeitä. John Horton Conwayn Game of Life-malli esitteli, miten monimutkaisiin systeemeihin soluautomaatti kykenisi. Hyvin yksinkertaisilla säännöillä kyettiin tuottamaan monimutkaisia spatiaalisia hahmoja. Siihen saakka spatiaalisten systeemien mallinnuksessa oli pääosin käytetty jatkuvia differentiaaliyhtälöitä. (Benenson & Torrens 2004)

Von Neumannin jälkeen soluautomaattien kehitystyössä suuria edistysaskeleita otti Stephen Wolfram. Laskennallisesta universaalisuudesta luovuttiin ideatasolla. Alunperin soluautomaattien suunniteltiin kykenevän suorittamaan kaikki mahdolliset funktiot siinä tietyssä laskennallisuuden ryhmässä mihin ne oli suunniteltu. Yksinkertaistamalla soluautomaatteja ja niiden alkuperäistä konseptia, kyettiin suuntaamaan kehitys kohti laajempaa sovelluksien joukkoa. S. Wolframin tekemän kehitystyön johdosta kyettiin huomaamaan soluautomaattien soveltuvuus monimutkaisten matemaattisten funktioiden laskennassa. Boolean-arvojen implementoinnin avulla kyettiin lisäksi tuottamaan tarkempia tuloksia kuin numeroarvoja käyttämällä. (Chopard & Droz 1998) Tämän johdosta soluautomaatteja suositaan matemaattisesti raskaissa simulointiprosesseissa. (Avolio et al. 2012)

Ennen Wolframin kehitystyötä soluautomaateille oli asetettu tietty määrä eri tiloja, joihin solu voisi muuntua. Ajan saatossa tutkijat pyrkivät yksinkertaistamaan soluautomaattia vähentämällä mahdollisten tilojen lukumäärää. Wolfram ehdotti omassa mallissaan, että solulla voisi olla vain totuusarvoisia tiloja. Tämä yksinkertaisti mallia, mutta mahdollisti silti kompleksisten systeemien mallintamisen. Käytännössä totuusarvojen hyödyntäminen solun tilan kuvaamisessa tarkoittaa sitä, että solun tila vaihtelee arvojen 0 ja 1 välillä. (Chopard & Droz 1998)

### 3.1.2 Soluautomaatin yleiset ominaisuudet

Soluautomaatti koostuu säännöllisestä yksi- tai kaksiulotteisesta ruudukosta. Jokaisella ruudulla tai solulla on tietty arvo, joka määräytyy lokaalisti. Arvo kuvastaa mallinnettavan aineellista olotilaa solussa. Jokaisella solulla on naapurusto, joka sisältää itse solun sekä tietyn määrän sen lähimpiä muita soluja. Solun sisäinen arvo päivittyy kaikissa soluautomaatin soluissa yhtä aikaa kunakin ajanhetkenä. (Karafyllidis et al. 1996)

Soluautomaatin luontainen vahvuus liittyy sen topologiaan. Solujen vierekkäisyys ja ilmiön vaikutuksen ilmentyminen lokaalisti soluissa soveltuu spatiaalisten suhteiden esittämiseen mallissa. (Dahal & Chow 2014) Monimutkaista ilmiötä kyetään mallintamaan soluautomaatin sääntöjen sisällä ja tietokoneet kykenevät simuloimaan samaa mallinnusperiaatetta. Useat muut simulointimetodit, kuten osittaisdifferentiaaliyhtälöihin perustuvat simulointimetodit, ovat paljon monipuolisempia ja tarkempia mallinnuksessa, mutta niiden heikkoutena on varsinkin jatkuvilla pinnoilla syntyvä muuttujien arvojen teoriassa rajoittamaton määrä. Kyseisestä ominaisuudesta on hyötyä, kun aikaa mallinnetaan jatkuvana, mutta usein aika ja tila halutaan pitää diskreettina eli aika etenee askeleittain. (Benenson & Torrens 2004)

Soluautomaatin luontainen kyky kuvata spatiaalisuutta vastaa hyvin myös Waldo Toblerin maantieteen ensimmäistä lakia: lähellä olevat asiat ovat samanlaisempia kuin kaukana toisistaan olevat. (Benenson & Torrens 2004)

Soluautomaatilla on viisi tärkeätä ominaisuutta jotka määrittävät sen toiminnan ja prosessin:

1. Spatiaalisten ulottuvuuksien määrä  $n$ . Tyypillisin soluautomaatti koostuu kahdesta ulottuvuudesta, mutta myös yksiulotteisia soluautomaatteja on tutkittu. Yksi- ja kaksiulotteisella soluautomaatilla on runsaasti samankaltaisuuksia niiden muodostamien kuviodien ja joukkojen kesken.
2. Jokaisen solun sivun leveys. Kuten rastereiden muodostuksessa, myös soluautomaateilla yksittäisen solun sivun leveys vaikuttaa merkittävästi siihen, mitä ollaan tutkimassa ja millaiselle aineistolle prosessi suoritetaan.
3. Solun naapuruston leveys. Käsittää sen kuinka monta muuta solua muodostaa solun naapuruston. Naapurustoa käytetään yhdessä lokaalin funktion kanssa määrittämään solun tilaa kullakin ajanhetkellä. Solu itsessään lukeutuu myös naapurustoonsa.
4. Solun tila. Tällä ominaisuudella kyetään kuvaamaan lähes mitä tahansa attribuuttiarvoa tai totuusarvoa, jota tarvitaan analyysissa.
5. Soluautomaatin lokaali sääntö, jonka perusteella solun tila määräytyy naapuruston mukaan. Yleensä kuvataan funktiolla  $f$ .

(Benenson & Torrens 2004)

Näiden viiden kohdan lisäksi soluautomaatille olennaista on temporaalisen tilan eli ajan määrittäminen. Solujen tilat vaihtuvat kun aika muuttuu ja kuten aiemmin esiteltiin, jokaisen solun tila muuttuu ajan edetessä samanaikaisesti. Tällä tavoin saadaan yhdistettyä malliin tärkeä määre, koska ajan mukaan tapahtuva muutos on tärkeä työkalu, kun halutaan ennustaa tulevaa. Johtuen soluautomaatin lähtökohdista, eli alkuasetelman edistymisen toistamalla asetettua sääntöä, on soluautomaatin simuloitujen mallien ajallisesti muotoutuvia. (Li & Yeh 1998)

### 3.1.3 Formaali kuvaus soluautomaateille

Soluautomaatit kuvaillaan formaalisti samalla tavalla nykyisin kuin von Neumann kuvaili niitä vuonna 1951. Soluautomaatti on yksi- tai kaksiulotteinen ruudukko, joka sisältää automaatin solut. Jokainen solu on samanlainen toistensa kanssa. Jokainen automaatti  $A$  määrittää joukon tiloja  $S = \{S_1, S_2, S_3, \dots, S_n\}$  ja joukon tilasääntöjä  $T$ :

$$A \sim (S, T) \quad (1)$$

Tilasääntö määrittää automaatin tilan  $S_{t+1}$  ajanhetkellä  $t + 1$  riippuen automaatin tilasta hetkellä  $t$  sekä syötteenä  $I$  saadusta ulkoisesta datasta:

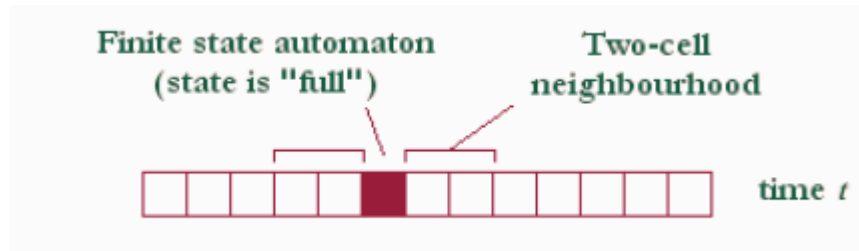
$$T: (S_t, I_t) \rightarrow S_{t+1} \quad (2)$$

Soluautomaatti ja sen solukko määritellään soluautomaatiksi, kun syötteet määräytyvät naapuruston solujen tiloilla. Naapurustoon lukeutuu ne solut, jotka määritellään automaatin ruudukon perusteella. Tällöin automaatti  $A$  voidaan määritellä:

$$A \sim (S, T, R) \quad (3)$$

missä  $R$  on automaatin naapurusto, josta haetaan syöte  $I$ , ja syötettä  $I$  käytetään yhdessä tilasäännön  $T$  kanssa.

Naapurusto  $R$  on yksiulotteisessa soluautomaatissa tyypillisesti solun kummallakin puolella olevat solut, mutta myös laajempia naapurustoja voidaan esittää.



Kuva 4 Yksiulotteinen soluautomaatti, jolla on kahden solun kokoinen naapurusto (Torrens 2000)

Kaksiulotteiset soluautomaatit koostuvat tyypillisesti neliöruudukosta. Naapurusto näissä on tyypillisesti tarkasteltavan solun neljä tai kahdeksan viereistä solua. Naapurustot on nimetty von Neumann ja Mooren naapurustoiksi:



Kuva 5 Kaksi naapurustotyyppiä, (a) von Neumann ja (b) Moore (Wolfram 1985)

Mahdollisten tilojen lukumäärä riippuu käytössä olevasta soluautomaatista. Von Neumann määritteli tilojen lukumääräksi 29. (Wolfram 1983) Yleisesti voidaan laskea mahdollisten tilasääntöjen määrä kaavalla  $N^{N^K}$ , jossa  $N$  on tilojen lukumäärä ja  $K$  on naapuruston solujen lukumäärä. Tällä tavoin yksiulotteisessa, kahden naapurin soluautomaatissa, jossa soluilla voi olla vain kaksi mahdollista tilaa, saadaan mahdollisten tilasääntöjen lukumääräksi  $2^{2^3} = 256$ . Tällä tavoin voidaan huomata, että naapuruston kokoa tai solun tilojen lukumäärää kasvattamalla voidaan nopeasti kasvattaa mahdollisten tilasääntöjen lukumäärää. (Benenson & Torrens 2004)

### 3.1.4 Luokittelu

Stephen Wolfram suoritti joukon kokeita ja tutkimuksia liittyen soluautomaattien mahdollisiin kokoonpanoihin, kun lähtökohdat ja säännöt tilojen muuttumisille ovat yksinkertaiset. Wolfram (1994) ja myös Langton (1992) kuvailivat neljä erilaista luokkaa, joihin soluautomaatit voidaan luokitella niiden muodostamien hahmojen ja käyttäytymisen perusteella:

- Luokka 1; homogeeninen joukko, jossa kaikilla soluilla on sama arvo. Jos soluilla voi olla kaksi mahdollista tilaa (0 ja 1), on loppujoukon kaikilla soluilla tilan arvona 0.
- Luokka 2; toisistaan erottuneita, jaksottaisia rakenteita globaalisti jakautuneena. Muistuttaa rakenteeltaan vertikaalisia raitoja.
- Luokka 3; kaoottinen rakenne sekä lokaalisti että globaalisti. Ei välttämättä täysin satunnaisesti generoitunut, vaan itsejärjestäytynyt.
- Luokka 4; kompleksisia rakenteita. Solut järjestäytyvät mahdollisesti pitkäikäisiksi systeemeiksi.

Luokittelua ei ole tarkoitettu yleispäteväksi metodiksi kunkin soluautomaatin tyyppin kuvailuun, koska kyseiset luokittelut pätevät vain yksinkertaisiin automaatteihin, joiden säännöt käsittävät vain muutaman tapahtuman liittyen solun kehittymiseen ja taantumiseen. Luokista voidaan silti havainnoida soluautomaatin käyttäytymisen arkkityyppejä, jotka voivat päteä myös monimutkaisempiin soluautomaatteihin tai sen alueellisiin osiin. Tutkimusten perusteella on pystytty osoittamaan soluautomaattien kykenevän mallintamaan kompleksisia systeemeitä, koska yksinkertaisilla säännöillä on saatu aikaiseksi biologisia systeemeitä muistuttavia hahmoja. (Wolfram 1994)

## 3.2 Soluautomaattien pääominaisuudet

### 3.2.1 Naapurusto

Naapuruston (neighborhood) määrittely on tärkeä osuus soluautomaatin toimintaa. Alunperin naapurustoksi ehdotettiin neljää pääilmansuuntaa, eli pohjoinen, eteläinen, läntinen ja itäinen solu pääsolun vierellä lukeutuisivat pääsolun naapurustoon. Tätä kutsutaan tekijänsä mukaisesti von Neumann-naapurustoksi. Myöhemmin käytössä on ollut laajempia naapurustoja, joista kahdeksan pääsolun viereistä solua on yleisin, ja kyseistä naapurustoa nimitetään Moore-naapurustoksi. (Wolfram 1984) Nykyisin naapuruston koko ja muoto voivat vaihdella käyttötarkoituksen mukaan. Varsinkin kaupunkiympäristön mallinnuksessa solut voivat käsittää tietyn osan naapurustoa tai kaupunginosaa, jolloin niiden muoto ei ole yhtenevä. Formaalisti ja säännönmukaisesta soluautomaatin muodosta on siis osittain luovuttu, kun mallinnetaan luonnollisia

ympäristöjä. Haasteena onkin ollut realistisen ympäristön mallintaminen samalla pitäen mallin säännöt riittävän yksinkertaisina. (Altartouri 2015)

Naapuruston käsite soluautomaateissa sekä myös paikkatiedossa on kaksijakoinen. Topologisessa merkityksessä vierekkäisyys (adjacency) on yksi perusrelaatio. Objektien tai sijaintien vierekkäisyys toisiinsa voidaan tulkita joko fyysisenä vierekkäisyytenä, joka pystytään mittaamaan, tai sitten vaikuttavana vierekkäisyytenä. (Gorodetzka 2013) Fyysisen vierekkäisyyden mallintaminen soluautomaatissa naapurustolla ei ole ongelmallista ja mallinnuksessa kyetään pitämään suhteellinen yksinkertaistus. Tämänkaltainen mallinnus naapuruston avulla oli perinteisen soluautomaatin idean mukaista. (Benenson & Torrens 2004)

Kun siirrytään tutkimaan spatiaalisen vaikuttavuuden mallintamista naapurustolla, joudutaan naapurustoa miettimään monimutkaisempaan alueena, jossa vaikutukset eivät välttämättä ole spatiaalisesti yhteneväisiä. Tämä johtaa siihen, että naapurusto on pikemminkin koko se alue, johon solulla on vaikutusta. Naapuruston valinnassa vaikuttaa siis erityisesti se, että mitä ilmiötä ollaan mallintamassa ja millaiselle alueelle sillä on vaikutusta. Tämä johtaa siihen, että koko solujako voidaan joutua muokkaamaan sen mukaan mitä ilmiötä tutkitaan. Esimerkiksi kaupunkiympäristön mallinnuksessa eri vetovoimatekijöiden vaikutukset asuinalueille voivat kattaa erisuuruisia alueita, jolloin solujen jako ja niiden naapurustot eivät välttämättä käsitä pelkästään maantieteellisiä naapurustoja. Tällainen ajattelu johtaa kauemmas perinteisestä soluautomaatin ideasta, mutta on lähempänä modernin geoinformaatiojärjestelmien menetelmiä mallintaa vaikutuksia sekä spatiaalisuutta. (Couclelis 1997) Haluttujen tuloksien tarkkuuden kannalta naapurustolla on myös merkitystä, koska suurempi naapurusto merkitsee suurempaa määrää soluja, joita käytetään pääsolun tilan määrittämisessä. Eri soluautomaatit omaavat erilaisen herkkyyden solun naapuruston muodolle ja naapurisolujen määrälle. Onkin ehdotettu, että naapuruston vaikutukset ja herkkyyksien analyysi tulisivat olla osa soluautomaatin mallinmuodostusprosessia. (Dahal & Chow 2014)

Pienemmät, epäsäännöllisen solukon omaavat mallit ovat olleet nykyaikaisen tutkimuksen kohteena. Shen et al. (2009) mukaan kaupunkisuunnittelussa soluautomaatteja on alettu käyttää mallintamaan mikrotasolla kaupungin kasvua ja maan käyttöä. Tyypillisesti soluautomaatteja on ollut käytössä laaja-alaisesti mallintamaan lähes koko kaupunkia sekä siihen liittyviä ilmiöitä, kuten kasvua sekä väestön liikkumista. Pienemmässä mittakaavassa kyetään jakamaan soluautomaatti esimerkiksi naapuruston ja talojen perusteella solukoksi. Erickson ja Lloyd-Jones (1997) jakoi solukon työssään erikseen rakennuksiin ja tiesoluihin. Shen et al. (2009) taas jakoi naapuruston soluihin, joiden erottajana toimi tieverkosto. Pienemmässä mittakaavassa joudutaan myös ottamaan huomioon tarkasteltavan alueen rajoittaminen ja sen merkitys, kun osuus kaupungista erotetaan omaksi itsenäiseksi alueeksi. (Shen et al. 2009) Reuna-alueiden vaikutusta voidaan huomioida erilaisilla puskureilla (Torrens 2000). Toinen ongelma syntyy esimerkiksi Suomen maakuntien tyypillisissä kylissä, joissa välimatkat naapureihin voivat olla pitkiä. Tällöin voidaan käyttää esimerkiksi Voronoi-diagrammia ja Euklidisia etäisyyksiä määrittämään naapurussuhteita (Erickson & Lloyd-Jones 1997).

Kun soluautomaatteja alettiin hyödyntämään ja tutkimaan kaupunkisuunnittelussa tarkemmin, huomattiin pian, että tiukat säännöt naapuruuksien koosta ja muodosta ovat riittämättömiä kuvastamaan esiintyviä dynaamisia suhteita ja ilmiöitä. Myös naapurussuhteiden kuvaaminen vain vierekkäisyydellä on huomattu ongelmalliseksi

tilanteissa, joissa spatiaaliset kohteet muuttuvat ajan myötä ja täten mahdollisesti myös niiden geometria. Tämän vuoksi naapuruuksia kuvataan mieluummin kontekstissaan eikä suoranaisesti vain läheisyydellä tai topologialla. (Marceau & Benenson 2011)

Toinen ongelma nykyaikaisessa kaupunkimallinnuksessa liittyy resoluutioon ja eri tarkkuuden tasoihin. Ilmiötä voidaan tarkastella yhdellä tasolla, mutta ajalliset muutokset voivat aiheuttaa sen, että ilmiön laajuus vaatii laajempialaisen tarkastelun ja täten resoluution muuttamisen. J. Van Vilet et al. (2011) ehdotti, että soluautomaateilla voisi olla vaihtelevia resoluutioita eri tarkastelutasoilla. Tällöin samaa geograafista aluetta voitaisiin esittää hierarkkisesti järjestetyillä tasoilla, joiden solukoot ovat erikokoisia. Peruseriaate on, että spatiaalisesti merkittävät alueet olisivat korkeampiresoluutioisia kuin kauempana olevat alueet, jotka esitetään matalaresoluutioisina. Tällöin kauempana olevat vaikutukset esitetään yleisemmällä tasolla eikä niiden vaikutukset ole yhtä merkittäviä kuin lähempänä olevien kohteiden. Kyseinen jako jakaa vaikutukset karkeasti lokaaleihin, pienempien tekijöiden vaikutuksiin, sekä globaaleihin, yleisempiin vaikutuksiin. (Marceau & Benenson 2011)

### 3.2.2 Tila

Solun tila määrittää sen mitä ilmiötä mallinnetaan. Yksinkertaisessa mallissa solulla voi olla kaksi eri tilaa: kehittynyt tai kehittymätön, tai toisin sanoen elävä tai kuollut. Kyseinen ajattelumalli on alun perin johdettu Game of Life-mallista, jota on tutkittu runsaasti sen yksinkertaisten sääntöjen, mutta samalla monimuotoisten hahmojen ja ilmiöiden mallina. Yksinkertaisena sääntönä tilan kehittymiselle voidaan pitää mahdollisuutta, johon vaikuttaa solua ympäröivien muiden solujen kehittyneiden solujen määrä. (Li et al. 2000)

Solun tilan määrittää kaksi asiaa: tilan sääntö eli funktio sekä naapurusto. Tilafunktio on yksi tärkeimmistä yksittäisistä tekijöistä, kun suunnitellaan soluautomaattia. Tilafunktio suoraan määrittää soluautomaatin toiminnan ja haluttavan ilmiön leviämisen mekaniikat. Tyypillisin tapa kuvata ilmiön leviämistä on laskea solun potentiaalisuus kehittyä riippuen sen naapuruston kehittymisen asteesta. Potentiaalisuus kehittymiselle lasketaan tyypillisessä tapauksessa summa-arvona kaikista naapuruston solujen potentiaaliarvoista, jotka lasketaan tilafunktion avulla. (Wolfram 1984) Totalistinen (totalistic) soluautomaatti ottaa huomioon pelkästään solun naapuruston solun, kun uutta arvoa lasketaan. Ei-totalistinen soluautomaatti käyttää käsiteltävän solun nykyistä arvoa. (Wolfram 1994)

Edellä kuvattu peruseriaate solun arvon laskemiseen toimii pohjana, jonka päälle voidaan lisätä tarkentavia sääntöjä. Monissa tapauksissa halutaan ottaa huomioon maaston tai ympäristön vaikutus ilmiön leviämiseen. Tämä tapahtuu esimerkiksi painoarvojen avulla. Painoarvot ovat kertoimia, jotka liitetään summa-arvon sisälle merkitsemään kyseisen sijainnin soveltuvuutta (suitability) ilmiön edistymiselle. Esimerkiksi, jos tietyssä sijainnissa halutaan estää soluautomaatin tilan tai käyttäytymisen leviämisen, voidaan antaa sille sijainnille 0-arvoinen painoarvo, joka mitätöi solun potentiaalinarvon, jos soluautomaatti etenee sen kohdalle. Painoilla voidaan kuvastaa siis sitä, miten ilmiö käyttäytyy eri sijainneissa. (Dahal & Chow 2014) Kuvassa 5 on esimerkki siitä, miten painoarvoja voidaan käyttää useista eri lähteistä ja etsiä solun naapurustosta potentiaalinen solu kehittymiselle.

