

Vurdering av romlige mønstre i Stavanger Kommune ved bruk av skadedata og multivariat regresjon

- En GIS studie

Masteroppgave i Geografi

Øystein Hole

Norges Teknisk-Naturvitenskapelige Universitet

Fakultet for samfunnsvitenskap og teknologiledelse

 NTNU

Geografisk institutt

Veileder: Jan Ketil Rød

Bi-veileder: Eirik Mannsåker

Abstract

This paper investigates spatial patterns regarding the intensity of incidents and insurance payments related to water damage in the municipality of Stavanger. The identified hot spots answered the question of *where*, while the following ordinary least squares regressions (OLS) sought to answer *why* insurance payments varied. The why was explained using different terrain parameters, among them slope, aspect and curvature. Furthermore, flow accumulation and variables related to the characteristics of the sewerage system were included. Two regressions were conducted, one for the municipality as a whole and one for the identified problem area. The first one derived the terrain parameters from a digital elevation model (DEM) of a 10x10 meter resolution. While the second one derived, the same parameters from a DEM converted from a LiDAR point cloud using a 1x1 meter resolution. The results were respectively a 10 percent and a 27 percent explanation rate in the variation of the dependent variable.

Sammendrag

Denne oppgaven utforsker romlige mønstre vedrørende intensiteten av hendelser og erstatningsutbetalinger knyttet til vannskader i Stavanger Kommune. Hot spots som ble identifisert, svarte på spørsmålet om *hvor*, mens de følgende minste kvadraters metode (MKM) svarte på *hvorfor* erstatningsutbetalinger varierte. Dette ble forklart ved bruk av ulike terrengparametere som skråning, aspekt og kurvatur. Videre ble en akkumuleringsvariabel og egenskaper tilknyttet avløpssystemet inkludert. To regresjonsanalyser ble gjennomført, en med utgangspunkt i hele kommunen og en med fokus på det identifiserte problemområdet. Den første ekstraherte terrengparameterne fra en digital elevasjonsmodell med en 10x10 meter oppløsning. Mens den andre ekstraherte de samme parameterne fra en DEM med 1x1meter oppløsning som var konvertert fra en LiDAR punktsky. Resultatene var henholdsvis 10 prosent og 27 prosent forklaringssevne i variasjonen til den avhengige variabelen.

Forord

Ferdigstillelsen av denne masteroppgaven betyr slutten på mine fem år som student ved Geografisk Institutt, NTNU. Oppgaven er skrevet i samarbeid med Stavanger Kommune og har som hensikt å nyttiggjøre seg av tilgang på skadedata fra Finans Norge. Prosessen har vært lærerik og krevende og vil være god ballast å ha med seg videre inn i arbeidslivet.

Først vil jeg rette en stor takk til min veileder Jan Ketil Rød for å ha pekt meg i retning av dette fagfeltet, i tillegg til gode faglige råd og engasjement gjennom hele studieløpet.

Takk også til mine kontaktpersoner i Stavanger Kommune. Først og fremst til min bi-veileder Eirik Mannsåker for tilrettelegging av datamaterialet og god støtte underveis. Takk også til Hugo Kind og Bjørn Zimmer Jacobsen for gode forslag og idémyldring i startfasen.

Takk også til Ivar Svare Holand for kommentarer og tips på prosjektbeskrivelsen.

Trondheim, 10. mai 2015

Øystein Hole

Innholdsfortegnelse

<i>Abstract</i>	<i>iii</i>
<i>Sammendrag</i>	<i>iii</i>
<i>Forord</i>	<i>iv</i>
<i>Innholdsfortegnelse</i>	<i>vi</i>
<i>Figurliste</i>	<i>x</i>
<i>Tabelliste</i>	<i>xii</i>
<i>Forkortelser</i>	<i>xiii</i>
1. INNLEDNING	1
1.1 PROBLEMSTILLINGER	2
1.1.1 Hovedspørsmål	2
1.1.2 Delspørsmål	2
1.2 BAKGRUNN (FNO PROSJEKT)	2
1.3 STUDIEOMRÅDE	4
1.4 TIDLIGERE ARBEID FOR STAVANGER KOMMUNE	5
1.5 DATAGRUNNLAG FOR SKADEDATA	6
1.5.1 Annet datamateriell.....	6
1.6 OPPGAVENS RELEVANS OG AKTUALITET	7
1.6.1 Regress	9
1.6.2 Stavangerdommen	9
1.6.3 Ansvarsfordeling	10
1.7 OPPGAVENS OPPBYGNING	11
2. TEORI OG METODE	13
2.1 GEOGRAFISKE INFORMASJONSSYSTEMER	13
2.1.1 Vektor og raster.....	13
2.2 HØYDEMODELLER	13
2.2.1 LiDAR	14
2.2.2 Interpolering	15
2.2.3 Romlig autokorrelasjon	17
2.3 ROMLIGE MØNSTER	17
2.4 FLOM OG OVERFLATEAVRENNING.....	19
2.4.1 Dreneringsretning.....	20
2.4.2 Forsenkninger	21
2.4.3 Akkumulering.....	22
2.5 TERRENGPARAMETERE.....	22
2.5.1 Skråning	23

2.5.2	Kurvatur.....	23
2.5.3	Aspekt.....	24
2.6	AVLØPSLEDNINGER	25
2.7	REGRESJONSANALYSE	26
2.7.1	Minste kvadraters metode (MKM)	27
2.7.1.1	Begreper og tester.....	28
3.	DATABEARBEIDING	30
3.1	SKADEDATA	30
3.1.1	Identifisering av romlige mønstre i skadedata.....	30
3.2	BEARBEIDING AV PUNKTSKY	33
3.2.1	Konvertering og interpolering/cellestørrelse.....	35
3.3	BYGGING AV HYDROLOGISK MODELL.....	36
3.3.1	Forsknninger	36
3.3.2	Heving av bygninger	40
3.3.3	Akkumulering.....	42
3.4	OLS.....	44
3.4.1	Enhet for kobling av statistiske data	44
3.4.2	Transformering av variabler	47
3.4.3	Oppsummering av variabelbearbeiding	49
4.	ANALYSE OG RESULTATER.....	52
4.1	HOT SPOT ANALYSE FOR ALLE ADRESSEPUNKTENE	52
4.1.1	Antall hendelser / erstatning	52
4.2	FLOMVEIKART	56
4.3	OLS RESULTATER (HELE KOMMUNEN).....	57
4.4	OLS RESULTATER (UTSNITT).....	66
5.	DISKUSJON.....	73
5.1	EVALUERING AV DATAGRUNNLAG OG ANALYSER.....	73
5.1.1	Evaluering av skadedata.....	73
5.1.2	Evaluering av Hot Spot analysene	74
5.1.3	Evaluering av høydemodeller	76
5.1.4	Evaluering av terrengparametere	76
5.1.5	Evaluering av akkumuleringsraster	77
5.1.6	Evaluering av OLS	79
6.	KONKLUSJON	86
6.1.1	VIDERE ARBEID.....	88