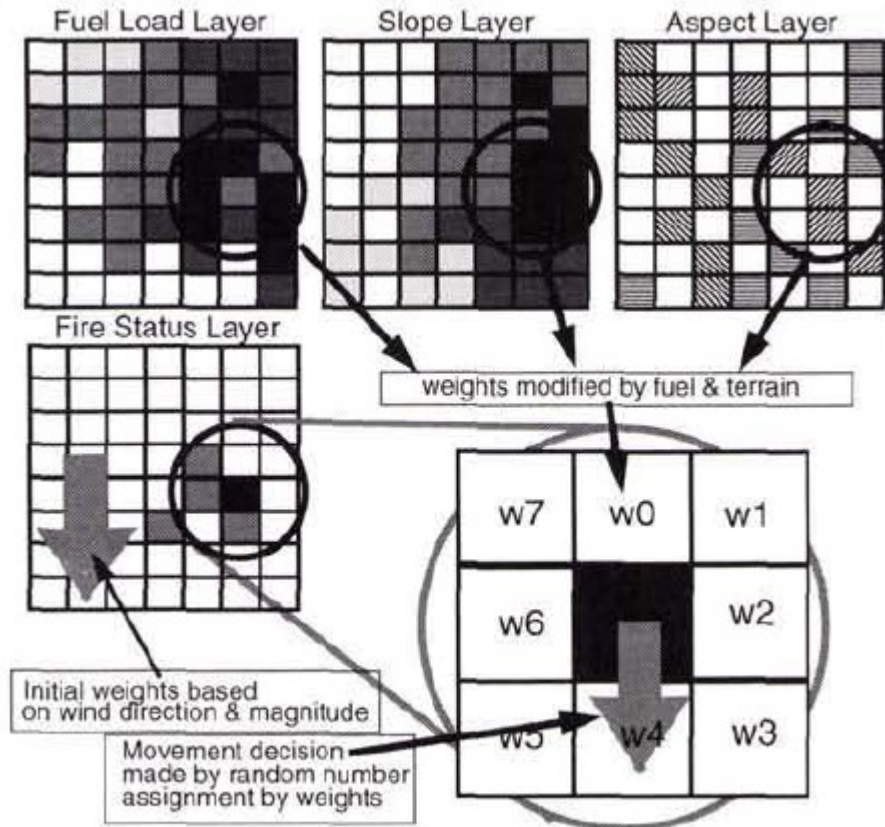


Figure 5. Firelet motion decision are made using weights from data layers.

Kuva 6 Esimerkki painoarvojen ja soveltuvuuden käytöstä soluautomaatissa. Useasta eri datatasosta koostetaan painoarvot soluautomaatissa ilmiön etenemisen suunnalle. (Clarke et al. 1994)

Monimutkaisempia implementointeja tilaan ja sen sääntöihin on kehitetty riippuen ilmiön toiminnan ymmärryksestä ja halutunlaisesta tarkkuudesta. Kun mallinetaan eteneviä luonnonilmiöitä, kuten metsäpaloa tai tauteja, voidaan ottaa huomioon myös suunta, josta soluautomaatin mallintama ilmiö yrittää levitä muihin soluihin. Esimerkiksi metsäpalojen tapauksessa tuli leviää puissa helpommin ylämäkeen kuin alamäkeen. Tällöin tulen leviämissuunnalla on merkitystä riippuen maaston pinnan muodoista (Karafyllidis et al. 1996) Monissa tapauksissa malliin voidaan liittää myös satunnaisuutta kuvaamaan pitkällä aikavälillä tapahtuvaa luonnollista vaihtelua. Satunnaisvaihtelu johtuu yleensä tekijöistä, jotka ovat niin pieniä merkitykseltään, ettei niiden mallintaminen malliin omana osana olisi kannattavaa. Toisaalta satunnaisuus mallissa saattaa lisätä epähaluttua kohinaa. Täten kohinaa voidaan pitää epähaluttuna ominaisuutena, kun halutaan mallilta selkeyttä kuvata ilmiöiden jakautuneisuutta eri aikaväleillä tai kun mallin realistisuuden vuoksi tietty määrä kohinaa on hyväksyttävissä. Satunnaisuutta voidaan myös kuvastaa lisäämällä tilafunktion normaalijakaumaa noudattava osatekijä. (Dahal & Chow 2014)

### 3.2.3 Säännöt

Soluautomaatin tilasäännöt olivat von Neumannin kehittämässä mallissa deterministisiä eli ehdottomia, tutkijan päättämiä sääntöjä. Melko pian ehdotettiin, että säännöt voisivat myös olla stokastisia, eli todennäköisyyteen perustuvia. 1900-luvulla tilastollisen matematiikan myötä tutkijoiden keskuudessa suosituksia tullutta Markovin prosessia alettiin implementoimaan myös soluautomaatteihin. Tämä johtui osittain siitä, että

soluautomaattien tilasäännöt ovat pohjimmiltaan hyvin samankaltaisia diskreettien stokastisten prosessien kanssa. (Benenson & Torrens 2004)

Markovin prosessissa ainoastaan muuttujan nykytilalla on merkitystä, eikä muutoksien historiaa tarkastella. Muuttujan tila määräytyy tilasäännön sijasta todennäköisyysmatriisilla, joka määrittää todennäköisyyden muuttujan tilan muuttumisesta. Samanlainen prosessi tapahtuu soluautomaateissa, joissa ajanhetkellä  $t+1$  solun tila muuttuu edellisen ajanhetken  $t$  tilan perusteella. Toisaalta soluautomaateissa todennäköisyydet eivät ole vakioita, vaan ovat riippuvaisia naapurustosta ja sen solujen tilasta. (Mondal & Southworth 2010) Kuten edellä mainittiin, on naapurustojen muoto ja laajuus muuttuneet soluautomaattien ja kaupunkimallinnuksen yhteydessä. Muutos johtaa siihen, että todennäköisyydet tilan muutokselle eivät ole lineaarisia, vaan riippuvaisia ympäristön solujen tiloista sekä muista myötävaikuttavista kertoimista. (Benenson & Torrens 2004)

Markovin prosessin implementointi soluautomaatteihin tuotti samankaltaisia tuloksia kuin oli myös huomioitu kompleksisista systeemeistä. Tutkijat huomasivat, että soluautomaateissa käytetyt stokastiset tilasäännöt tuottivat epälineaarisia tuloksia, mutta myös jaksottaisia tilamuutoksia automaatissa. Nämä kaksi huomiota kumpikin täyttivät ne havainnot mitä oli myös tehty kompleksisille systeemeille. (Benenson & Torrens 2004)

Markovin prosesseista Markovin ketju määritellään alkutilan todennäköisyysjakaumalla ja tilamuutosten todennäköisyysmatriisilla. Markovin ketju on stokastinen prosessi, joka ilmaisee miten todennäköistä on, että solun tila muuttuu toiseksi. Kun tunnetaan solun tila ajanhetkillä  $t$  ja  $t-1$ , voidaan näiden kahden tilan avulla laskea muutoksen todennäköisyys. (Benenson & Torrens 2004) Markovin ketju ei ota huomioon spatiaalisuutta, joten soluautomaatin avulla kyetään tuottamaan tehokas kehittyviä systeemeitä kuvastava malli. Markovin prosessin periaatetta voidaan myös käyttää mallin tarkkuuden ja varmuuden tarkastamiseen. (Brown et al. 2000) Prosessista on hyötyä, kun validoidaan tilasääntöjä ja ympäristön myötävaikuttavia kertoimia. Tämä käänteinen prosessi toimii luontaisesti soluautomaateissa ja on vahvuus verrattuna muihin malleihin. (Benenson & Torrens 2004)

Turner (1988) oli kehittänyt mallin, joka hyödynsi sekä soluautomaatteja että Markovin prosesseja solun tilan määrittämisessä. Kuten aiemmin todettiin, on Markovin prosessi stokastinen prosessi, jossa pelkästään muuttujien nykyarvolla on merkitystä. Soluautomaatit aloittavat prosessin naapuruston vaikutuksen tarkastelulla ja tilasäännöt ottavat huomioon ympäristönsä vaikutuksen. Markovin prosessit taas arvioivat todennäköisyysmatriisin arvot ja lisäävät niihin naapuruston vaikutuksen. Turner kalibrooi Markovin prosessin mallia siten, että solun tilan muutoksen todennäköisyys on muutoksen todennäköisyyden ja naapuruston, seuraavassa tilassa olevien, solujen kertoma. Kertoman tuloksen perusteella solun tila muuttuu sen perusteella kumpi todennäköisyys on suurempi. Turner hyödynsi malliaan oikean elämän maankäytön muutoksen mallinnukseen ja päätyi erittäin lähelle oikean elämän tilannetta. Kyseinen malli oli ensimmäisiä hyväksi todettuja maantieteellisten soluautomaattien malleja ja on toiminut pohjana nykyaikaiselle mallinnustekniikalle. (Turner 1988)

Batty (2013) esitti soluautomaatille yksinkertaisen algoritmin, jonka avulla kyetään laskemaan iteratiivisesti solun kehittyminen:

*IF any cell  $\{x_{+1}, y_{+1}\}$  is already developed*

*THEN  $P_d\{x, y\} = \sum P_d\{i, j\}/8$*

*&*

*IF  $P_d\{x, y\} > \text{threshold value}$*

*THEN cell  $\{x, y\}$  is developed with some other probability  $p\{x, y\}$*

$P_d\{x, y\}$  on todennäköisyys kehitymiselle solussa  $\{x, y\}$ , solut  $\{i, j\}$  ovat naapurisoluja Mooren naapurustossa  $\Omega$ . Solu  $\{x, y\}$  sisältyy tähän naapurustoon.

Kyseinen malli kehittää soluja siten, että jos tarkasteltavan solun naapurustossa on kehittynyt solu, lisääntyy todennäköisyys kehitymiselle. Tällä tavoin kehittyminen lähtee liikkeelle silloin, kun solun naapurustoon tulee uusi kehittynyt solu. (Li et al. 2010) Algoritmi on yksinkertainen, mutta tehokas ja se ottaa huomioon Markovin prosessin mukaiset todennäköisyydet.

Toinen kaupunkien kasvua ja kehittymistä kuvastava malli ottaa huomioon myös sijainnin sopivuuden (suitability) kehitykselle. White et al. (1997) ja Wu & Webster (1998) kehittivät malliinsa sopivuutta kuvastavan määreen, joka ilmaisee kuinka sopiva sijainti kyseinen solu on kehitymiselle. Solut, joilla on korkea sopivuuden kehityksen arvo, valitaan ensimmäisenä iteroinnissa kehitettäväksi. Kyseinen tekniikka käyttää myös yleensä multi-criteria evaluation -tekniikoita sijainnin kehityksen sopivuuden arvioimiseen. Maankäytössä sopivuuden arvo voi kuvastaa, miten otollinen maa-alue on kestävä kehityksen kannalta ja täten heikon maa-aineksen alueille sijoitetaan enemmän kasvua, jotta ne tulisivat tehokkaammin käyttöön. (Li et al. 2010)

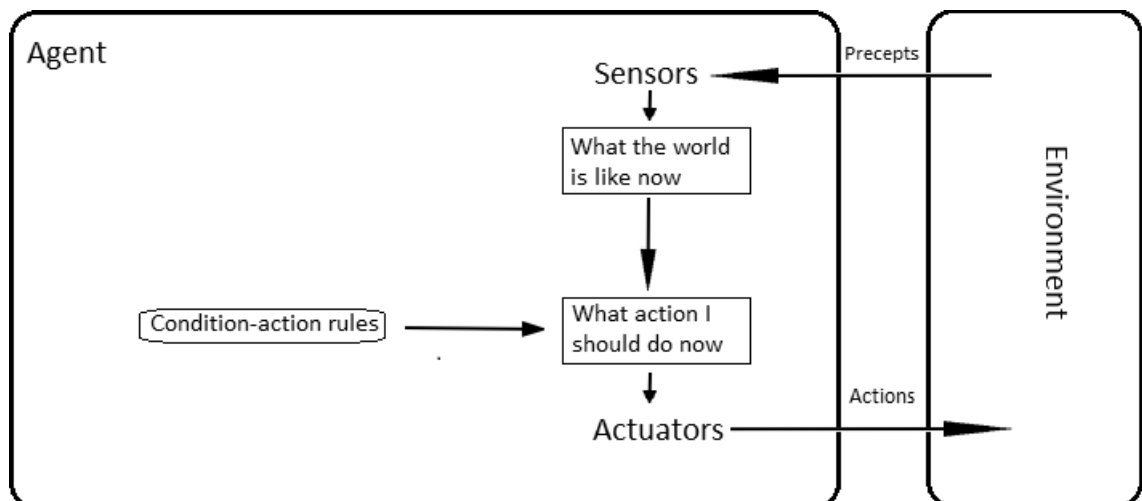
## 4 Agentit

### 4.1 Yleistä

#### 4.1.1 Tyypilliset piirteet

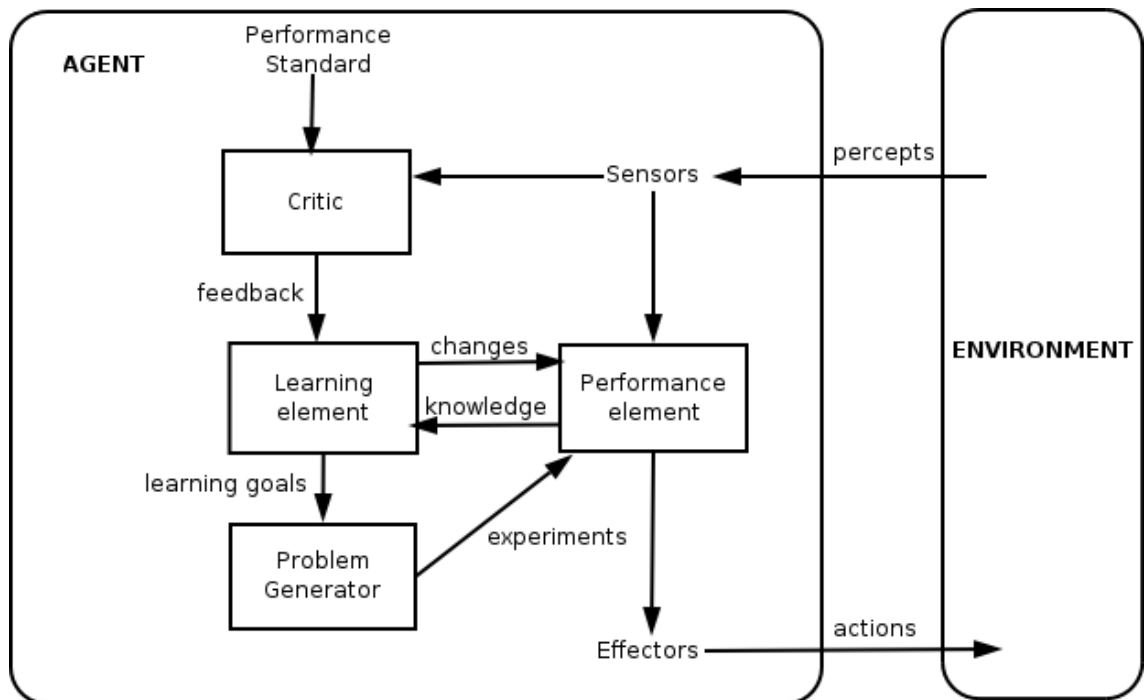
Agentin määritelmästä on useita muotoja, mutta Eric Bonabeau (2002) totesi, että agentit ovat pikemminkin ajattelutapa kuin spesifinen teknologia. Yhtäläisyyksiä agentteihin voidaan huomata keinoälyssä, olio-ohjelmoinnissa sekä koneoppimisessa. Geosimuloinnissa agentit ovat käytännössä ohjelmoituja olioita, jotka toimivat niiden käytössä olevan informaation perusteella. Tyypillisesti agentti on entiteetti mallinnuksessa jostakin mallinnettavassa ympäristössä olevasta objektista. Agentteja käytetään geosimuloinnissa, koska ne soveltuvat hyvin mallintamaan heterogeenisiä ilmiöitä. (Russel & Norvig 1995)

Agenteista ja niihin liittyvistä systeemeistä käytetään useita erilaisia nimityksiä kirjallisuudessa. Yleisin termi, joka ilmenee geosimuloinnin yhteydessä, on multi-agent systems. Termin lisäksi käytössä on agent-based model/system. Tässä työssä käytetään suomennettua termiä agenttipohjaiset systeemit.



Kuva 7 Yksinkertainen Simple Reflex Agent -tyyppinen agentti. (Russel & Norvig 1995)

Kuvassa 3 on havainnollistava kuva yksinkertaisen agentin toiminnasta. Agentti kerää informaatiota ympäristöstä ja toteuttaa toiminnon, jos toiminnolle asetettu ehto täyttyy. Agentin suorittama toiminto taasen vaikuttaa ympäristöön, jossa agentti toimii. (Russel & Norvig 1995)



Kuva 8 Tyypillinen Learning agent -tyyppinen malli agentista. (Russel & Norvig 1995)

Kuvassa 8 on monimutkaisempi malli oppivasta agentista. Agentti kerää informaatiota ympäristöstään ja käyttää informaatiota luodakseen tehokkaampia toimintoja aikaisempien kokemusten kautta. Agentissa oleva Critic -elementti arvioi agentin toimintaa ja tehokkuutta, ja neuvoo Learning -elementtiä. Learning -elementti taasen tekee arvion siitä miten tehokkuutta voidaan parantaa ja muuttaa Performance -elementtiä. Performance -elementti toimii kuten Simple Reflex -agentti, eli käyttää ulkopuolelta saatua informaatiota ja päättää toiminnosta. Problem generator -elementti pyrkii kokeilemaan eri skenaarioita ja ratkaisuja saadun informaation ja aikaisemmin opitun tiedon kautta. (Russel & Norvig 1995)

Agenttipohjaiset systeemit ovat pohjimmiltaan automaatteja soluautomaattien ohella, kun agenteja käytetään geosimuloinnissa. Verrattuna muihin automaatteihin, agentit soveltuvat hyvin ihmismäisen käyttäytymisen mallintamiseen. Ihmisten käyttäytymisen myötä tapahtuu kaupunkiympäristössä muutoksia, jotka voivat johtaa eri ilmiöihin, kuten kaupungin kasvuun ja kehittymiseen tai ympäristön näivettymiseen. Tutkimukset käyttäytymisestä ja niihin liittyvistä attribuuteista toimivat pohjana agenteille. Agenttipohjaisella systeemillä pyritään mallintamaan globaalit vaikutukset kaupunkiympäristöön. (Benenson & Torrens 2004) Geosimuloinnin ja muiden sovellusalueiden pohjalta on huomattu, että itsenäisyys järjestelmän sisällä on agenteille tyypillinen piirre. Tämä tarkoittaa, että agentit eivät tarvitse diskreettejä käskyjä. Tällöin agentit toimivat järjestelmässä ilman käyttäjän suoraa interaktiota, ja eri agenteilla voi olla hyvin erilaiset tavoitteet ja motiivit toisiinsa verrattuna. (Wooldridge 2009)

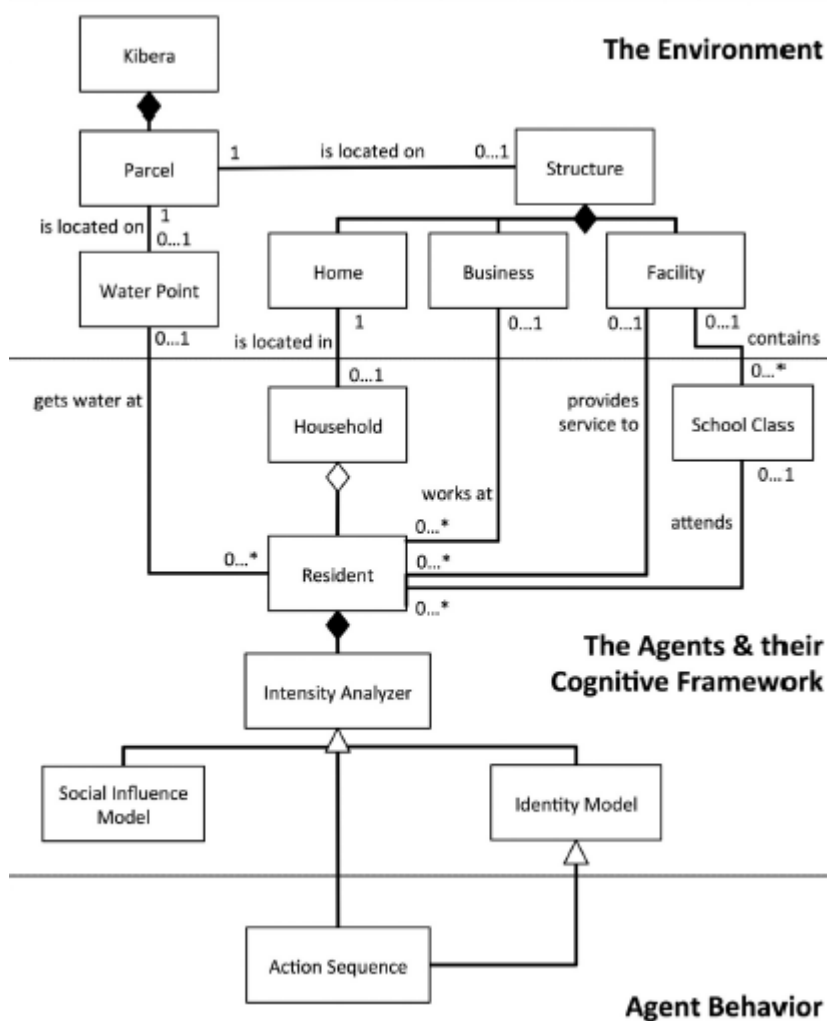
Russel ja Norvigin (1995) mukaan agentit ovat työkalu systeemien analysoimiseen, eivätkä absoluuttinen taho, jonka avulla voitaisi jakaa maailma niihin, mitkä ovat agenteja ja mitkä eivät ole. Tämän vuoksi agenteja voidaan pitää eräänlaisena keinoälyn johdannaisena. Geosimuloinnin ja GIS:n yhteydessä agenteille määritellään silti tiettyjä ominaisuuksia, jotta agentit kykenevät ympäristössään toteuttamaan haluttua käyttäytymistä. (Vahidnia et al. 2015) Agenteilta puuttuu silti diskreetti notaatio siitä miten ne toimivat missäkin tapauksessa. Agenttien käytön suunnittelu kuuluu käyttäjälle ja on osa suunnitteluprosessia (Benenson & Torrens 2004).

Verrattuna muihin automaatteihin agenteille on määritelty neljä ominaisuutta, jotka erottavat ne muista. Franklin ja Graesserin (1996) mukaan:

1. Agentti on automaatti, joka toimii ympäristössään.
2. Agentti aistii ympäristönsä, jossa toimii.
3. Agentilla on oma agenda.
4. Agentti toimii sen perusteella, mitä se ennustaa tulevaisuudesta.

Soluautomaatit kattavat kaksi ensimmäistä kohtaa, mutta kahdesta jälkimmäisestä kohdasta ei voida olla täysin varmoja. Erityisesti tulevaisuuteen näkeminen uupuu soluautomaatilta täysin, koska solut ”katsovat” vain naapurustoonsa, joten automaatti ei kykene tulkitsemaan, mitä naapurustosolukon ulkopuolella tapahtuu. Jalostamalla soluautomaattien naapurustoja, sisällyttämällä niihin esimerkiksi puskurin, jonka vaikutusalue kattaa naapuruston ulkopuolisia alueita, voidaan tuottaa älykkäämpiä ja ennakoivampia soluautomaatteja. Mutta Wolframmin ja von Neumannin kuvailemassa perusmuodossaan soluautomaatit eivät sisällä näitä tekniikoita. (Benenson & Torrens 2004) Kompleksisissa systeemeissä, kuten kaupunkien kohdalla, sisältyy agenttien toimimiseen tulevaisuutta ajatellen haasteita. Seuraamuksien käsite ja valintojen tekeminen niiden perusteella liittyy vapauden käsitteeseen. Jotta agenteilla voisi olla ihmismäinen päättelyketju seuraamuksien kannalta, on niiden toiminnoille myönnettävä vapaus toimia systeemissä tietyllä laajuudella. Vapaus valita ja toimia taasen liittyyvät agendaan, joka agentille myönnetään. Agenttien tekemien valintojen ja seuraamuksien ongelmaa sanotaan myös oikeudenmukaisuuden ongelmaksi. (de Jong 2008) Francez (1986) kuvasi ongelmaa kompleksisissa systeemeissä siten, että lokaalisti tehdyllä valinnalla voi olla kauaskantoiset seuraukset, joita on hankala ennustaa. Jos systeemin ja sen osallisten käyttäytyminen tunnetaan, kauaskantoisia ennustuksia voidaan tehdä, mutta ongelma muodostuu tällöin aikaikkuna, minkä aikana agentin on tehtävä päätös. Jos aikaa ei ole määritelty, agentti voi harkita eri vaihtoehtoja tarpeeksi pitkään, mutta suurimmassa osassa tilanteita tämä ei ole mahdollista. (Wooldridge 2009)

Kun agentit toimivat joukkona tietyssä ympäristössä, kutsutaan tätä kokonaisuutta agenttipohjaiseksi systeemiksi. Jotta systeemi toimii, ovat agentit vuorovaikutuksessa keskenään. Ihmismäistä käyttäytymistä varten agenttien välinen vuorovaikutus on tärkeää, kuten kommunikointi, neuvottelu ja tehtävien koordinointi. Geosimulaation tapauksessa agenteja lähestytään puhtaasti laskennallisina, keinotekoisina olioina. (Wooldridge 2009)



Kuva 9 UML (Unified Modeling Language)-kaavio agenttipohjaisesta systeemistä ja siihen kuuluvista osatekijöistä. Tässä mallissa agenteja ovat asukkaat (Residents) ja kotitaloudet (Household). (Crooks 2017)

Agenttipohjaiset systeemit tuottavat omat vahvuutensa muihin automaattisysteemeihin verrattuna, mutta täyttä paremmuutta ei ole pystytty todistamaan metodien välillä. Pohjimmiltaan eri automaattitekniikoilla kyetään toteuttamaan samanlaisia skenaarioita ja saamaan samankaltaisia tuloksia, mutta lopputulos riippuu paljon tutkijasta sekä hänen henkilökohtaisesta kokemuksestaan. Agenttien monikäytännöllisyyden vuoksi niitä voidaan käyttää samoissa tehtävissä kuin soluautomaatteja. Agenttien pohjimmainen tarkoitus, mallintaa ihmiskäyttäytymistä, antaa silti todennäköisesti paremman tuloksen, kun niitä käytetään systeemeissä, missä ihmismäisen käyttäytymisen mallinnus asetetaan tarpeelliseksi tavoitteeksi. Tämän vuoksi vastuu jää käyttäjälle tunnistaa ilmiön taustat ja miten paljon ne ovat yhteydessä yksilöiden päätöksentekoon ja agendojensa ajamiseen. (Benenson & Torrens 2004)

#### 4.1.2 Agenttien luokittelu

Agenttien luokittelu ei ole välttämätön vaihe agenttipohjaisen systeemin suunnittelussa, mutta luokittelusta on hyötyä systeemin ymmärtämisen kannalta. Kun agenttipohjaista systeemiä käytetään kaupungin mallintamisessa, voi eri kohdeluokkia ja niiden entiteettejä syntyä paljon. Ongelmaksi muodostuu tällöin, että eri kohteisiin liittyvät ilmiöt johtuvat lukuisasta joukosta vaikuttavia tekijöitä. Tällöin mallinnettavan systeemin kompleksisuus kasvaa. Monimutkaisen systeemin mallintamista varten suunnittelussa on

hyödyllistä osata jaotella ja luokitella agentteja. Yksi tapa jaotella agentteja systeemin sisällä eri hierarkioihin on agentin sisäisen ajan skaala. Kun mallinnetaan esimerkiksi yhtiötä agenttina, lasketaan agentin aikaa kuukausissa, kvartaaleissa tai vuosissa riippuen yhtiön koosta. Taasen yksittäisen asukkaan aikaa voidaan laskea minuuteissa tai sekunneissa. (Benenson & Torrens 2004)

Agenttien säännöt laaditaan pääosin joko käyttämällä empiiristä dataa tai teoreettiselta, kokeelliselta pohjalta. Sääntöjen laatiminen agenteille on yksi haastavimmista kohdista agenttien käyttämisessä. Agentit voidaan jakaa sääntöjen puolesta joko reaktiivisiin agenteihin tai aktiivisiin agenteihin. Reaktiiviset agentit ainoastaan reagoivat muihin tapahtumiin ympäristössään tai muiden agenttien toimiin. Aktiiviset agentit taasen suorittavat toimintoja, jotka edistävät niiden tilaa lähemmäksi niille asetettua päämäärää. (Crook & Heppenstall 2012) Kyseinen jako ei ole sitova, mutta on agenttien suunnitteluvaiheessa hyvä lähtökohta.

## **4.2 Agenttien erityispiirteet**

Kuten luvun alussa todettiin, ei agenteilla ole olemassa perusmallia, joka olisi muodostunut standardiksi agentteja käyttäville järjestelmille. Agenteille voidaan silti listata joukko piirteitä, jotka ovat niille ominaisia ja erottavat agentit omaksi systeemikseen.

Agentit eroavat perinteisestä ajatuksesta kuvata jokaista yksilöä ryhmänsä keskiarvoisena ominaisuuksien suhteen. Agentit ovat kukin yksilöitä ja vaikka ne toimisivat systeemissä yhtenä ryhmänä, on jokainen agentti luotu riippumattomaksi joukon muista yksilöistä. (Gorodetskaya 2013) Kun kompleksisessa systeemissä esiintyy epälineaarisuutta, on erittäin hyödyllistä, että kyetään osoittamaan yksilöllisiä agentteja, jotka toimivat itsenäisesti ja täten vaikuttavat epälineaarisesti. Tästä on hyötyä myös, kun halutaan tarkastella systeemin epälineaarisuuden vaikutusta sijoittamalla agentteja muutoin lineaariseen systeemiin. Agentit ovatkin siksi heterogeenisiä. (Benenson & Torrens 2004)

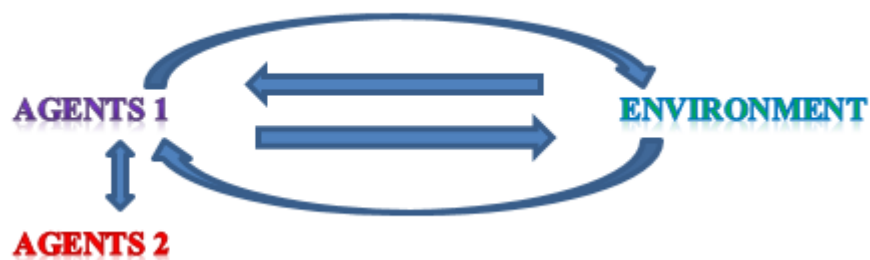
Agenteilla on yleensä tietty tavoite, johon ne pyrkivät. Tämän johdosta agentin toimintatavat riippuvat tästä tavoitteesta ja ne toimivat tämän tavoitteen puitteissa. Tämän vuoksi voidaan sanoa, että agenteilla on oma agenda, mihin ne pyrkivät. Vastaavanlaista ominaisuutta ei löydy soluautomaatilta. (Benenson & Torrens 2004) Jotta agentit kykenevät toimimaan agendansa pohjalta, on niillä oltava myös rationaalisuutta tehdä valintoja. Valintojen tekemistä varten agentilla on oltava kyky kerätä tietoa tulevasta ja tehdä saadun informaation pohjalta sekä lyhyen että pitkän aikavälin ennustuksia. Ennustuksien pohjalta agentti voi päättää toimia siten, että maksimoi saadun hyödyn ja pääsee lähemmäs tavoitettaan. Ongelmallinen piirre simuloinnissa syntyy tarjotun informaation pohjalta, koska agenteille on rajattava oikeudet vain tiettyyn informaatioon. (Gorodetskaya 2013)

Agentit havainnoivat ympäristöään ja käsitys naapurustosta eroaa tämän vuoksi soluautomaateista. Havainnointi ei silti rajoitu pelkästään naapurustoon, vaan inhimillisen käyttäytymisen vaatimuksena agentit tunnistavat ympäristöstään myös käyttäytymiseen vaikuttavia tekijöitä. (Benenson & Torrens 2004) Riippuen onko agentti reaktiivinen vai aktiivinen toiminnaltaan rajoittuu sen havainnointikyky joko maantieteellisesti alueeseen, johon agentilla on vaikutusta, tai agentin keräämän tiedon perusteella koostettuun ns. havaintokarttaan. Päätöksiä tehdessä agentti reagoi muihin havaintokykynsä alueella oleviin objekteihin tai muihin agenteihin. Myös edellä kuvattu

agenteille saatavissa oleva informaation määrän ongelma liittyy agentin havainnointikykyyn. (Gorodetskaya 2013)

Kommunikaatio on myös agenttien erityispiirre. Aktiivisesti toimivat agentit kykenevät vaihtamaan tietoa toistensa välillä suorittaakseen mahdollisesti yhteistä tavoitetta. Kommunikaatio voi myös tarkoittaa sen puutetta eli agentit kykenevät suodattamaan sen tiedon, mikä ei ole niille tarpeellista sillä hetkellä. Tämän vuoksi agenteja voidaan myös pitää mukautuvina, koska uuden informaation myötä agentit voivat muuttaa omia sisäisiä sääntöjään soveltumaan paremmin uusiin olosuhteisiin. (Benenson & Torrens 2004) Toisaalta agentit voivat myös pysyä oppimisen suhteen staattisina eli ne eivät käytä saatavilla olevaa informaatiota sisäisten sääntöjen muokkaamiseksi. (Gorodetskaya 2013)

Viimeinen ja ehkä tärkein ominaisuus agenteille kaupunkimallinnuksessa on niiden kyky liikkua tilassa. Toisin kuin soluautomaatit agentit voivat liikkua mallin tilassa joko vapaasti, liikkua ainoastaan niille rajoitetulla alueella tai pysyä staattisesti paikoillaan. Huomattavaa staattisuudessa on, että agentit kykenevät silti vaikuttamaan spatiaalisesti kauempanakin oleviin kohteisiin, vaikka niiden maantieteellinen sijainti pysyisi tietyssä paikassa. Yleensä soluautomaateilla staattisten kohteiden spatiaalinen vaikutus vähenee välimatkan lisääntyessä. Agentit kykenevät sen sijaan vaikuttamaan joko lähellä tai kaukana oleviin muihin kohteisiin tai muihin agentteihin. Tämän johdosta agenttien vaikutus on sekä paikallista että globaalia. (Gorodetskaya 2013)



Kuva 10 Agenttien interaktio muiden agenttien ja ympäristön kanssa (Gorodetskaya 2013)

## 5 Geosimulointi

### 5.1 Yleistä

#### 5.1.1 Yleiset ominaisuudet

Geosimulaatio on kattotermi, joka kattaa joukon eri tekniikoita ja metodeita spatiaalisen, suuriresoluutioisen datan käsittelyyn ja simulointimallin prosessointiin. Geosimulointi on varsin uusi tekniikka spatiaalisen datan käsittelyssä. (Benenson & Torrens 2004) Pääasiassa geosimulointia käytetään kaupunkiympäristön mallinnuksessa, mutta lähivuosina myös muita sovellusalueita on ilmennyt, kuten väkijoukkojen liikkuminen sekä tautien leviäminen. Geosimuloinnin suosion kasvun myötä myös muita tekniikoita paikkatietoalalta on tutkittu geosimulaatioiden tueksi sekä laajentamaan käyttökohteiden määrää. (Shen et al. 2009)

Yleisperiaate geosimuloinnille on, että mallit luodaan aloittaen pienemmistä objekteista ja niiden käyttäytymisestä kohti yleispätevämpää suurta kokonaisuutta. Kaupunkiympäristöissä on runsaasti kohteita, joiden käyttäytymisen seuraamiseksi joudutaan tekemään lukuisia kompleksisia ja dynaamisia systeemeitä, jotta lopputulos olisi riittävän realistinen. (Benenson & Torrens 2004) Tämä asettaa geosimuloinnille tavoitteeksi mallintaa kohteiden ja objektien dynaamisia suhteita toisiinsa spatiaalisesti ympäristössä, jossa kompleksisuus ja epävarmuus ovat isoja. Epävarmuustekijät ja kompleksisuus johtavat kaupunkiympäristön suunnittelussa siihen, että alkuperäinen konsepti voi olla hyvin erilainen lopputuotteen kanssa, koska projektin aikajänne on pitkä. (Shen et al. 2009)

#### 5.1.2 Geosimulointimallien pääpiirteet

Neljä pääpiirrettä erottaa geosimulointimallit muista malleista, joita käytetään kaupunkimallinnuksessa.

Kaupunkimalleissa eri objektien ja kohteiden kokonaisuutta kuvastetaan yhtenä ryhmänä, jota voidaan mallintaa yleistämällä niiden keskinäiset ominaisuudet. Kun otetaan maantiede huomioon, voidaan ryhmät jakaa spatiaalisesti tutkittavaan ympäristöön. Geosimuloinnissa taasen kohteet ovat harvoin liikkumattomia ja spatiaalisesti muuttumattomia. Tällaisia kohteita ovat esimerkiksi yksittäiset kotitaloudet ja ajoneuvot. (Benenson & Torrens 2004)

Kaupunkimallien spatiaaliset riippuvuudet kuvataan usein malleilla, joissa on vetovoimatekijöitä. Liikenteen ja ihmisten virta kulkeutuu kohti tiettyjä keskuksia ja etäisyys vetovoimaisiin kohteisiin vähentää niiden merkittävyyttä. Geosimulointi omaksuu huomattavasti yksityiskohtaisemman lähestymistavan: objekteilla ja kohteilla on oma käyttäytymismalli, joka aiheuttaa niissä riippuvuutta lukuisiin muihin kohteisiin. Suuremmassa skaalassa kohteiden väliset interaktiot ja käyttäytyminen ovat syynä eri ilmiöille. Vahvuutena tässä mallissa on paljon monimuotoisemmat interaktiot eikä vetovoimatekijät riipu yhtä paljon etäisyydestä. (Benenson & Torrens 2004)

Ajan käsite perinteisissä kaupunkimalleissa liittyy eri ilmiöiden omiin sisäisiin kelloihin. Syklisten ilmiöiden toistumiset johtuvat eri aikaskaaloista, joissa systeemit toimivat ja suurempia systeemeitä tarkastellaan pitemmällä aikavälillä. Geosimuloinnissa kohteiden

käyttäytyminen on hyvin paljon yksityiskohtaisempaa, koska ajankäsite on monimuotoisempi. Karkea jako ajan kuvastamiseen on, että kohteet voivat toimia synkronoidusti tai epäsynkronoidusti. Synkronoidut kohteet suorittavat tietyn käyttäytymistoiminnon samanaikaisesti ja epäsynkronoidut taas voivat tuottaa käyttäytymisen mukaisen toimintansa satunnaisesti tietyissä sekvensseissä tai ennalta määrätyn tapahtuman johdosta. Tällä tavoin kuvastettu aika seuraa kohteiden toimintojen mukaista aikataulua. Kyseinen tapa ajoittaa tapahtumien kulkua on soluautomaattien tapa toimia kuten aiemmin työssä mainitaan. (Benenson & Torrens 2004)

Geosimulaatio pyrkii muodostamaan uskottavan kuvaus oikean maailman tapahtumista, mutta myös tuottamaan mallin, jolla kyetään luomaan abstrakteja tilanteita ja skenaarioita. Samaa työkalua pystytään siis käyttämään uusien ideoiden jalostamisessa mallintamalla ja simuloimalla mahdollisia tapahtumia.

Merkittävin ero muihin simulointimenetelmiin on geosimuloinnin tapa selittää ilmiöitä yksittäisten kohteiden käyttäytymisen kokonaisuutena pikemminkin kuin niiden summana. Ilmiötä ei pyritä jakamaan pienempiin osatekijöihin ja sitä kautta loogisiin komponentteihin, vaan kuten on osoitettu Conwayn Game of Lifessa, yksinkertaiset säännöt ja alkuasetelmat kykenevät tuottamaan monimuotoisen järjestelmän, mikä ei ollut alun alkaen odotettavissa. Tämä tapa havainnoida odottamattomia vaikutuksia on yksi pääsyy siihen, miksi geosimulointia pidetään merkittävänä simulointimenetelmänä.

### 5.1.3 Geosimuloinnin haasteet

Geosimuloinnin määrittelyssä esiintyy tekijöitä, jotka aiheuttavat suuren joukon haasteista geosimulointimallin muodostamisessa. Benenson ja Torrensin (2004) mukaan geosimulointi on objektilähtöinen spatiaalisesti tarkka mallinnus dynaamisista systeemeistä. Objektilähtöinen ajattelu tarkoittaa diskreettejä, systeemissä esiintyviä objekteja, jotka voivat olla joko elollisia tai elottomia. On huomattu, että heterogeenisissä systeemeissä, kuten sosiaalisissa ja ekologisissa systeemeissä, yksittäiset itsenäiset toimijat yhteisessä tilassa ovat tärkeässä roolissa kompleksisen systeemin toiminnassa. (Grimm & Railsback 2005) Agentteihin perustuvat systeemit on havaittu tehokkaaksi keinoksi mallintaa tämän kaltaista itsenäisistä toimijoista muodostuvaa systeemiä ympäristössä, jossa malli mukautuu ja kehittyy muutoksien myötä. (Benenson & Marceau 2011)

Toisekseen spatiaalisesti tarkka maailma ja sen mallinnus tarkoittaa maailman dynamiikan ja sen tekijöiden mallinnusta riittävän tarkalla tasolla. Spatiaalisuuden esiintymiseen ja sen havainnointiin on luotu joukko työkaluja erityisesti geoinformaatiojärjestelmien alalla. Pääasiallinen tavoite mallinnukselle on siis riittävän realistinen kuvaus maailmasta sellaisena kuin se on. Mallin on siis kestävä vertailua oikean maailman tilanteisiin ja havaintoihin, jotta sitä voitaisiin pitää riittävän tarkkana. Soluautomaattien näennäisesti yksinkertainen tapa kuvata maailman säännönmukaisuuksia rasteripohjaisia systeemeitä muistuttavalla mallinnustavalla on koettu toimivaksi. (Benenson & Marceau 2011)

Geosimulointimallin rakentamisen tärkeä vaihe on validointi, mallin hienosäätö ja verifiointi. Malli ei ole uskottava tai käytännöllinen, jos ei ole osoittaa sen toimivuutta sille tarkoitettuun käyttöympäristöön ja –tarkoituksessa. Puhutaan myös mallin asiaankuuluvasta määrittelyalueesta. Muilla tieteenaloilla, joissa käytetään soluautomaatteja ja agentteja, kuten sosiaalitieteissä ja taloustieteessä, on validoinnin ja verifiointin prosesseja tutkittu paljon, ja samoja periaatteita voidaan hyödyntää

geosimulointimallien vastaavissa kriittisissä prosessivaiheissa (Benenson & Marceau 2011).

Hienosäätö tehdään, jotta malli olisi lähempänä oikean elämän vastinetta parametrien ja muuttujien osalta (Benenson & Marceau 2011). Verifioinnissa luotua mallia verrataan konseptuaaliseen kuvaukseen ja täten mallin tulisi täyttää ne tavoitteet, mitä sille on asetettu. Validoinnissa sen sijaan tarkistetaan, miten hyvin malli sopii siihen tarkoitukseen, mihin se on rakennettu. (Benenson & Marceau 2011)

Validoinnille voidaan muodostaa kolme kategoriaa: konseptuaalinen, numeerinen ja toiminnallinen. Numeerinen validointi on metodeista helpoin. Paikkatiedon aineistojen ja luodun mallin samankaltaisuuden vertailuun voidaan käyttää tilastomatematiikassa käytettyjä metodeita, kuten Cohenin Kappa-indeksiä. (Louie & Carley 2008) Numeerisia arvoja vertailemalla voidaan helpoiten hienosäätää mallin parametrien arvoja, jotta lopputulos olisi lähempänä oikean elämän vastinetta.

Konseptuaalinen validointi sen sijaan on huomattavasti haastavampaa geosimulointimalleille. Osa tutkijoista on jopa esittänyt, että ihmismäisen käyttäytymisen mallit ovat liian kompleksisia, jotta ne voisivat kuvastaa oikeaa elämää tarpeeksi tarkasti. Tämä johtaa siihen, että konseptuaalinen validointi ei ole mahdollista riittämättömän tarkkuuden vuoksi. Tällöin mallin tekijän tehtäväksi jää määrittellä mallin parametrien rajat ja uskottavuus sillä tasolla, että yhteensopivuus on maksimoitu asetetuissa rajoissa. (Holling 1978) Toisenlainen evaluointi mallille on sen herkkyys eri parametreille ja aloitusehdoille. Eri herkkyydet erilaisilla resoluutioilla antaa arvokasta tietoa siitä, millaisilla tasoilla malli on vähemmän herkkä vaihtelevuudelle.

## **5.2 Agentit geosimuloinnissa**

Geosimuloinnin yhteydessä agentit tulkitaan automaateiksi, koska agenteilla on samoja ominaisuuksia kuin soluautomaateilla: naapurusto, soluja, tiloja sekä tilasääntöjä. Kuten edellä työssä kerrottiin, pääasiallinen ero soluautomaatteihin on agenttien kyky toimia inhimillisesti. Kyseisen ominaisuuden mallintaminen pelkästään tiloja ja tilasääntöjä käyttämällä on haastavaa. Inhimillisyyteen kuuluu kauaskantoisten päätösten tekeminen, mutta myös niistä joustamista odottamattomissa tilanteissa. Lukuisat pienet tekijät vaikuttavat päätöksiin ja niiden mallintaminen sekä hallinta tulee ongelmalliseksi mallin kasvaessa. Näiden lisäksi inhimillisessä käyttäytymisessä on epävarmuutta sekä ”virheellistä” käyttäytymistä esimerkiksi vahingon tai erehtymisen vuoksi. (Benenson & Torrens 2004)

Koska agentit ovat hyvin monitulkintaisia ominaisuuksiltaan, voidaan laatia joukko yleisominaisuuksia agenteille, jotka kattavat perusjoukon käyttäytymiseen vaadittavia tekijöitä mallinnusta varten. Kun yleisominaisuudet on katettu mallissa, kyetään agenteja jalostamaan erityisempiä tilanteita varten, sekä samalla pitää yhteneväinen pohja koko joukolla. (Benenson & Torrens 2004) Esimerkiksi eri ihmisagentteja varten voidaan koota joukko attribuutteja, jotka pätevät jokaiseen yksittäiseen ihmiseen riippumatta siitä, mikä kyseisen agentin tehtävä on. Crooks et al. (2012) mukaan agentit voivat kuvata mitä tahansa ilmentymää kohteesta, jolla on autonomia. Tähän lukeutuvat ihmiset, kommuunit, yhtiöt, rakennukset, eläimet tai jopa osakemarkkinoiden osakkeet. Luomalla pohjaksi perusominaisuudet kattavan agentin kyetään tehokkaammin koostamaan yhteneväisen joukon agenteja eri kohteille. Agenttien yhteneväinen pohja edistää myös niiden keskinäistä toimimista systeemissä.

### 5.3 Geosimuloinnin moderni malli

Benenson ja Torrens (2004) ehdottivat kirjassaan mallia geosimuloinnille, joka yhdistää agenttipohjaisen systeemin sekä soluautomaatit yhdeksi malliksi. Mallin tavoite oli kuvata maantieteellisessä tilassa olevia objekteja automaateina, jotka yhdessä muodostavat systeemin, jossa toimivat. Mallin nimi on Geographic Automata Systems (GAS).

#### 5.3.1 GAS

GAS on viitekehys, joka yhdistää sekä soluautomaatit että agentit yhdeksi systeemiksi. Tässä kappaleessa määritellään GAS:n perusominaisuudet.

GAS koostuu maantieteellisistä automaateista, joille on määritelty tiloja ja tilasääntöjä kuten muillekin automaateille. Toisin kuin normaaleiden soluautomaattien naapurustot, GAS:n automaateilla on määritelty joukko georeferoituja sääntöjä, jotka määrittävät sen sijainnin yhteisessä avaruudessa. Näiden lisäksi naapurussäännöt määrittävät millainen solun naapurusto on sen sijaan, että jokaiselle solulle määriteltäisiin yhteneväinen, muuttumaton naapurusto. Jotta kaupunkimalleille tyypillinen dynaamisuus täyttyy, on lisäksi joukko liikkumissääntöjä, jotka määrittävät miten solut kykenevät liikkumaan simuloitussa tilassa.

Edellä mainitut säännöt täyttävät vaatimukset, joita uudenmalliselle geosimuloinnille on asetettu. Mobiilit solut ja dynaaminen naapurusto ovat huomattavan realistisempi mallinnus kuin perinteiset mallit. Tällä tavoin kyetään soveltamaan soluautomaattien vahvuuksia kohteille, joita ei välttämättä olisi voitu mallintaa yhtä tehokkaasti perinteisin soluautomaatiotekniikoin. Lisäksi soluautomaateille on yleensä määritelty ennalta alueet joko geometrisenä tasajakona ruudukkoon tai muiden maantieteellisten piirteiden perusteella tehty jako. GAS ei määrittele näitä alueita ennalta, vaan georeferoidut säännöt määrittävät alueet. Tällä tavoin mallin käyttäjän vastuulle ei jää päättää sopivien alueiden jakoa ennen simulaation suorittamista.

Benenson määrittelee GAS:n funktiona  $G$ , joka sisältää komponentit:

$$G \sim (K; S, T_S; L, M_L; N, R_N) \quad (4)$$

$K$  on automaattien tyypit tai ontologiat. Loput kolme paria,  $S, T_S, L, M_L$  ja  $N, R_N$ , kuvastavat spatiaalisia tai ei-spatiaalisia piirteitä kolmessa eri pääkomponentissa.

$S$  on automaatin tilat ja  $T_S$  on tilasäännöt automaateille. Seuraava pari  $L$  ja  $M_L$  kuvaavat paikkatietoa.  $L$  on automaatin georeferoitu sijainti systeemin sisällä ja  $M_L$  on automaatin liikkumissäännöt.  $N$  ja  $R_N$  kuvaavat automaatin naapurustoa. Ne ottavat huomioon automaatin syötteenä saadun tiedon  $I$ . Syötteen  $I$ , naapuruston  $N$  ja naapurustosääntöjen  $R_N$  perusteella automaatti voi muuttaa tilaansa.

Dynaamisen ja spatiaalisesti riippuvaisen automaatin muutos ajanhetkellä  $t+1$  riippuu siis tilasäännöistä, liikkumisesta sekä naapurustosäännöistä  $T_S, M_L$  ja  $R_N$ . Tällä tavoin automaatti ottaa huomioon temporaaliset muutokset. Jokaiseen  $G$ :n automaattiin pätee seuraavat säännöt:

$$\begin{aligned}
T_S: (S_t, L_t, N_t) &\rightarrow S_{t+1} \\
M_L: (S_t, L_t, N_t) &\rightarrow L_{t+1} \\
R_N: (S_t, L_t, N_t) &\rightarrow N_{t+1}
\end{aligned}
\tag{5}$$

GAS muodostuu useasta eri tyyppisestä automaattista. Korkeimmalla tasolla jako perustuu siihen, onko automaatti liikkumaton vai liikkeessä. Liikkumattomat automaattit eivät muuta sijaintiaan ajan myötä, jolloin ne vertautuvat perinteisiin soluautomaatteihin. Ne käyttävät kaikkia komponentteja GAS:ssa paitsi liikesääntöjä  $M_L$ . Liikkeessä olevat automaattit taas voivat muuttaa sijaintiaan systeemissä. Yleensä kaupunkiympäristö koostuu useista liikkumattomista ja liikkeessä olevista automaateista, jotka yhdessä kuvastavat kokonaista kohdetta geosimuloidussa ympäristössä. Esimerkiksi asuinkerrostalon kiinteistö olisi liikkumaton automaatti, jonka tiloina toimii useita rakennuksen ominaisuuksia, kuten huoneistojen määrä, kerrosten määrä tai kivijalassa olevien mahdollisten toimitilojen määrä. Liikkuvat automaattit kyseisessä kohteessa olisivat asukkaat, joiden tiloina olisivat perheenjäsenten määrä, taloudellinen tila tai ikä. Yhdessä nämä kaksi eri tyyppistä automaattia määrittävät esimerkiksi asuinkerrostalon arvon ja väestön määrän naapurustossa. (Benenson & Torrens 2004)

Seuraavissa luvuissa käsitellään tarkemmin GAS:n pääkomponentteja, eli automaatin tiloja, naapurustoa sekä liikkumista.

### 5.3.2 GAS tilat

Tilat määrittävät automaatille sen keskeiset ominaisuudet, sekä mitä automaatti simuloi ympäristöstään systeemissä. Tilaan vaikuttavat tilasäännöt  $T_S$ , sekä ulkoiset muuttujat. Automaattiin vaikuttavat ulkoiset muuttujat voivat olla käytännössä mitä tahansa muuttujia, joilla on vaikutusta automaatteihin. Myös lähes mistä tahansa muuttujasta voidaan muodostaa automaatille tila. Tilojen on kuvastettava automaatin keskeisiä ominaisuuksia ja oltava simuloidulle ympäristölle ja ilmiölle vaikuttavia. Esimerkiksi liikkuvalle ajoneuvolle keskeisiä tiloja ovat sen ajanhetken nopeus, suunta ja päämäärä. Taasen asuinrakennukselle voidaan määrittellä tiloiksi kerrosten määrä, rakennuksen korkeus ja kiinteistön rajat.

Toisin kuin soluautomaateilla, GAS:n tilasäännöt ottavat huomioon kaikenlaiset automaattit, jotka vaikuttavat solun tilaan. Tämä tarkoittaa, että solun tila määräytyy soluun vaikuttavien muiden tekijöiden pohjalta pikemminkin kuin vain sen lähiympäristössä vaikuttavien tekijöiden mukaan. Tällöin liikkeessä olevat solut, tai agentit, voivat vaikuttaa myös solun tilaan. Kaupunkiympäristön dynaamisuus on siis pohjana kyseiselle eroavaisuudelle verrattuna perinteisiin soluautomaatteihin. Kyseisen tilan määrityksen vuoksi solujen muuttuvat tilat eivät etene samalla tavalla kuin soluautomaatissa, jossa tilat muuntuvat vähitellen edeten soluautomaatissa. (Benenson & Torrens 2004)

### 5.3.3 Georeferoidut säännöt

GAS:n  $L$  kuvaa millä tavoin maantieteellinen automaatti merkitään systeemin tila-avaruuteen. Jos automaatti ei liiku, sen merkitsemiseen riittää kaksi sijaintikoordinaattia, joiden avulla sijainti kyetään määrittämään riittävällä tarkkuudella. Riippuen mallinnettavan ympäristön koosta ja tarkkuudesta, eri määrä kohteisiin liittyvää sijaintitietoa on tarpeellista merkitä malliin. Jos skaala mallissa on tarpeeksi pieni, voidaan joutua merkitsemään esimerkiksi rakennuksille niiden rajat luonnossa, keskipiste, korkeus tai muuta tarpeellista sijaintitietoa. Suuremmalla skaalalla taas riittää, jos kohteita pidetään pistemäisinä. Tarkoituksena on, että sijaintitiedon

tarkkuudella kyetään riittävän tarkasti määrittämään kohde. Jos mallinnettavaa ympäristöä pidetään abstraktina, voi solupohjainen aluejako riittää. (Benenson & Torrens 2004)

Mallinnettavat kohteet voidaan jakaa kahteen ryhmään: edellä mainitut liikkumattomat kohteet sekä dynaamiset, liikkuvat kohteet. Liikkuvien kohteiden kanssa joudutaan ylläpitämään niiden sijaintikoordinaatteja ympäristössä, mutta myös niiden suhteellista sijaintia muihin kohteisiin nähden. Esimerkiksi taloudessa asuvien asukkaiden sijainti voi liittyä asuinkiinteistön osoitteeseen ja sitä kautta kiinteistön sijaintiin. Erityisesti liikkuvien kohteiden kanssa epäsuora georeferointi on hyödyllistä simulaation edetessä. (Benenson & Torrens 2004)

### 5.3.4 GAS:n naapurustot

Automaatin naapurustot,  $N$ , ja siihen vaikuttavat säännöt  $R_N$ , määrittävät millaisessa suhteessa automaatti on muihin automaatteihin nähden. Naapurustoja on määritelty joukko erityyppisiä solujen tilasääntöjen  $T_S$ , ja liikkumissääntöjen  $M_L$ , toiminnan kannalta. Naapuruussuhde ja sitä kautta naapuruussäännöt voidaan esittää usealla eri tavalla liikkumattomille kohteille. Naapuruutta voidaan ilmaista esimerkiksi topologisilla suhteilla (vierekkäisyys), verkkomaisesti solujen ollessa kohteita tilassa tai läheisyydellä. Onkin käyttäjistä kiinni miten naapuruussääntöjä  $R_N$  käyteään suhteiden ilmaisemiseen. Huomionarvoista on myös ajan vaikutus naapuruussuhteisiin. (Benenson & Torrens 2004)

Liikkuvat kohteet ovat haastavampia naapuruuden ilmaisemisessa. Etäisyyden ja lähimmän naapurin perusteella määritetty suhde muodostuu monimutkaiseksi, kun kohteiden määrä kasvaa. Tällöin naapuruutta voidaan ilmaista epäsuorasti. Koska mobiilit automaattit ovat yleensä agenteja, voidaan hyödyntää muita automaatteja agenttien välisen naapuruussuhteen määrittelyssä. Esimerkiksi kaksi agenttia voidaan määrittää naapureiksi, jos agentit sijaitsevat staattisissa automaateissa, jotka ovat naapureita keskenään. (Benenson & Torrens 2004)

## 5.4 Soluautomaattien ja agenttien heikkoudet

Soluautomaatteja ja agenttipohjaisia systeemeitä on käytetty erikseen sekä erinäisin yhdistelmin aiemminkin kaupunkimallinnuksessa ja geosimuloinnissa. Aiemmissä malleissa on silti jouduttu tekemään kompromisseja sen suhteen, mitä ominaisuuksia kummastakin tekniikasta on hyödynnetty ja pyritty paikkaamaan toisen tekniikan heikkouksia. Tällöin esimerkiksi soluautomaattien tiukkaa solujakoa on pyritty joustamaan sekä antamaan niille agenttimaisia piirteitä tai agentit on tulkittu soluautomaatista irrallisina tekijöinä, jotka ovat käyttäneet joitakin soluautomaatin solujen ominaisuustietoja hyväkseen. Ongelma näissä käyttötapauksissa on ollut se, että mallin ongelmaa on lähdetty tarkastelemaan heikkouksien kompensoinnilla eikä teoreettisella pohjalla. Tämä tarkoittaa, että kaupunkien dynamiikkaa ja erityisominaisuuksia ei ole huomioitu ja niiden pohjalta rakennettu mallia. Kummassakin tekniikassa, soluautomaateissa ja agenttipohjaisissa systeemeissä, on omat ongelmat ja heikkoudet, jotka ilmenevät kaupunkimallinnuksessa. Seuraavassa kappaleessa kerrotaan tarkemmin näistä heikkouksista, mutta myös aiemmissa luvuissa on yleisellä tasolla kerrottu tekniikoiden heikkouksista ja miten kyseisiä heikkouksia on huomioitu erinäisissä malleissa.

Soluautomaattien suurin heikkous kaupunkimallinnuksen näkökulmasta on solujen formaali ja joustamaton tapa kuvata tilaa. Tasainen ruudukko on helppo tapa jakaa tila

osuuksiin, mutta se ei välttämättä vastaa sitä jakoa, mikä oikeassa maailmassa on tilanteena. Kuten soluautomaatteja koskevassa luvussa kerrottiin, on kaupunkimallinnuksessa ollut tyypillistä käyttää epätasaista jakoa soluissa, kun on pyritty kuvaamaan tilaa esimerkiksi naapurustoina, jotka eroavat toisistaan. Tasaisesta ruudukosta luopuessa on huomioitava taasen naapuruussuhteet ja läheisyyden periaatteen pohjalta se, mitkä solut ovat keskenään vaikuttavia. Naapuruston määrittelyssä on ollut käytössä etäisyys-, vierekkäisyys- ja yhdistyvyyspohjaisia tulkintoja.

Solujaon ja naapuruussuhteiden ohella iso heikkous soluautomaateissa on ollut niiden kykenemättömyys kuvastaa liikkuvia objekteja dynaamisessa ympäristössä. Kaupunkimallinnukselle oleellista on objektien liike muutoinkin kuin vain vierekkäisiin soluihin.

Agenttipohjaisten systeemien pääasiallinen ongelma niiden kehityksessä on ollut spatiaalisuuden ja tilan puute. Vaikka sosiaalisten tieteiden tutkimuksissa on pyritty kehittämään agenteista inhimillistä käytöstä kuvaavia dynaamisia päätöksentekijöitä, ei paikkatietoa ole otettu huomioon agenttipohjaisissa malleissa sosiaalitieteiden alalla. Tällöin myös agenteille luodut työkalut ovat olleet epäspatiaalisia. (Benenson & Torrens 2004) Stanilovin (2012) mukaan useissa agenttipohjaisissa malleissa tilaa ja ympäristöä ei pidetä merkittävänä tekijänä mallissa. Tila on pikemminkin useille agenteille ympäristö vaikuttaa toistensa kanssa. Toinen ongelma on ollut luotujen agenttipohjaisten mallien yleisluontoisuus. Koska agentit kykenevät toimimaan useissa ympäristöissä ja suorittamaan lukuisia eri toimintoja, on luotujen mallien yleisluontoinen kuvaus toiminut niitä vastaan kun agenttipohjaisia malleja on yritetty saada toimimaan kaupunkimallinnuksessa. (Benenson & Torrens 2004) Yleisluontoisuuden lisäksi useat agenttipohjaiset mallit on luotu tiettyä, spesifiä tarkoituserää varten. Tällöin malli suoriutuu hyvin siinä ympäristössä, johon se on luotu. Toisaalta tällöin mallin implementointi muita simulointiskenaarioita varten ei ole helppoa.

Suunnitellessa agenttipohjaista systeemiä, kohdataan ongelma mallin tavassa yleistää todellisuutta. Kun pyritään ratkaisemaan ongelmaa simuloimalla tutkittavaa ympäristöä, joudutaan tekemään abstraktioita. Tällöin mallin kehittäjän tehtäväksi jää päättää millainen abstraktion taso on riittävä, jotta malli vastaa tarpeeksi todellisuutta tarkkuudeltaan, mutta on myös samalla kompleksisuudeltaan ratkaistavissa. Kyseinen tasapaino tulee kyseeseen, kun suunnitellaan mitä ominaisuuksia tahdotaan mallintaa attribuuteilla ja miten paljon mallissa käytetään agenteja. (Heppenstall et al. 2012)

## 6 Geosimuloinnin uusien tutkimuksien arviointi

Tässä osiossa työtä tarkastellaan kahta eri ajankohtaista geosimulointiin liittyvää tutkimusta. Tavoitteena on ottaa selvää siitä, millä tavoin kumpikin tutkimus hyödyntää aiemmin tarkasteltuja geosimuloinnin piirteitä ja hyödyntää niitä eri käyttötarkoituksiin. Erityisesti kiinnitetään huomiota, miten tutkimukset ja niiden kehittämät mallit huomioivat soluautomaattien ja agenttien ominaisuudet ja ratkaisevat tyypillisiä ongelmia, jotka liittyvät kyseisiin metodeihin. Töiden valinnassa huomioitiin niiden käyttämät tekniikat, ajankohtaisuus sekä sovellusala. Vertailun kannalta valittiin töitä, joista kumpikin keskittyy urbaaniin kaupunkimallinnukseen, mikä on geosimuloinnin tyypillisin sovelluskohde. Toinen motivaatiokohde töiden vertailulle ja niiden arvioinnille on huomioida, miten geosimulointimallien kehitys on vaikuttanut käytettyjen tekniikoiden kehittymiseen. Kyseinen tietämys on arvokasta jatkokehitykselle.

Arviointia varten on ensin huomioitava, mitä haasteita ja heikkouksia geosimulointimalleilla tyypillisesti on ja miten tarkasteltavat mallit huomioivat nämä. Geosimuloinnin haasteita on käsitelty jo aikaisemmissa luvuissa, mutta tässä luvussa ne kerrataan ja koostetaan arviointia varten listaksi. Toisekseen mallien edistyneempiä tekniikoita arvioidaan sen kannalta, miten ne edistävät perusmallin tekniikoita. Koska tämän kappaleen tavoitteena on ottaa selvää nykyaikaisista tekniikoista, on syytä vertailla, miten mallien käyttämät tekniikat eroavat työn aiemmissa kappaleissa esitellyistä peruseräilyistä.

Ensimmäisessä luvussa esitellään ne geosimuloinnin haasteet ja heikkoudet, joihin arvioinnissa kiinnitetään huomiota. Seuraavaksi esitellään tarkasteltavat työt. Kolmannessa luvussa tehdään esitellyjen töiden vertailu niiden heikkoustehtävien suhteen, mihin on kiinnitetty huomiota. Neljännessä luvussa tehdään yhteenveto kummankin työn vertailun tuloksista ja suoritetaan pohdintaa aiheesta yleistasolla.

### 6.1 Tutkimuksien ja mallien arvioinnin kriteerit

Tässä osuudessa työtä on tarkoituksena tarkastella kahta nykyaikaista geosimulointiin liittyvää tutkimusta, huomioida mitä haasteita ja ongelmia nykyisin kohdetaan agentti- ja soluautomaattipohjaisissa geosimulointimalleissa ja miten kaksi tutkittavaa työtä ovat huomioineet kyseiset ongelmat.

Lähtökohdaksi arvioinnissa otetaan Andrew Crooks et al. (2007) työ Key Challenges in Agent-Based Modelling for Geo-Spatial Simulation. Työssä esitellään seitsemän ongelmakohtaa tai haastetta joihin voidaan törmätä agenttipohjaisissa simulaatiomalleissa. Soluautomaattien implementoinnin haasteet huomioidaan siinä määrin mitä Benenson & Marceau (2011) kyseisestä aihepiiristä käsittelevät kirjassaan.

Agenteille ehdotetut seitsemän haastetta ovat:

- mallin tarkoitus ja mitä varten se on kehitetty
- kuinka paljon malli pohjautuu olemassa olevaan teoriaan
- miten hyvin mallia voidaan käyttää eri tarkoituksiin
- miten mallia voidaan validoida, verifioida ja kalibroida
- miten mallin dynamiikka esiintyy agenttien välisinä interaktioina
- mallin toiminnallisuus
- mallin jakaminen

Seuraavassa kappaleessa käydään lävitse tarkemmin jokainen kohta ja millä tavoin haasteita hyödynnetään vertailussa.

### 6.1.1 Haasteet tarkemmin

Agenttipohjaiset mallit ovat lähtökohtaisesti hyvin avoimia ja geneerisiä lähtökohdiltaan, koska agentit ovat pikemminkin mallinnuksen ja simuloinnin tyyli kuin tarkka teoreettinen tai käytännöllinen malli. **Mallin tarkoituksen** selvittäminen on sen vuoksi tärkeätä, jotta voidaan arvioida, miten hyvin tällainen avoin ja moneen käyttötarkoitukseen soveltuva tekniikka täyttää mallin tarpeet. Agenteja voidaan käyttää malleissa joko esittämässä tai ennustamassa ilmiötä ja toisinaan tuottamassa yllättäviä tuloksia.

**Mallien pohjautuminen tunnettuun teoriaan** ja teorian testaaminen malleissa on vähentynyt geosimuloinnin alalla. Agenttipohjaiset mallit pyrkivät pikemminkin edistämään aikaisemmin kehitettyä mallia eri käyttöympäristöissä ja –tilanteissa lisäten siihen samalla pieniä muutoksia. Tämän vuoksi malleja voidaan pitää geneerisinä, koska niitä ei ole sidottu pohjalla olevaan teoriaan. Tämän kaltainen *ad hoc* mallintaminen on aiheuttanut teoreettisen pohjan heikentymisen malleissa. Mallin jatkokehityksen ja tieteenalan kehittymisen kannalta on olennaista, että mallilla on olemassa teoriapohja, joka sitoo sen tietentekijöiden konseksukseen. Varsinkin kaupunkimallien kohdalla malliin voidaan liittää osioita useista eri tieteenaloista kattavamman ja suurempiskaalaisen mallin myötä, jolloin yhteisen ymmärryksen saavuttamiseksi kompleksisissa malleissa on teoriapohjan huomioiminen arvokasta.

**Mallin soveltaminen muihinkin käyttötapauksiin tai –tarkoituksiin** vahvistaa luottamusta mallin tuomaan uutuusarvoon sekä tieteelliseen kontribuutioon. Jos malli on liian monimutkainen ja yrittää kontrolloida suurta määrää muuttujia, on malli yleensä rakennettu vain tiettyä käyttökohdetta varten. Tästä johtuen mallin soveltaminen muihin skenaarioihin on hankalampaa. Tästä johtaa myös se, että mallien välinen vertailu tulee hankalammaksi, koska hyvinkin samanlaiset mallit käyttävät saatavilla olevaa dataa moninaisin tavoin. Mallin soveltaminen liittyy myös teoriapohjaan, koska perinteisesti ajateltuna malli toteuttaa saatavilla olevaa teoriaa ja osoittaa sen paikkansapitävyyden, jos mallia voidaan soveltaa useisiin eri tilanteisiin.

Malli voi olla tuottamatta konkreettisia tuloksia. Osa malleista rakennetaan vain kuvaamaan ongelmaa tai ilmiötä, mutta suurin osa agenttipohjaisista malleista tuottaa sellaista dataa, jota voidaan käyttää kalibroinnissa ja testaamisessa. Tässä työssä on aikaisemmin jo käsitelty **verifioinnin** ja **validoinnin** eroja, mutta lyhyesti verifiointilla testataan onko logiikka mallin taustalla paikkansapitävä. Validoinnilla testataan taasen miten hyvin malli sopii dataan ja testaus tapahtuu vertailemalla mallin tuottamaa dataa vertailuaineistoon. **Kalibroinnissa** taasen mallin parametreja ja arvoja hienosäädetään sillä tavoin, että malli toimii optimoidusti ja ennaltamääritellyn laadun mukaisesti. Ongelmia syntyy agenttipohjaisissa malleissa niissä tapauksissa, kun pyritään mallintamaan ihmiskeskeisiä systeemeitä, joihin kuuluu ennalta-arvaamaton käyttäytyminen. Tästä johtuen validointia varten vaadittaisiin mittava määrä dataa kattamaan mallin heterogeenisyyden prosesseissa.

Malleissa ilmentyvien **agenttien välisten interaktioiden** ongelma liittyy agenttien määrittelyyn. Jokainen malli tulkitsee agenttien käsitteen eritavalla ja tällöin agentin määrittelyssä syntyy eroja siinä, mitkä entiteetit tulkitaan malleissa agenteiksi. Yksi mahdollinen rajoite agenttien määrittelylle on se, että agentilla on oltava kyky liikkua

tilassa. Kun määritellään agentteja joko ryhmänä, yksittäisinä toimijoina, abstrakteina määritteinä tai kokonaisina objektien joukkona, voidaan kadottaa näkemys siitä, mitkä ovat relevantteja prosesseja ja interaktioita, joita vaaditaan agenteilta. Toinen ongelma liittyy agenttien kattamaan skaalaan maailmassa. Mitä yksityiskohtaisemmaksi agentit tulevat, sen monimutkaisempia systeemit ovat. Vaikka tietoteknisissä ratkaisuissa voi löytyä riittävästi prosessitehoa, on tutkijoiden ja kehittäjien ymmärryksen kannalta tarpeellista, etteivät agenttipohjaiset mallit kasva liian monimutkaisiksi. Agenttimalleja ei voida myöskään tehokkaasti tutkia näytepohjaisesti, koska monimutkaisen systeemin ymmärtäminen alkutekijöiden tasolla ja sen jälkeen sen yksinkertaistaminen ei ole vielä kehittynyt varteenotettavaksi metodiksi.

Käytetty ohjelmisto ja pohjamalli määrittelevät sen miten paljon aikaa kehittäjän ja tutkijan on käytettävä mallintamisessa muuhun kuin konkreettiseen tutkimustyöhön. Yksi etu generisimmissä agenttipohjaisissa malleissa on se, että niitä voidaan **soveltaa ja adaptoida** eri käyttötarkoituksia varten, eikä tutkijoiden välttämättä tarvitse lähteä rakentamaan mallia täysin tyhjästä. Ohjelmistojen tarjoamien työkalujen myötä kyetään sisällyttämään malliin monimutkaisia, mutta yleisiä osia moduulimaisesti, jotka esiintyvät tyypillisesti agenttipohjaisissa malleissa. Geneerisemmät mallit mahdollistavat suuremman yhteisön käyttöönottoa mallin ja implementoida sitä eri tarkoituksiin.

Viimeinen haaste, joka mainitaan, liittyy agenttipohjaisten **mallien jakamiseen ja tiedottamiseen**. Tuloksien lisäksi mallin sisäisen rakenteen, prosessien, syötteiden ja kalibroinnin esittäminen esimerkiksi visuaalisin keinoin antaa mallille lisäarvoa. Tällöin voidaan myös informoida mallista henkilöille, jotka eivät välttämättä ole alan asiantuntijoita, mutta voivat silti tuoda lisäarvoa tietämyksellään muilta aloilta. Mallin jakamisessa eteenpäin voidaan käyttää internetin eri portaaleita, joissa käyttäjät voivat kohdata ja jakaa luomiaan malleja käytettäväksi eteenpäin.

### 6.1.2 Soluautomaattien haasteita

Ulam, von Neumannin, Conwayn ja Wolframmin määrittelemä perusmuoto soluautomaatista ei ole säilynyt alkuperäisessä muodossaan, kun soluautomaatteja on hyödynnetty kaupunkimallinnuksessa. Adaptaatiot mallista ovat muokanneet soluautomaatin naapurustoja, tilasääntöjä, solujen kokoa ja muotoa ja jopa soluautomaatin ajallisia ulottuvuuksia. Tällöin on syntynyt kysymys siitä, että onko kaupunkimallinnuksessa käytetty soluautomaatti ollenkaan enää formaalisti soluautomaatti.

Soluautomaattien muuttumisen myötä useita erilaisia malleja ja adaptaatioita on syntynyt. Vertailu tässä työssä pohjautuu haasteisiin, joita Benenson ja Marceau (2011) sekä Torrens ja O'Sullivan (2001) käsittelevät töissään. Ensinnäkin **soluautomaatin spatiaalinen skaala** rajoittuu solun kokoon ja sen naapurustoon. Ilmiöt, jotka vaikuttavat laajalla alueella tai esimerkiksi kaupunkimallinnuksen tapauksessa kaupunkien välillä, huomioidaan muiden metodien kautta leviävänä ilmiönä. Homogeeniset, rajatut solut eivät ole siis välttämättä riittäviä osoittamaan ilmiön leviämistä eri spatiaalisilla etäisyyksillä. Tällöin joudutaan turvautumaan solujen naapurustoissa erilaisiin ratkaisuihin.

Toinen ongelma liittyy **soluautomaatin tilasääntöihin ja olemassa olevaan teoriaan**. Kaupunkien kasvusta on kehitetty useita teorioita eri tieteenaloilla, joten teorian sitominen sääntöihin edistää teorian tutkimusta. Pääsääntöisesti tilasäännöt perustuvat

ennemminkin mallinnuksen metodeihin ja hyvän mallin rakentamiseen sen sijaan, että otettaisiin huomioon teoreettinen pohja.

Kolmas ongelma on **soluautomaattimallien validointi**. Validointi on pääasiassa perustunut hahmopohjaisiin tekniikoihin ja niiden yhteen sopimisen hyvyyteen kuten khiin neliö-testiin tai kappi-testiin. Ongelma kyseisissä testeissä on siinä, että ne eivät ota huomioon soluautomaatin muita ominaisuuksia kuin hahmopohjaiset ilmiöt. On ehdotettu, että validoinnissa voitaisiin käyttää Monte Carlo –keskiarvoistamista, kompleksisuuden mittaamista tai spatiaalisuuteen liittyviä lukuarvoja. Yksinkertaisempaan validointiin voidaan käyttää soluautomaatin vertailua olemassa olevaan aineistoon.

Kahden tutkimustyön arviointi asetettujen kriteerien mukaisesti taulukoituna.

**Taulukko 1 Suoritetun arvioinnin tulokset agenttipohjaisille malleille**

	<i>Agent Based Simulation for Understanding Urban Dynamics</i>	<i>Geosimulation of urban growth and demographic decline in the Ruhr</i>
Mallin tarkoitus	Kaupungin dynamiikan ja sisäisten systeemien ymmärtäminen.	Kutistuvan kaupungin kehittymisen simulointi AI-tekniikoita hyödyntäen.
Pohjautuminen teoriaan	Osittainen keskuspaikkateoria.	Nykyiset teoriat kaupunkien, väestön ja talouden geografiasta.
Replikoitavuus	Parametreilla muunneltavat skenaariot eri alueille.	Suunnattu tietyille alueille Euroopassa.
Validointi, verifiointi, kalibrointi	Mallissa ei ole käytössä oikean maailman dataa verifiointia tai kalibrointia varten.	Kalibrointi ja validointi sekä UGMr että ReHoSh mallissa.
Agenttien välinen dynamiikka mallissa	Reaktiivisia ja proaktiivisia agentteja.	ReHoSh mahdollistaa agenttien väliset interaktiot.
Toiminnallisuus mallissa	NETLOGO ja Cobb-Douglas funktion parametrien muokattavuus.	UGMr-SVM ja ReHoShin loose coupling ja kummankin tekniikan parametrien muokkaaminen.
Mallin jakaminen ja kommunikaatio	Kevyt selitys mallin perustoiminnasta.	Yksityiskohtainen raportti koko mallin toiminnasta.

Taulukossa 1 on esitetty yhteenveto agenttipohjaisten mallien asetetuista tutkittavista arvioinnin kriteereistä sekä siitä, miten kumpikin esitetty tutkimus oli ongelmat ratkaissut tai käsitellyt niitä työssä.

Soluautomaattien osalta vastaava yhteenveto on taulukossa 2.

**Taulukko 2 Soluautomaattien kriteereiden tulokset**

	<i>Agent Based Simulation for Understanding Urban Dynamics</i>	<i>Geosimulation of urban growth and demographic decline in the Ruhr</i>
Spatiaalisen skaalan käsittely naapurustossa	Mooren naapurusto ja NetLogon vaikutus kiinnostavuuteen.	SVM tukemassa UGMr:a.
Tilasääntöjen ja teorian yhteneväisyys	Tilasääntöjä ei ole sidottu teoriaan.	UGMr-SVM sitoo teorian kaupunkien kehittymisestä solujen kehitykseen.
Validointi	Ei validointia.	Validointi suoritetaan.

## 6.2 Tutkittavien töiden esittely

### 6.2.1 Työ 1: Agent Based Simulation for Understanding Urban Dynamics, Kiranjith, Jonu John Thomas Giridhar K. 2016

Työssä käsitellään kaupunkien kasvun analysointia ja mallinnusta. Motivaatio työlle oli se tosiasia, että ihmiset asuvat entistä suuremmassa määrin kaupungeissa ja väestön määrän kasvun myötä myös kaupunkialueiden kokonaisala tulee kasvamaan. Maankäytön suunnittelun ja hallinnan avulla voidaan vaikuttaa paremmin siihen, millä tavoin kaupungit kehittyvät. Käytännön ongelmat, jotka syntyvät hallitsemattomasta kaupunkialueiden kasvusta ovat esimerkiksi asuntojen puute ja toimimaton infrastruktuuri. Kuten Johnson (2002) totesi, on kaupunki suurempi kuin osiensa summa. Tämän johdosta ei pystytä helposti ennustamaan kaupungin kehitystä kokonaisuutena, vaikka yksittäisten palasien ja osa-alueiden kehittymistä voitaisiinkin ennustaa.

Kaupunkien mallinnuksessa ja niiden toiminnan simuloinnissa pohjimmainen periaate on mallintaa yksilöiden toimintaa ja tätä kautta useiden yksilöiden muodostaman kokonaisuuden käyttäytymistä. Osasy tällaiseen yksilökeskeiseen lähestymistapaan on se, ettei ihmisen käyttäytymistä pystytä mallintamaan samalla tavalla kuin oikean maailman ilmiötä perinteisin menetelmin. Aikaisemmin kehitetyt kaupunkimallit keskittyivät tarkemmin kaupungin toiminnan ymmärtämiseen pikemminkin kuin analysointiin ja mallintamiseen. Vaikka kaupunkien toimintaa ja sen osien dynamiikkaa opittiin ymmärtämään, ei saatua tietoa kyetty hyödyntämään suunnittelussa. Tästä johtuen uudemmat kaupunkimallit on kehitetty suunnittelijoiden, päättäjien, rakennuttajien ja poliitikkojen työkaluiksi päätöksentekoa tukemaan.

Kun on tutkittu yleisiä ilmiöitä ja ongelmia, jotka ilmaantuvat kaupungin kasvun myötä, kuten asuntojen riittämättömyys, maankäytön tehokkuuden hallinta, liikenteen ruuhkautuminen ja infrastruktuurin riittävyys, on huomattu, että yksilöiden käyttäytymisen muodostamalla kokonaisuudella on merkittävä vaikutus. Tämän johdosta ongelmaa voidaan pitää kompleksisena. Kompleksisen ilmiön mallintamista varten työssä hyödynnetään agenteja. Agentit on huomattu tehokkaaksi tavaksi mallintaa ympäristössä toimivia yksilöitä ja kyseinen tapa mallintaa kaupungeja on osoittautunut suosituksi.

Työn keskeinen tavoite on luoda malli, joka yhdistää parviälyä sekä agenteja. Parviälyä on aiemmin käytetty luonnollisten ilmiöiden mallinnuksessa, mutta kaupunkimallinnuksessa kyseistä tekniikkaa ei ole tutkittu kattavasti. Toisekseen pyritään tuottamaan algoritmi, joka huomioi sekä kvalitatiiviset että kvantitatiiviset stigmergiat

agenttien välillä. Stigmergia on viestinnän muoto, jossa informaatiota välitetään ympäristöä muuttamalla. Stigmergiassa yhden toimivan yksilön toiminta vaikuttaa epäsuorasti muiden yksilöiden toimintaan. Tällöin syntyy epäsuora vaikutussuhde ryhmän yksilöiden välille.

Parviällyn ja agenttien yhdistämisessä käytetään kehitettyä algoritmia. Agenttien käytöksen ja päätöksenteon ohjaamiseksi algoritmin lisäksi käytetään laskennallisia sääntöjä. Simulointia varten hyödynnetään myös elementtejä geosimulointimallista, jossa Torrens (2006) yhdisti samaan malliin soluautomaatteja, agenteja sekä GIS:n. Työssä tehdään muutama oletus mallin suhteen. Agenttien oletetaan toimivan tehokkaasti ja rationaalisesti valintojensa suhteen. Tämä tarkoittaa, että agentit pyrkivät maksimoimaan hyötynsä ja toimintansa sillä tavoin, että saavutetaan paras arvo maalle ja asunnoille. Jokaisella agentilla on myös jonkinlainen asunto, joka on riittävän lähellä työpaikkaa sekä hinnaltaan kohtuullinen.

Urbaanin ympäristön muodostaa mallissa NetLogo ja sen perusmallin pohjalle rakennetaan muut komponentit. NetLogon kehittäneet Liévano (2001), ja Felson sekä Wilensky (2007) sisällyttivät perusmalliin kaksi pääkomponenttia: teollisuus ja työvoima. Kaupungin kasvun kannalta tärkeimpinä komponentteina voidaan pitää sen asukkaita sekä taloutta, ja niiden kehittymistä rajoittavat luonnolliset resurssit. Tässä tapauksessa luonnollisena resurssina voidaan pitää maa-alaa.

Maata mallinnetaan NetLogolla, joka jakaa mallinnettavan alueen ruudukkoon. Jako muistuttaa samanlaista jakoa, joka tehdään soluautomaateilla. Jokainen ruutu tai solu, on tilaltaan joko kehittynyt tai kehittymätön. Solulla on myös joukko attribuutteja: kehittyneen maan arvo, kiinnostavuus eli vetovoima, hinta, rakennuksien lukumäärä maa-alueella sekä etäisyys lähimmästä työkeskuksesta.

Mallissa on vain kahdentyyppisiä agenteja: kehittäjiä ja talouksia. Talous-agentit etsivät asumusta ja ne toimivat dynaamisesti sekä reaktiivisesti. Kehittämis-agentit taas muuntavat käyttämätöntä maata joko asumukseksi tai teollisuudeksi. Lisäksi kehittäjät on jaettu sekä alhaiselle että korkealle tulotasolle. Kehittäjä-agenteilla on mahdollisuuksia suorittaa toimintoja, kuten metsän hakkuu, maiseman ehostaminen, asumistiheyden kasvattaminen ja maan arvoon vaikuttaminen. Tällöin kummatkin agenttityypit toimivat keskenään epäsuorassa vaikutussuhteessa, jossa kehittäjä-agentti vaikuttaa maan arvoon ja asuntojen määrään ja talous-agentit reagoivat kyseiseen muutokseen.

Kehittäjä-agenttien käyttäytymistä ohjaa kaksi eri tekijää: Cobb-Douglas funktio sekä satunnaisuustekijä toiminnan suhteen. Vaikka kehittäjät pyrkivät käyttämään toimintaansa mahdollisimman tehokkaasti, on agenteille määritelty satunnaisuutta ja epätäydellistä toimintaa toteuttavat funktiot. Tästä johtuen kehittäjä-agentit eivät pyri kehittämään jokaista solua, vaikka niillä on tietämys jokaisen solun tilasta, vaan kehittäjät preferoivat tiettyjä soluja hyödykeperusteisesti.

Mallissa on soluautomaatin ja agenttien lisäksi joukko parametreja, joita käyttäjä voi hyödyntää simulointituloksen tarkkuuden parantamisessa tai eri skenaarioiden luomisessa. Parametrit ovat väestön kasvun määrä, töiden määrä, jonka tietty teollisuudenala luo, simulointikerran aikana luotojen töiden kokonaismäärä sekä yhden talouden henkilömäärä.

Simulointi suoritetaan askel kerrallaan ennaltamääriteltyjen ajanmuutosten  $\Delta t$  välein. Jokaisella ajanhetkellä suoritetaan seuraavat toimenpiteet:

- Kysyntä asunnoille kasvaa, koska väestömäärä kasvaa.
- Kehittymättömiä maa-alueita kehitetään kehittäjä-agenttien toimesta.
- Kehitetyn maa-alueen solun attribuuttitiedot päivitetään.
- Riippuen kysynnästä alueella, maa-alueet otetaan käyttöön ja maan hinta päivitetään.

Kohdistetut muutokset suoritetaan soluihin soluautomaatin naapurustoa käyttämällä. Solujen tila voi vaihtua myös, jos niiden naapurustossa olevan solun tila tai attribuutit vaihtuvat. Jos solulle on määritelty kyseinen vaikutussuhde, vaihtuu solun attribuutit tai tila, jos sen naapurustossa olevan solun tila tai attribuutit muuttuvat. Kyseinen spatiaalinen riippuvuus vaikuttaa tällöin maan hintaan tehtyjen maankäyttöön liittyvien päätöksien lisäksi.

### **6.3 Työ 2: Geosimulation of urban growth and demographic decline in the Ruhr: a case study for 2025 using the artificial intelligence of cells and agents. Rienow, Andreas. Stenger, Dirk. 2014.**

Työssä tutustutaan Ruhrin alueen urbaaniin kehittymiseen. Vanha teollisuuskaupunkien alue on aikoinaan rakennettu monofunktionaaliseksi teollisuuden ja kaivostoiminnan keskittymäksi, mutta on sittemmin kehittynyt suureksi asuinkaupunkien alueeksi. Ilmiötä kutsutaan kaupunkialueen hajoamiseksi, koska kaupungit ovat kehittyneet nopeasti jättäen rakennetut asuinalueet väljiksi ja väestötiheydeltään alhaisiksi. Kyseisen ilmiön mallintaminen perinteisin tilastotieteisiin perustuvilla menetelmillä ei tuota tarpeeksi tarkkoja malleja kompleksisuutensa vuoksi. Aihe on myös ajankohtainen, koska kyseinen samanaikainen kasvu ja teollisuuden vähentyminen tapahtuu muissakin vanhoissa Euroopan kaupungeissa.

Mallissa käytetään loose coupling -menetelmää, jossa mallinnetaan kaupungin kasvua sijaintiin perustuvilla tekijöillä sekä alueiden välisillä asuntomarkkinoilla. Ruhrin alue koostuu 15 kaupunkialueesta, joten mallissa ei keskitytä pelkästään yhden kaupungin mallinnukseen. Perimmäinen metodi mallinnuksessa on soluautomaatti SLEUTH. SLEUTH ottaa huomioon aikaisemmat kaupungistumisilmiön hahmot ja hyödyntää niitä kasvuasteen kertoimien laskennassa. Mallin tarkennukseksi käytetään koneoppimisen SVM -tekniikkaa makrotasolla maa-alueiden sopivuuden ja kannattavuuden laskemisessa.

Kaupunkialueiden dynaamisten muutoksien seuraamiseksi ja aiheutuneiden vaikutuksien kartoittamiseksi käytetään ReHoSh (Residential Mobility and the Housing Market of Shrinking City Systems) metodia, jossa otetaan huomioon yksittäisten talouksien mieltymykset että myös potentiaalisten asumisalueiden kehitys sekä saatavilla olevat asunnot. Sekä SLEUTH-SVM että ReHoSh toimivat erikseen ja niiden tuottamat tulokset yhdistetään mallissa. Kumpikin metodi tuottaa dataa eri skaaloissa, jolloin yhdistämisessä on otettava huomioon laaja-alainen skaalaus tuotetulle datalle.

Tutkittavat ongelmat ovat:

1. Ruhrin alueen kaupunkialueiden kasvun spatiaalisesti tarkka mallinnus vuoteen 2025 saakka käyttäen SLEUTH-SVM –metodia soluautomaattina.
2. Alueen keskuksien ja asuinalueiden välinen väestön liikkuminen ja asuntojen hintojen kehitys agenttipohjaisella ReHoSh –metodilla.
3. Kummankin AI pohjaisen metodin yhdistäminen loose coupling –tyyppisesti ja tuloksien analysointi asuntojen hintojen ja tyyppien jakautumisesta spatiaalisesti.

Mallin soluautomaatti perustuu Clarke et al. (1997) SLEUTH-malliin, joka on suunniteltu mallintamaan kaupungin kasvua. SLEUTH:a kutsutaan myös UGM-malliksi (Urban Growth Model). SLEUTH tarkoittaa kuutta attribuuttia, jotka malli ottaa syötteenä: kaltevuus, maankäyttö, kiellot, liikenne ja rinnevarjostus. Näiden lisäksi on viisi kasvun osatekijää, jotka määrittävät neljä kasvusääntöä: spontaani kasvu, uusi laajeneva kasvukeskus, kasvu kaupunkialueen laidoilta kuvastaen kaupunkialueen hajoamista sekä teihin perustuva kasvu. Työssä mallia mukautettiin R. Goetzken eXtendable Unified Land use Modeling Platform (XULU) mallilla. XULU muuttaa kalibrointivaiheen siten, että viiden syötteenä otetun attribuutin sijasta käytetään vain kahta. Kokeissa on huomattu, että yksinkertaistettu malli tuotti paremman tuloksen. Mallissa ilmenevää satunnaisuutta kompensoitiin hyödyntämällä Monte Carlo iteraatiota.

Yksi aikasykli SLEUTH-mallissa vastaa yhtä vuotta, jolloin tapahtuu aiemmin mainitut neljä kasvusääntöä. Jokaista uutta valittua solua verrataan paikan kaltevuuteen, kieltoihin sekä satunnaisarvoon. Kasvusääntöihin liittyvät kasvuosatekijät testataan kalibrointivaiheessa ja niiden jokainen parametrikombinaatio saa arvon välillä 0-100. Kalibrointia jatketaan kunnes sopiva tasapaino on saavutettu.

Mukautetun SLEUTH-mallin ohjaamiseksi työssä hyödynnettiin SVM:a. SVM on oppimisalgoritmeja sisältävä malli, jota käytetään yleistämiseen, luokitteluun ja regressioanalyysiin. Merkittävin hyöty SVM:a on sen kyky ratkaista epälineaarisia luokitteluongelmia. Työssä SVM:a käytettiin useiden sosioekonomisten, geofysikaalisten ja muiden vetovoimatekijöiden sekä ilmiöiden adaptointiin SLEUTH-mallin ohjaamiseksi. Ilmiöistä kerätty ajankohtainen data koskien kaupunkien hajoamista koostettiin rasteritasoiksi. Käytetty ohjelmisto valitsi tärkeimmät tekijät kaupungin kasvun kannalta ja käytti vähemmän merkittäviä tekijöitä datan kouluttamiseen. Tuloksena on todennäköisyyskartta siitä, mille alueille mahdollinen kaupungin kasvu ja kehittyminen suuntautuu. Mukautettua SLEUTH/UGM-mallia kutsuttiin työssä UGMr-malliksi. Yhdistettyä UGMr-SVM mallia kalibroitiin ja validoitiin vertaamalla ennakoitua kasvua toteutuneeseen kasvuun tietyllä alueella.

Kaupunkialueen hajoamisen mallinnusta varten jouduttiin käyttämään myös agenteja, koska UGMr-SVM malli mallintaa vain kaupungin kasvua. Agentit tässä työssä kuvastavat abstrakteja entiteettejä, kuten yhteisöjä sekä kotitalouksia. Väestön määrän vähentymiseen keskittyviä malleja ei ole kehitetty paljoa eikä varsinkaan Keski-Euroopan kaupunkialueiden muuttamiseen keskittyneitä malleja tai tutkimuksia. Työssä valittiin käytettäväksi ReHoSh-mallia, joka mallintaa alueidenvälisten asuntomarkkinoiden dynamiikkaa. Kyseisen ilmiön dynamiikat antavat mahdollista tietoa kehittyvistä asuinalueista. Mallissa simuloidaan kolmea ilmiötä: väestön jakautumista, asuntojen hintoja sekä asuntojen kysyntää väestöllisesti heikkenevässä kaupungissa. Malli on nykyaikainen ja käyttää hyväkseen teorioita sekä empiirisiä tutkimuksia nykyaikaisesta väestön ja kaupunkialueiden dynamiikasta.

ReHoSh:ssa on kuusi oletettavaa. Ensinnäkin, simulaatiossa huomioon otetut interaktiot rajoittuvat temaattisesti vain asumuksiin liittyviin. Toisekseen, kaikilla kotitalouksilla on tietämys kaikista mahdollisista asumuksista. Tällöin etäisyystekijät eivät vaikuta asuntoa etsivän tietämykseen, vaan kaikki tutkimusalueen saatavilla olevat asunnot riippumatta kaupungin rajoista on kaikkien tiedossa. Muita oletettavia ovat asuntojen hintojen lineaarinen muutos, hinnat eivät jousta asuntojen kysynnän mukaan, saatavilla oleva asuntojen määrä on simulaation alussa oletettu eikä perustu oikeaan tilanteeseen ja viimeiseksi simulaatiossa oletetaan Ruhrin alueen kokonaisväestömäärän vähenevän. Näiden oletettavien lisäksi kotitalouksien rakenne oletetaan pysyvän samana eli väestön rakenne ei muutu.

ReHoSh-malli voidaan jakaa kolmeen pääkomponenttiin: agentit, agenttien väliset interaktiot sekä ympäristö. Agentit ja agenttien väliset interaktiot sijaitsevat tutkittavassa ympäristössä, tässä tapauksessa se on Ruhrin ympäristö. Ympäristöön kuuluu sen demografiset piirteet sekä jokaisen agentin spatiaalinen sijainti. Kotitaloudet ja kaupungit alueella ovat taasen agenteja. Kaupungit käsittävät hallinnollisten rakenteiden lisäksi myös kiinteistöväliittäjät. Asunnot taasen omaavat erinäisiä ominaisuuksia, kuten hinta, laatu sekä saatavuus. Näistä ominaisuuksista muodostetaan saatavilla olevat asumisalueet sekä kiinteistöt. Kotitaloudet sekä kaupungit interaktivoivat kyseisten asuntojen kesken. Lisäksi sekä kaupunkeja että kotitalouksia voi olla useita erilaisia tyyppisiä eikä agenttien ryhmää voida pitää homogeenisenä. Kotitaloudet etsivät uutta asuntoa ja kilpailevat niistä keskenään. Kaupungit taasen reagoivat asuntojen etsintään muuntamalla asuntoihin liittyviä ominaisuuksia. Kaupungit reagoivat myös muiden kaupunkien toimiin. Ajanjakso on vuosi ja edellä kuvatut interaktiot ovat ainoita joita agentit voivat suorittaa ajanjakson aikana.

Agenttien käyttämisen perimmäinen tarkoitus on mallintaa ihmismäistä käytöstä päätöksenteossa. Työssä asuntoa etsivää agenttia ohjaa kolmivaiheinen päätöksentekoketju, joka on johdettu nykyaikaisesta teorioista liittyen väestön muuttamiseen ja kaupungeissa asumiseen. Vaiheet ovat: aikomus muuttaa eri asuntoon, uuden asunnon etsintä sekä päätös asunnon hankinnasta. Päätöksentekoon vaikuttavat etäisyys, asunnon tyyppi sekä sijainti. Näiden lisäksi työssä käytetään kaavaa kaupungin kokonaisvetovoiman määrittämiseksi:

$$totalA^t = \frac{priceF * priceA^t + quantityF * quantityA^t + qualityF * qualityA^t}{priceF + quantityF + qualityF} * households \quad (6)$$

Kaavassa  $totalA$  on kaupungin kokonaisvetovoima,  $priceA$  on hinnan kiinnostavuus,  $quantityA$  on saatavilla olevan tarjonnan kiinnostavuus,  $qualityA$  on saatavilla olevan tarjonnan laadun kiinnostavuus.  $priceF$ ,  $quantityF$  ja  $qualityF$  ovat taasen painokertoimia ja  $households$  on kaupungin kaikkien asuntojen määrä.

Sosiaalitieteistä johdetaan myös kaava asuntojen saatavuudelle sekä siitä aiheutuvalla hinnankkehitykselle. Näiden lisäksi työssä käytetään kaavoja asuntojen laadulle sekä naapurikaupunkien tekemien muutoksien vaikutuksille ympäröivien kaupunkien asuntojen saatavuudessa.

ReHoSh kalibroinnissa käytettiin Ruhrin alueelta kerättyä dataa asuinalueiden ja maankäytön kehityksestä. Ongelmaksi muodostui Ruhrin alueen uusien asuinalueiden

vähäinen määrä, joten kalibroinnissa ja validoinnissa käytettiin myös virallisia ennusteita väestön kehityksestä, maankäytöstä sekä asuntomarkkinoista. ReHoSh kalibroititiin vuodesta 2010 vuoteen 2025. Ainoa yliestimoitu tekijä kalibroinnissa oli kotitalouksien määrän vähentyminen.

## **6.4 Tutkimuksien arviointi**

### **6.4.1 Agenttipohjaisten metodien arviointi**

Kumpikin tutkimus hyödynsi erilaisia tekniikoita sekä soluautomaatin että agenttien esittämiseksi. Huomattavaa oli, että kumpikin tutkimus oli hyödyntänyt jo olemassa olevaa mallia ja mukauttanut sitä sopimaan kyseisen tutkimuksen skenaarioon. Kaupunkitutkimuksien osalta kumpikin työ kosketti erilaista aihetta. Kiranjith et al. (2016) pyrkivät työssään simuloimaan kaupungin kehitystä alueella, jossa on varallisuuseroja. Rienow & Strenger (2014) taas keskittyivät ennustamaan heikentyvästä kaupunkirakenteesta aiheutuvan kehityksen. Suuri eroavaisuus tutkimuksissa oli niiden käsiteltävässä skenaariossa. Kun Kiranjith et al. (2016) eivät suunnanneet tutkimusta mitään olemassa olevaa kaupunkialuetta kohtaan, oli taas Rienow & Strenger (2014) tutkimus suunnattu selkeästi Ruhrin kaupunkialueelle. Tällöin tutkimuksien lähtökohtainen tarkoitus on selkeästi erilainen. Voidaan myös olettaa, että Kiranjith et al. (2016) malli on luonteeltaan geneerisempi ja ei ota välttämättä huomioon jokaista yksityiskohtaa, vaan jättää jatkokehityksen huoleksi hienosäätää malli sopimaan käsiteltävään skenaarioon. Tästä johtuen myös tietyn kaupunkisysteemin täydellinen imitointi ei ole välttämättä tavoitteena, kun pyritään löytämään tuoreita hahmoja tai piirteitä aineistosta (Teodorovic 2010). Toisaalta, kumpikin tutkimuksesta liittyi aiheeltaan kaupunkien kasvuun. Kaupunkien väestön kasvaessa asuminen leviää laajemmalle alueelle aiheuttaen tarpeita maan, veden, sähkön ja infrastruktuurin suhteen. Kaupunkialueiden leviämisen myötä joudutaan ottamaan huomioon erilaiset maankäyttöön liittyvät säännökset sekä asuinympäristöjen muoto. Kun Kiranjith et al. (2016) malli keskittyy kasvavaan kaupunkiin, on Rienow & Strengerin (2014) malli taas suunnattu kaupunkiympäristöön, jossa kasvu on jo tapahtunut.

Kiranjith et al. (2016) pyrkivät löytämään työssään yllättäviä tuloksia kaupunkien kompleksisuudesta. Toinen uniikki piirre työssä on joukkoälyn konseptin hyödyntäminen. Teodorovic (2010) esitti, että sekä agentit että joukkoäly kuuluvat kumpikin keinoälyn tutkimuksen piiriin. Yleinen jako agenttipohjaisille systeemeille on jakaa agentit joko reaktiivisiin agenteihin tai tietämyksellisiin agenteihin. Joukkoäly taas nähdään joukkona pelkäästään reaktiivisia agenteja. Jako ei ole kirjaimellinen, sillä malleissa voidaan hyödyntää monenlaisia agenteja eikä tarkoituksena ole luoda tarkkaa jakoa. Joukkoälyn konseptia hyödynnetään Kiranjith et al. (2016) työssä jakamalla agenteja eri tyyppisiin niiden toiminnan ja interaktioiden perusteella.

Rienow & Strenger (2014) käyttävät työssään ReHoSh-mallin mukaelmaa. Toisin kuin Kiranjith et al. (2016), agenteja pidetään proaktiivisina soluautomaatteina, joilla on kyky liikkua, adaptoitua ja omaavat tietouden. Koska työssä mallinnetaan laskevan väestönmäärän monikeskuksisia kaupunkialueita, oli agenttisysteemin valinnan sovelluttava kyseiseen skenaarioon. Tällöin asuntomarkkinoiden dynamiikkaa mallintava ReHoSh on soveltuva valinta. Erityisvaatimuksena oli väestöryhmien pienentymiseen kykenevä malli, mutta myös useiden kaupunkikeskuksien välinen vaikutus oli oltava mahdollista ottaa huomioon. Malli oli myös sidottu nykyaikaisiin tutkimuksiin ja teorioihin kaupunkiväestön kehityksestä, joten malli oli ajankohtainen. Myös SVM

pohjautui teoriaan, koska SVM-soluautomaatti mallinsi niitä tekijöitä, jotka määrittävät lokaalisti kaupungistumisen soveltuvuutta.

Kun tutkitaan töiden teoriapohjaa, löytyy eroja teorian ajankohtaisuudessa. Kiranjith et al. (2016) työ ei sitoudu tarkasti mihinkään olemassa olevaan teoriaan, mutta keskuspaikkateoriaa sivutaan ja siitä on otettu työssä vaikutteita. Pääasiallinen vaikutus teoriolla on jakaa asumukset korkean- ja matalantulotason mukaisesti. Itse keskuspaikkateoriaa on kritisoitu siitä, ettei se sellaisenaan toimi jälkitekollistuneen ympäristön mallinnuksessa. Toisaalta Kiranjith et al. pyrkii työssä simuloimaan ja esittämään segregatiota naapurustojen ja tuloerojen välillä. Tällöin kyetään esittämään mallissa alueellisia jakoja, vaikka keskuspaikkateoriaa ei sellaisenaan kokonaan implementoida. Keskuspaikkateoria on Walter Christallerin vuonna 1933 esittämä teoria, joten se ei ole täysin ajankohtainen.

Rienow & Strenger (2014) taasen käyttävät mallissaan edellä mainittua ReHoSh-mallia, joka perustuu nykyaikaisiin teorioihin kaupunkien väestön kehityksestä ja asuntomarkkinoista. ReHoSh asettaa joukon eri oletettavia malleja, jotka on kuvattu aiemmin. Olettamat rajoittavat mallin kykyä ottaa huomioon kaikkia kaupunkikuvan hajoamiseen liittyviä ilmiöitä ja niiden taustoja, mutta samalla mallin käyttäjät kykenevät paremmin tarkkailemaan niitä tekijöitä, joista ollaan kiinnostuneita. Rajoittaminen liittyy mallin kompleksisuuteen. Malli on myös tehty simuloimaan ja ennustamaan tiettyä ilmiötä eikä yleispäteväksi simulointimalliksi tietylle kaupunkityypille. Tällöin voidaan rajoittamalla vaikuttavia tekijöitä ja yleistämällä ilmiöiden kompleksisuutta saavuttaa kontrolloidumpi ympäristö.

Agenttien interaktiossa mallit eroavat monimutkaisuudessa ja toiminnallisuudessaan. Kiranjith et al. (2016) jakaa agentit kahdeksi ryhmäksi, joista toinen on reaktiivinen. Reaktiiviset agentit etsivät asumusta ja reagoivat siihen, miten kaupunki ja kehittäjäagentit vaikuttavat asuntojen hintoihin. Kehittäjäagentit taasen pyrkivät toimimaan rationaalisesti ja mahdollisimman tehokkaasti.

Rienow & Strenger (2014) jakavat agenttisysteemin kolmeen eri lohkoon: agentteihin, interaktioihin sekä ympäristöön. Agentit jakautuvat taasen kaupunkeihin ja kotitalouksiin. Toisin kuin Kiranjith et al. (2016) työssä, kotitalousagentit etsivät uutta asuntoa ja kilpailevat keskenään. Kaupungit sen sijaan reagoivat kotitalouksien tekemiin päätöksiin sekä muiden kaupunkien tekemiin päätöksiin vaikuttamalla asuntoihin liittyviin tekijöihin kuten hintoihin, määrään ja laatuun. Selkeä ero Kiranjith et al. (2016) työhön on, että päätäntövallassa hierarkisesti alempana olevat agentit eli asukkaat, vaikuttavat toimillaan, johon päättävä korkean tason elin, eli kaupunki, reagoi. Tässä yhteydessä kaupunki käsittää myös kiinteistövälittäjiä ja -myyjiä.

Kiranjith et al. (2016) työssä ei kuvata validointia ja verifiointia, mutta NetLogossa on sisäänrakennettuna mahdollisuus suorittaa kyseiset vaiheet. Osaltaan verifiointia ei voida suorittaa, koska tutkittu malli ei perustu olemassa olevaan dataan, vaan on puhtaan abstrakti. Oikean datan puutteesta huolimatta mallia voidaan käyttää uusien näkökulmien tuomiseen tunnettuihin ilmiöihin ja jatkokehittämisen kannalta geneerisempi malli voi inspiroida tutkijoita. Kalibroinnin kannalta työssä esitetään parametreja, joiden käsittely simulointikertojen välissä antaa erilaisia tuloksia.

Rienow & Strenger (2014) käsittelevät työssään tarkasti kalibroinnin, validoinnin ja verifiointin sekä UGMr että ReHoSh-malleissa. Työssä on käytössä laajasti dataa Ruhrin

alueelta vuosikymmenien ajalta sekä virallisia ennustuksia väestön kehityksestä ja asuntojen hintojen kehityksestä muiden ennusteiden lisäksi. Malli oli tehty ennustamista varten, joten malli jouduttiin kalibroimaan vuoden 2010 datasta vuoteen 2025. Kaiken kaikkiaan validoinnissa voidaan havaita, että ReHoSh tuottaa kotitalouksien liikkuvuuksien ja asuntojen hintojen kehityksen 90% minimivarmuudella.

Operationaalisuus eli toiminnallisuus töissä esiintyy eri tavalla. Kiranjith et al. (2016) työssä esitetään joukko parametreja sekä Cobb-Douglas funktiolle että NetLogolle kaupunkiympäristön mallinnuksen kohdalla. Cobb-Douglas funktio määrittelee kehittäjäagenttien tuottavuutta sekä kuinka agentit preferoivat kehittäessään sekä alemman että ylemmän tulotason asuntoja alueella. NetLogo soveltuu hyvin mallin adaptoimiseen muihin tarkoituksiin, koska NetLogo tarjoaa joukon moduulimaisia lisäosia, joiden avulla kyetään kattamaan mallinnuksen eri käyttötärpeisiin tarpeellisia toiminnallisuuksia. Esimerkiksi vaikka Kiranjith et al. (2016) ei käsittele validointia ja verifikointia työssään, on NetLogossa tarpeelliset käyttöominaisuudet näitä varten. Lisäksi kun NetLogo huolehtii mallinnuksesta ja koostuu yhdestä kokonaisuudesta, jolla katetaan sekä ympäristön että agenttien mallinnus. Kyseinen kokonaisuus takaa helpomman implementoinnin, kun ei tarvitse huolehtia eri mallinnusmetodeiden yhteensovittamisesta.

Rienow & Strenger (2014) työssä taasen käytetään loose coupling-menetelmää yhdistämään ReHoSh-agenttisysteemi sekä UGMr. ReHoSh huolehtii agenttien muodostamisesta ja niiden välisestä interaktiosta mallintamaan keskuksien välistä muuttoilmiötä sekä asuntojen hintojen kehitystä. UGMr taasen mallintaa ympäristön ominaisuuksia keskuksissa soluautomaatin tekniikoiden avulla. Kummankin tekniikan tulokset yhdistettiin yhdeksi kartaksi käyttäen normaaleita spatiaalisia liitäntäsääntöjä. Tällöin menetelmät toimivat toisistaan riippumattomasti ja vasta lopussa kummankin tulokset yhdistettiin. Tiukemmin yhdistetyssä mallissa kumpikin menetelmä toimisi yhdessä jo mallintamisvaiheen aikana. Kumpikin käytetty menetelmä on tunnettuja ja testattuja, jolloin niiden soveltuvuus työn aiheeseen oli varmistettu. Toisaalta toiminnallisuuden kannalta menetelmät on sidottu vahvasti käytettyyn skenaarioon, jolloin niiden soveltaminen muihin tilanteisiin tuottaa käyttäjälle haasteita. Syy kahden eri menetelmän käytölle johtui Ruhrin alueen väestön vähentymisen ja kaupunkirakenteen hajoamisen ilmiöstä, jossa kaupunki samaan aikaan kasvaa sekä leviää ohuemmaksi. Ainutlaatuinen skenaario vaati tällöin ainutlaatuisen ratkaisun.

Viimeinen esitetty haaste agenttipohjaisille järjestelmille liittyy siihen kuinka mallia on jaettu ja miten siitä on kerrottu. Mallien jakamisesta ei ole tarkempaa tietoa, joten tässä kohtaa keskitytään vain siihen miten malleista on kerrottu kussakin työssä. Kiranjith et al. (2016) on töistä selkeästi kevyemmin selitetty ja käyty lävitse mallinnuksen eri vaiheet sekä yksityiskohdat. Malli on luonteeltaan geneerisempi, jolloin yksityiskohtainen käsittely ei ole välttämättä työlle tarpeellista. Esimerkiksi kun työssä mainitaan algoritmit agenttien yhteydessä, ei työssä avata tarkemmin, millä tavoin algoritmit vaikuttavat agenttien käytökseen. Työn tuloksia esitellään lyhyesti muutaman kuvan kanssa. Kiranjith et al. (2016) toisaalta käy lävitse työn ja geosimulointitekniikoiden taustoja sekä niihin johtaneita tutkimuksia vuosikymmenien ajalta. Työ vaikuttaa sen vuoksi alkuosuudeltaan kirjallisuustutkielmalta.

Rienow & Strenger (2014) taasen käsittelee työssä hyvin kattavasti ja seikkaperäisesti työn eri osuudet sekä koko ilmiön taustat. Koska malli oli suunnattu tiettyä käyttötapauksia varten, on työssä käyty lävitse tarkasti mallinnettavan ilmiön taustatiedot sekä ongelmat.

Kun ongelmat ovat selvillä, pystytään perustelemaan paremmin, miksi työssä tehdyt ratkaisut on tehty. Tämä luo varmuutta työn tuloksia kohtaan. Työ on visuaalisesti myös selkeä useiden kaavio- ja selityskuvien myötä.

#### **6.4.2 Soluautomaattien heikkouksien arviointi**

Kiranjith et al. (2016) käyttää työssään normaalia, soluautomaateille tyypillistä neliöruudukkoa ympäristön jakamiseen soluihin. Kyseinen jako on tarkoitettu tuottamaan mallille spatiaalinen ulottuvuus solujen jaon ja solujen tilojen myötä. Solun naapuruus koostuu Mooren naapurustosta eli kahdeksasta solua ympäröivästä muusta solusta. Naapurisuhteet perustuvat tällöin vierekkäisyyteen eikä laajempaa spatiaalista skaalaa käsitellä naapuruudessa. Ilmiön leviäminen tapahtuu tällöin naapurisoluihin tapahtuvan kehittymisen perusteella. Mallissa on silti käytössä NetLogosta tuotu malli kiinnostavuudelle ja vetovoimatekijöille, jolloin malli kykenee vaikuttamaan soluihin epäsuorasti naapuruuden ulkopuolellakin.

Rienow & Strenger (2014) käyttivät työssään mukautettua UGM-mallia nimeltä UGMr. UGMr on soluautomaatti, mutta kasvu tapahtuu automaattissa vaiheittain neljän kasvusäännön mukaisesti. Tällöin kasvua ei oltu sidottu soluissa naapureihin, vaan mallissa otetaan huomioon useat eri kaupungeissa kasvuun liittyvät tilanteet. UGMr tueksi malliin sisällytettiin SVM, joka ohjaa soluautomaattia todennäköisyyskartan avulla. SVM ohjaa tällöin soluautomaatin kehittymistä alueille, jotka eivät välttämättä ole lokaaleita. Tällöin myös soluautomaatin spatiaalinen skaala on laaja.

Rienow & Strenger (2014) sitoivat alueiden kehittymisen SVM:n avulla teoriapohjaan. SVM ottaa huomioon vetovoimatekijät, jotka vaikuttavat alueiden kehittymiseen kaupungeissa. Tekijöitä voidaan mallissa kalibroida ja tuottaa täten tarkempia tuloksia. Kokonaisuutena UGMr-SVM sitoo teorian hyvin soluautomaatin solujen kehittymisen sääntöihin.

Kiranjith et al. (2016) ei käsittele työssään miten teoria on sidottuna soluautomaatteihin. Soluautomaatteja käytetään pikemminkin keinona mallintaa ympäristöä yksinkertaisilla säännöillä ja toimia alustana, joka tuottaa spatiaalisen ympäristön agenteille. Toisaalta Kiranjith et al. (2016) malli ei perusta muutenkaan mallin simuloimaa ilmiötä olemassa olevaan teoriaan, vaan simuloi mahdollisia skenaarioita.

Kiranjith et al. (2016) ei käsittele työssään mallin validointia. NetLogo sisältää mahdollisuuden mallin validointiin, mutta työssä validointiin ei keskitytä. Validointia ja verifiointia käsiteltiin kyseisen työn osalta myös agenttien kohdalla. Rienow & Strenger (2014) taasen ottavat validoinnin huomioon mallissaan. Mallille on asetettu tarkat tavoitteet ja ennustuksen tuloksia arvioidaan kerättyihin tilastoihin ja virallisiin ennustuksiin väestön kehitymisestä alueilla.

### **6.5 Loppupäätelmät vertailusta**

Vertailussa huomattiin merkittäviä eroja mallien välillä useissa osa-alueissa. Ensinnäkin, töiden tarkoitus ja perimmäinen tavoite tutkimuksessa erosivat toisistaan. Kiranjith et al. (2016) työ oli pohjimmiltaan kirjallisuustutkimus geosimuloinnista sekä geosimulointimallin esittely. Tavoite työssä keskittyi enemmän tietämyksen tavoitteluun ja kokeilemalla löytää uudenlaisia keinoja simulointiin. Rienow & Strenger (2014) taasen keskittyivät työssä ratkaisemaan selkeästi rajattua ongelmaa. Eroavaisuudet töiden lähtökohdissa osoittautuivat hyväksi vertailulle, koska voitiin huomata, mitkä tekijät ovat eri tasoissa töissä merkittävimpiä.

Vertailua varten muodostettiin kriteereitä, joiden tarkoituksena oli kiinnittää töissä huomiota erityisesti siihen, miten soluautomaatit ja agentit oli implementoitu töihin. Tämän työn yhtenä tutkimusongelmana tahdottiin selvittää, miten nykyaikaiset mallit käyttävät agentteja ja soluautomaatteja, joten kriteerit muodostuivat sen pohjalta, mitkä soluautomaattien ja agenttien ongelmat on tunnistettu. Ongelmien ratkaisun pohjalta voidaan siten arvioida miten hyvin malli soveltaa kyseisiä tekniikoita. Kriteerit eivät suoraan vastaa siihen, miten soluautomaatit ja agentit on yhdistettynä töissä, mutta niiden pääasiallinen tehtävä oli sen sijaan tuottaa kriittistä analysointia töiden rakenteesta.

Kiranjith et al. (2016) työ pyrki sisältämään kaupunkimallin ilmiön tutkimisen, jota ei ole tutkittu aikaisemmissa töissä merkittävästi. Eriarvoistumisen ja segregaatoin tutkiminen kaupunkiympäristössä ja asumisalueiden muodostuksessa on tärkeä aihe tutkittavaksi kaupunkien kasvun myötä. Tyypillinen ongelma aikaisemmin kehitetyissä malleissa on ollut, ettei kaupunkien kokonaiskehitystä ole huomioitu, vaan ainoastaan tietyt aspektit liittyen ongelmaan. Toinen ongelma on liittynyt ilmiön sosiaalisiin taustoihin. Tyypillisesti kaupunkimalleissa on lähtökohtaisesti tutkittu asuinalueiden jakautumista taloudellisista ja infrastruktuurisista syistä, kun taasen segregaatio pohjautuu syvempiin sosiaalisiin struktuureihin. Tämän johdosta mallia voidaan pitää uudenlaisena lähestymistapana kaupunkiympäristössä tapahtuvan ilmiön mallintamisessa.

Rienow & Strenger (2014) mallin tavoitteet olivat selkeät. Ilmiö, siihen vaikuttavat tekijät ja ongelmat oli tunnistettu työn alussa, ja mallin tuottama data sekä sen laatu tarkastettiin. Vaikka työssä käytettiin jo olemassa olevia malleja, kyettiin osoittamaan, että mallien implementointi ja muutokset tuottivat parempilaatuista dataa. Tällä tavoin kyettiin osoittamaan, miten geosimuloinnissa voidaan hyödyntää useita eri malleja keskenään, mutta tuottamaan silti hyviä tai parempia tuloksia. Yksi ongelma, joka on tunnistettu laajalti alan kirjallisuudessa, on että agenttipohjaisten mallien ja soluautomaattien yhdistetyistä malleista puuttuu standardisointi ja yhtenäinen peruspohja, jolta kaikki mallien kehittäjät voisivat edistää mallien tarkkuutta ja toimivuutta. Rienow & Strenger (2014) tunnistivat työssään, ettei käytetty kahden mallin yhdistelmä ollut täysin integroitunut ratkaisu. Geosimuloinnissa haasteena on ollut yhdistää soluautomaatit ja agentit yhtenäiseksi systeemiksi, jossa ympäristö, agentit ja agenttien interaktio toimivat yhdessä systeemissä. Rienow & Strenger (2014) esittelivät työssään ratkaisua yhtenäisen systeemin rakenteesta, jolloin voidaan olettaa, että työn tekijät ovat huomioineet agenttipohjaisten systeemien ja soluautomaattien yhdistelmän haasteet.

Suurin puute Kiranjith et al. (2016) mallissa on, ettei työ käytä mallissaan oikean maailman dataa, vaan pohjautuu puhtaasti oletetuille arvoille. Kyseinen rajoite rajaa mallin teoreettiseksi, jolloin validoinnin suorittaminen mallille ei onnistu yhtä hyvin kuin Rienow & Strenger (2014) mallille. Teoreettisuudesta huolimatta mallista voidaan huomioida, millä tavoin soluautomaatit ja agentit on implementoitu. Kiranjith et al. (2016) käyttivät parviällyn ja keskuspaikkateorian piirteitä agenttien käyttäytymisen. Vaikka piirteitä otettiin vain vähänlaisesti malliin mukaan, on huomioitavaa, miten pienillä muutoksilla kyetään tuottamaan tuloksia kompleksisen systeemin mallinnuksessa. Tällöin voidaan huomata myös potentiaali tekniikoissa, joita ei ole aiemmin hyödynnetty vastaavissa tilanteissa. Toisaalta mallissa voidaan myös huomata, ettei käyttäytymismalli pohjautu suoranaisesti olemassa olevaan teoriaan.

Oikean maailman datan puute Kiranjith et al. (2016) mallissa tuottaa myös kysymyksiä siitä, miten hyvin malliin tuodut ulkopuoliset komponentit oikeasti suoriutuvat. Kiranjith

et al. hyödynsivät NetLogon komponentteja agenttien mallinnuksessa sekä soluautomaatteja agenttien ympäristön määrittelyssä. Yksi mallintamisen ongelma, jota arviointi ja kriteerit eivät koskettaneet, on väärin olettamien tekeminen. Jo aiemmin tässä työssä on todettu, että kaupunki on kokonaisuutena suurempi kuin osiensa summa. Tällöin yksi mahdollinen ongelmatilanne mallintajalle on olettaa liaksi erillisten komponenttien ratkaisevan yhdessä kompleksisen systeemin tuottamat ongelmat. Vaikka toisessa mallissa toimiva komponentti olisi hyväksi todettu ratkaisu kyseisessä mallissa ilmentyneeseen ongelmaan, ei voida olla varmoja, miten komponentti sopii muihin malleihin. Tämä johtuu laajemmasta ongelmasta kaupunkien mallinnuksessa varsinkin agenttipohjaisten mallien kohdalla, koska vakiintunutta mallinnustapaa ei ole syntynyt alalle. Kiranjith et al. esittävät mallissaan teoreettisen idean ilmiön mallintamisesta, mutta tekevät tällöin samalla olettamia siitä, mitkä tekijät vaikuttavat ilmiön leviämiseen kaupunkissa.

Rienow & Strenger (2014) perustelivat työssään hyvin miksi valitsivat käyttöönsä komponentit mallintamaan ilmiötä ja miten hyvin komponentit suoriutuivat. Perustellut valinnat ja tuloksien esittely liittyy viimeiseen agenttien kriteeriin, joka tässä työssä otettiin osaksi arviointia. Mallin esittely ja kattava selvitys eri komponenttien välisestä toiminnasta on erittäin tärkeä tekijä, kun pyritään edistämään geosimulointia tieteenalana. Luottamus mallin toimintaan ja sen heikkouksien tunnistaminen ovat osa kokonaisprosessia. Jos verrataan perinteisempiin tieteenaloihin, on mallien päätavoitteena ollut riittävä yksinkertaistus ja abstraktiointi. Tästä kyseisestä ajattelutavasta on geosimuloinnissa siirrytty kohti kattavamman kuvan tuottamista monimutkaisesta, kompleksisesta kaupunkiympäristöstä.

Kummastakin tutkitusta mallista huomataan, että haasteita esiintyy yhä siinä, miten agentit ja soluautomaatit saadaan toimimaan samassa ympäristössä. Agenttipohjaisissa malleissa on haasteena ollut agentin määrittely mallin sisällä. Pelkästään eri agenttien tyyppien määrän selvittäminen on mallinnuksessa merkittävä haaste, kun mallinnettavan kompleksisen systeemin koko kasvaa. Toisekseen soluautomaatit ovat kokeneet muutoksia siinä, millä tavoin ne sidotaan ympäristöön. Yksinkertaisesta perusmallista on siirrytty joustavampiin ja monimutkaisempiin malleihin, joissa varsinkin rajoittavat naapurustot on koettu liian yksinkertaistetuiksi kuvaamaan yksityiskohtaisempia kaupunkimalleja. Tällöin on syntynyt kysymys siitä, voiko osaa soluautomaateista enää kutsua soluautomaateiksi niiden kokeman muutoksen myötä. Kompromisseja malleissa on jouduttu tekemään kompleksisuuden vuoksi ja varsinkin Rienow & Strenger (2014) mallissa joudutaan jättämään huomiotta lukuisia asuntomarkkinoiden mekaniikkoja.

## 7 Päätäntö

Tässä työssä tutustuttiin geosimulointiin ja siihen liittyviin kahteen simulointitekniikkaan, soluautomaatteihin ja agenttipohjaisiin malleihin. Työn ensimmäinen osuus oli kirjallisuustutkimus, jossa selvitettiin ensin paikkatiedon ominaisuuksia yleisellä tasolla sekä perusasioita mallintamisesta. Seuraavissa luvuissa käytiin lävitse soluautomaattien ja agenttipohjaisten mallien perusominaisuuksia sekä miten kyseisiä tekniikoita voidaan käyttää geosimuloinnissa. Tätä seuraavassa luvussa keskityttiin geosimulointiin. Perusteet ja taustat käsiteltiin, sekä myös geosimuloinnin erityispiirteistä ja haasteista otettiin selvää. Työn toisessa osuudessa vertailtiin kahta nykyaikaista tutkimusta liittyen geosimulointiin. Tutkimuksissa muodostettujen mallien vertailua varten tuotettiin kriteerit seitsemästä merkittävästä haasteesta, jotka kohdataan agenttipohjaisten systeemien kohdalla, sekä kolmesta haasteesta, joita kohdataan soluautomaattien yhteydessä. Arvioinnin tarkoitus oli tutustua nykyaikaisiin geosimuloinnin mallinnusmetodeihin.

Työn yhtenä tutkimusongelmana oli selvittää, miten soluautomaatteja ja agenttipohjaisia systeemeitä käytetään geosimuloinnissa. Tutustumalla kummankin tekniikan alkuperään sekä lähtökohtaiseen perusmalliin kyettiin muodostamaan ymmärrys tekniikoihin. Perusmallin lisäksi tutustumalla erityyppisiin sovelluksiin sekä tyypillisimpiin haasteisiin, oli tekniikoiden ongelma-alueiden tunnistaminen helpompaa. Kummastakin tekniikasta löytyi runsaasti materiaalia, mutta kaupunkimallinnukseen erikoistunutta kirjallisuutta löytyi pääosin vain agenttipohjaisista systeemeistä. Huomioitavaa kummassakin tekniikassa oli sovellettujen mallien suurilukuinen määrä. Tämän vuoksi kirjallisuustutkimuksessa pyrittiin esittelemään perusmuotoinen malli kummastakin tekniikasta, sekä muutamia eri erikoistuneempia malleja. Tämän avulla kyettiin esittelemään kuinka tekniikoita nykyisin on sovellettu.

Työn toisena tutkimusongelmana oli tunnistaa vaatimukset hyvälle geosimulointimallille. Hyvää mallia varten on tunnettava siinä käytetyt yleisimmät tekniikat eli soluautomaatit ja agenttipohjaiset mallit. Alan kirjallisuutta geosimuloinnista ei ole runsaasti, joten agenttipohjaisten mallien ja soluautomaattien osalta jouduttiin turvautumaan muidenkin tieteenalojen julkaisuihin. Kaupunkimallinnuksen ulkopuolelta oli mahdollista löytää kirjallisuutta kummankin tekniikan heikkouksista ja haasteista, joten pystyttiin muodostamaan kuva siitä, mitä ongelmia on ratkaistava, jotta voidaan muodostaa hyvä geosimulointimalli.

Kolmantena tutkimusongelmana oli selvittää miten sekä soluautomaatteja että agenttipohjaisia malleja käytetään nykyaikaisessa geosimulointimallissa ja miten kyseiset mallit ottavat huomioon hyvän mallin vaatimukset. Tutkimusta varten valittiin kaksi nykyaikaista geosimulointimallia sekä muodostettiin vertailua varten joukko haasteita, joita kohdataan geosimuloinnissa käytetyissä tekniikoissa. Geosimuloinnille ei ole vakiintunutta tieteellistä tapaa tai standardimallia, joten kumpikin valittu tutkimus oli muodoltaan hyvin erilaisia. Vertailussa silti huomattiin, että kummassakin työssä kohdattiin samoja haasteita ja töiden käyttötarkoitus ratkaisi pitkälti sen, miten haasteet oli töissä huomioitu.

Tämän työn lähtökohta geosimulointiin oli rajata aihe koskemaan kaupunkimallinnusta. Kaupunkimallinnus on yksi geosimuloinnin suurimmista sovellusaloista. Kaupunkeja on tutkittu laajasti useilla eri tieteenaloilla niiden kompleksisuuden ja monimuotoisuuden

vuoksi. Kaupungit ovat myös kasvamassa lähivuosisikymmeninä nopeasti, jolloin joudutaan ottamaan huomioon niiden vaikutus ympäristöön, ihmisiin ja sosiaalisiin rakenteisiin. Tällöin myös kaupunkisuunnittelun rooli muuttuu merkittävästi. Geosimulointi tarjoaa kaupunkisuunnittelulle tukea päätöksenteossa. Malleilla kyetään joko simuloimaan oikean elämän dataa käyttämällä ja ennustamaan tulevia muutoksia tai luomaan mahdollisia skenaarioita synteettisellä datalla. Geosimulointi toimii siis tärkeänä työkaluna kaupunkisuunnittelussa, vaikka nykyisin myös muissakin yhteyksissä geosimulointia voidaan hyödyntää.

## Lähdeluettelo

Altartouri, A. 2015. Application of advanced geospatial computation and technology to model and manage coastal environments. Yhdyskunta ja ympäristötekniikan laitos. Aalto-yliopiston Insinööritieteiden korkeakoulu. Otaniemi.

Avolio, M.V., Gregorio, S.D., Lupiano, V., Trunfio, G.A. 2012. Simulation of wildfire spread using cellular automata with randomized local sources. Cellular automata, 10th International Conference on Cellular Automata for Research and Industry. Vol 7495. S. 279-288. ISBN 978-3-642-33349-1.

Batty, M. 2013. The New Science of Cities. Cambridge, USA: The MIT Press. S. 520. ISBN 978-0262019521.

Benenson, I., Torrens, P. M. 2004. Geosimulation: Automata-based modeling of urban phenomena. 1. painos. London, UK: John Wiley and Sons. 312 s. ISBN 978-0470843499.

Bonebeau, E. 2002. Agent-based modeling: Methods and techniques for simulating human systems. Proceeding of the National Academy of Sciences of the United States of America. Vol. 99:3. S. 7280-7287. ISSN 1091-6490.

Bonhanm-Carter, G. 1994. Geographic Information Systems for Geoscientists: Modelling with GIS. Oxford, UK: Pergamon Press. S. 398. ISBN 9780080424200.

Brown, D.G., Pijanowski, B.C., Duh, J.D. 2000. Modeling the relationships between land use and land cover on private lands in the Upper Midwest, USA. Journal of Environmental Management. Vol. 59:4. S. 247-263. ISSN 0301-4197.

Chopard, B. & Droz, M. 1998. Cellular automata modeling of physical systems. Cambridge, UK: Cambridge University Press. 337 s.

Clarke, K.C., Hoppen, S., Gaydos, L. 1997. A Self-modifying cellular automata model of historical urbanization in the San Francisco Bay area. Environment and Planning B: Urban Analytics and City Science. Vol. 24:2. S. 247-261. ISSN 2399-8091.

Clarke, K.C., Brass, J.A., Riggan, P.J. 1994. A Cellular Automaton Model of Wildfire Propagation and Extinction. Photogrammetric Engineering & Remote Sensing. Vol 60:11. S. 1355-1367. ISSN 0099-1112.

Couclelis, H. 1997. From cellular automata to urban models: New principles for model development and implementation. Environment and Planning B: Planning and Design. Vol 24. S. 165-174. ISSN 2399-8083.

Crooks, A., Castle, C., Batty, M. 2007. Key challenges in agent-based modelling for geospatial simulation. Computers, Environment and Urban Systems. Vol. 32:6. S. 417-430. ISSN 0198-9715.

Crooks, A.T., Heppenstall, A.J. See, L.M., Batty, M. 2012. Agent-based Models of Geographical Systems. Berliini, Saksa: Springer-Verlag. S. 759. ISBN 978-90-481-8926.

- Crooks, A.T. 2017. Modeling the emergence of riots: A geosimulation approach. *Computers Environment and Urban Systems*. Vol. 61. S. 66-80. ISSN 0198-9715.
- Dahal, K.R., Chow, E. 2014. An agent-intergrated irregular automata model of urban land-use dynamics. *International Journal of Geographical Information Science*. Vol 28:11. S. 2281-2303. ISSN 1362-3087.
- de Jong, S. 2008. Fairness in multi-agent systems. *Proceedings of the 7<sup>th</sup> international joint conference on Autonomous agents and multiagent systems*. Vol. 2. S. 863-870. ISBN 978-0-9817381-1-6.
- Erickson, B., Lloyd-Jones, T. 1997. Experiments with Settlement Aggregation Models. *Environment and Planning B: Urban Analytics and City Science*. Vol 24:6. S. 903-928. ISSN 2399-8083.
- Etzion, O., Gal, A. 1998. A multiagent update process in a database with temporal data dependencies and schema versioning. *IEEE transactions of knowledge and data engineering*. Vol. 10:1. S. 21-37. ISSN 1041-4347.
- Felson, M., Wilensky, D.U. 2007. NetLogo urban suite – economic disparity model. Center for Connected Learning and Computer-based Modeling, Northwestern University, Evanston, IL, USA. Saatavissa: <http://ccl.northwestern.edu/netlogo/models/UrbanSuite-EconomicDisparity> (viitattu 14.3.2017)
- Ferber, J. 1999. *Multi-agent systems: an introduction to distributed artificial intelligence*. 1. painos. Reading, US: Addison-Wesley. 509 s. ISBN 0201360489.
- Francez, N. 1986. *Fairness*. Texts and monographs in Computer Science. Berliini, Saksa: Springer-Verlag. S. 297. ISBN 978-1-4612-9347-7.
- Franklin, S., & Graesser, A. 1996. Is it an agent, or just a program?: A taxonomy for autonomous agents. *Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages. Lecture Notes in Artificial Intelligence*. Vol: 1193. S. 21-35. ISBN 3-540-62507-0.
- Geosimulation. Katsottu 22.5.2017. <http://geosimulation.org/geosim/>
- Gorodetskaya, A. 2013. *Spatial agent-based models in urban applications*. Diplomityö. Maanmittaustieteiden laitos. Aalto-yliopiston insinöörityöiden korkeakoulu. Otaniemi. S. 54.
- Grimm, V., Rialsback, S.F. 2005. *Individual-based Modelling and Ecology*. Princeton, US: Princeton University Press. S. 448. ISBN 9780691096667.
- Hall, A. 2016. *Reasoning in Spatio-Temporal Analysis*. Väitöskirja. Rakennetun ympäristön laitos. Aalto-yliopiston insinöörityöiden korkeakoulu. Otaniemi. S. 191.
- Hartikainen, L. 2010. *Kolmiulotteiset kaupunkimallit ja niiden käyttömahdollisuudet kuntien toiminnossa*. Diplomityö. Maanmittaustieteiden laitos. Aalto-yliopiston insinöörityöiden korkeakoulu. Otaniemi. S. 63.

Heppestall, A.J., Crooks, A.T., See, L.M., Batty, M. 2012. Agent-based Models of Geographical Systems. Berliini, Saksa: Springer. S. 760. ISBN 978-90-481-8927-4.

Holling, C.S. 1978. Adaptive Environmental Assessment and Management. New Jersey, USA: Wiley & Sons. S. 377. ISBN 9781932846072.

JHS 162. Julkishallinnon suositus. 2012. Paikkatietojen mallintaminen tiedonsiirtoa varten. Julkisen hallinnon tietohallinnon neuvottelukunta.

Johnson, S. 2002. Emergence: The connected lives of ants, brains, cities, and software. New York, USA: Simon & Schuster. S. 288. ISBN 978-0684868769.

Karafyllidis, A., Thanailakis, A. 1996. A model for predicting forest fire spreading using cellular automata. Ecological modelling. Vol 99:1. S. 87-97. ISSN 0304-3800.

Kiranjith, J., Giridhar, K., 2016, Agent Based Simulation for Understanding Urban Dynamics. Journal of Interior Designing and Regional Planning. Vol. 1:3. S. 1-21. ISSN 1939-1668.

Koistinen, K. 2011. Paikkatiedon tietotuoteskeemojen ontologisointi tiedonhaun tueksi. Diplomityö. Maanmittatustieteiden laitos. Aalto-yliopiston insinööritieteiden korkeakoulu. Otaniemi. S. 53.

Langton, C. 1992. Life at the edge of chaos. Artificial Life II. Redwood, Kanada: Addison-Wesley. S. 41-93. ISBN 0226476553.

Li, X., Gar-On Yeh, A. 2010. Modelling sustainable urban development by the intergration of constrained cellular automata and GIS. International Journal of Geographical Information Science. Vol. 14:2. S. 131-152. ISSN 1365-8816.

Longley, P.A., Goodchild, M.F., Maguire, D.J., Rhind, D.W. 2011. Geographic Information Systems & Science. 3. painos. New Jersey, USA: John Wiley & Sons, Inc. S. 539. ISBN 978-0-470-72144-5.

Louie, M.A., Carley, K.M. 2008. Balancing the criticisms: validating multi-agent models of social systems. Simulation Modelling Practice and Theory. Vol. 16:2. S. 242-256. ISSN 1569-190X

Marceau, D.J., Benenson, I. 2011. Advanced Geosimulation Models. Hilversum, Alankomaat: Bentham Science Publishers. S. 142. ISBN 9781608052226.

Maria, A. 1997. Introduction to Modeling and Simulation. Proceedings of the 29<sup>th</sup> conference on Winter simulation. S. 7-13. ISBN 0-7803-4278-X.

Mondal, P., Southworth, J. 2010. Evaluation of conservation interventions using cellular automata-Markov model. Forest Ecology and Management. Vol 260:10. S. 1716-1725. ISSN 0378-1127.

Rienow, A., Strenger, D. 2014. Geosimulation of urban growth and demographic decline in the Ruhr: a case study for 2025 using the artificial intelligence of cells and agents. Journal of Geographical Systems. Vol. 16:3. S. 311-342. ISSN 1435-5949.

Russel, S., Norvig, P. 1995. Artificial intelligence: a modern approach. 1. painos. New Jersey, US: A Simon & Schuster Company. 1152 s. ISBN 978-0136042594.

Sanastokeskus TSK ry. 2014. Geoinformatiikan sanasto.

Shen, Z., Kawakami, M., Kawamura, I. 2009. Geosimulation model using geographic automata for simulating land-use patterns in urban partitions. *Environment and Planning B: Planning and Design*. Vol 36:5. S. 802-823. ISSN 2399-8083.

Torrens, P.M. 2000. How cellular models of urban systems work. (1. Theory). Centre for Advances Spatial Analysis Working Paper 28. University College London, UK. S. 68. ISSN 1467-1298.

Torrens, P.M. 2006. Simulating sprawl. *Annals of the Association of American Geographers*. Vol. 96:2. S. 248-275.

Turner, M.G. 1988. A Spatial Simulation Model of Land Use Changes in a Piedmont County in Georgia. *Applied Mathematics and Computation*. Vol. 27:1. S. 39-51. ISSN 0096-3003.

Vahidnia, M.H., Alesheikh, A.A., Alavipanah, S.K. 2015. A multi-agent architecture for geosimulation of moving agents. *Journal of Geographical Systems*. Vol 17:4. S. 353-390. ISSN 1435-5949.

Vliet, J. V., White, R., Dragicevic, S. 2011. Modelling urban growth using a variable grid cellular automata. *Computers, Environment and Urban Systems*. Vol. 33:1. S. 35-43. ISSN 0198-9715.

Wolfram, S. 1984. Cellular automata as models of complexity. *Nature*. Vol 311:5985. S. 419-424.

Wolfram, S. 1994. Cellular automata and complexity: Collected Papers. 1. painos. Colorado, USA: Westview Press. S. 608. ISBN 978-0201626643.

Wooldridge, M. 2009. An Introduction to Multiagent Systems. 2. painos. New Jersey, USA: Wiley & Sons. S. 484. ISBN 978-0470519462.

Wu, F., Webster, C.J. 1998. Simulation of land development through the integration of cellular automata and multi-criteria evaluation. *Environment and Planning B: Planning and Design*. Vol 25:1. S. 103-126. ISSN 2399-8083.

Yassemi, S., Dragicevic, S. & Schmidt, M. 2007. Design and implementation of an integrated GIS-based cellular automata model to characterize forest fire behaviour. *Ecological Modelling*. Vol 210:1-2. S. 71-84.