7.	REFERANSER	90
8.	VEDLEGG	101

Figurliste

FIGUR 1.1: ARBEIDSFLYT I PILOTPROSJEKTENE (BREVIK M.FL, 2015).....	3
FIGUR 1.2: TIL VENSTRE INTEGRERT BAKGRUNNSKART I ARCGIS. TIL HØYRE OVERSIKTSKART OVER KOMMUNEN (STATENS KARTVERK, 2015).....	4
FIGUR 1.3: UTSNITT AV KOMMUNEN BRUKT I DETALJERT ANALYSENIVÅ. (STATENS KARTVERK 2014; 2015)	4
FIGUR 1.4: UTVIKLING AV ANTALL SKADER OG ERSTATNINGER FOR VANNSKADER I PERIODEN 1992-2013 (FINANS NORGE, 2013)...	7
FIGUR 1.5: FORDELING AV VANNSKADER SORTERT PÅ ÅRSAKER I STAVANGER KOMMUNE.....	8
FIGUR 2.1: DRENERINGSRETNINGER I ALGORITMEN D8 (ESRI, 2012A)	20
FIGUR 2.2: 3X3 VINDU SOM BEVEGER SEG OVER HØYDEMODELLEN (OLAYA, 2009).....	22
FIGUR 2.3: ANTALL LEDNINGER INNENFOR GITTE ÅRSTALLSKATEGORIER I STAVANGER KOMMUNE.....	26
FIGUR 3.1: TOPPUNKTER HVOR ROMLIG AUTOKORRELASJON ER STØRST. DEN FØRSTE PÅ 667M BLE VALGT	31
FIGUR 3.2: TOPPUNKT HVOR ROMLIG AUTOKORRELASJON ER STØRST. DEN ENESTE PÅ 1061M BLE VALGT	32
FIGUR 3.3: SKJERMEDUMP FRA ARCGIS MED RESULTAT AV PUNKTFILINFORMASJON-RESULTATENE	33
FIGUR 3.4: FORSTØRRET OMRÅDE AV PUNKTSKYEN FOR Å VISE BYGNINGER OG VANN SOM ER FILTRERT BORT (HVITE OMRÅDER).....	34
FIGUR 3.5: RAMMEN (A) VISER OMRISSET AV KOMMUNEN, MENS (B) VISER SELEKSJON AV VANNTEMADATA OG (C) HVOR DISSE DATASETTENE ER KOMBINERT I VERKTØYET UNION. SKJERMEDUMP FRA ARCGIS (STATENS KARTVERK, 2015).....	37
FIGUR 3.6: GEOPROSSESSERING I FORM AV ET PYTHON SKRIPT. MODIFISERT ETTER RØD (2015)	38
FIGUR 3.7: SKJERMEDUMP FRA ARCGIS. RESULTAT AV VERKTØYET SINK. GRÅ/SORTE OMRÅDER INDIKERER OMRÅDER UTEN DEFINERT DRENERINGSRETNING. (STATENS KARTVERK, 2014)	39
FIGUR 3.8: SKJERMEDUMP FRA ARCGIS. VISER OMRÅDER SOM DRENERER TIL HVER SIN FORSENKNING. (STATENS KARTVERK, 2014)	39
FIGUR 3.9: SKJERMEDUMP FRA ARCGIS SOM VISER HØYDEMODELLEN TIL VENSTRE, BYGNINGSRASTEREN ØVERST TIL HØYRE OG DE TO KOMBINERT TIL EN HØYDEMODELL MED HEVEDE BYGNINGER NEDE TIL HØYRE. (STATENS KARTVERK 2014;2015)	41
FIGUR 3.10: BEREGNET FLOMVEI I ARCGIS OG MØLLEBEKKEN FRA FKB-DATA. (STATENS KARTVERK, 2015)	42
FIGUR 3.11: KART OVER FLOMVEIER MED EN AKKUMULERINGSVERDI PÅ OVER 5000 (STATENS KARTVERK, 2015).....	43
FIGUR 3.12: ASPEKTRASTER TIL VENSTRE OG SKRÅNINGSRASTER TIL HØYRE (1 GRADER). BEREGNET FRA HØYDEMODELL MED 10X10 METER OPPLØSNING (STATENS KARTVERK, 2014).....	45
FIGUR 3.13: KURVATURRASTER BEREGNET FRA HØYDEMODELL MED 10X10 METER OPPLØSNING (STATENS KARTVERK, 2014)	45
FIGUR 3.14: RUTENETT Plassert over skråningsraster og et utvalg adressepunkter (STATENS KARTVERK 2014)	46
FIGUR 3.15: AVHENGIG VARIABEL FØR TRANSFORMASJON. EN TREND MED POSITIVT SKJEV FORDELING	48
FIGUR 3.16: LOGARITMISK TRANSFORMERT AVHENGIG VARIABEL SOM ER NÆRMERE NORMALFORDELING	48
FIGUR 3.17: ARBEIDSFLYTEN FULGT FOR Å GJØRE DATASETTENE KLARE FOR REGRESJON	50
FIGUR 4.1: HOT SPOT ANALYSE AV ERSTATNINGSUTBETALINGER I STAVANGER KOMMUNE HVOR ALLE ADRESSEPUNKTER ER TATT MED. VISUALISERT TIL VENSTRE SOM PUNKTER OG INTERPOLERT TIL HØYRE. (STATENS KARTVERK, 2015)	52
FIGUR 4.2: HOT SPOT ANALYSE OVER ANTALL HENDELSER I STAVANGER KOMMUNE HVOR ALLE ADRESSEPUNKTER ER TATT MED. VISUALISERT SOM PUNKTER TIL VENSTRE OG INTERPOLERT TIL HØYRE. (STATENS KARTVERK, 2015)	53

FIGUR 4.3: HOT SPOT ANALYSE OVER ERSTATNINGSUTBETALINGER I STAVANGER KOMMUNE HVOR KUN SELEKTERTE ADRESSEPUNKTER ER TATT MED. VISUALISERT SOM PUNKTER TIL VENSTRE OG INTERPOLERT TIL HØYRE. (STATENS KARTVERK, 2015).....	54
FIGUR 4.4: HOT SPOT ANALYSE OVER ANTALL HENDELSER I STAVANGER KOMMUNE HVOR KUN SELEKTERTE ADRESSEPUNKTER ER TATT MED. VISUALISERT SOM PUNKTER TIL VENSTRE OG INTERPOLERT TIL HØYRE. (STATENS KARTVERK, 2015).....	55
FIGUR 4.5: KART OVER BEREGNEDE FLOMVEIER FOR UTSNITTET AV KOMMUNEN SOM TAR UTGANGSPUNKT I HØYDEMODELLEN MED 1X1 METER OPPLØSNING. TIL HØYRE ER DET IDENTIFISERTE PROBLEMMOMRÅDE FRA HOT SPOT ANALYSEN FORSTØRRET.	56
FIGUR 4.6: KART OVER BEREGNEDE FLOMVEIER FOR HELE KOMMUNEN SOM TAR UTGANGSPUNKT I HØYDEMODELLEN MED 10X10 METER OPPLØSNING. TIL HØYRE ER DELER AV DET IDENTIFISERTE PROBLEMMOMRÅDE FRA HOT SPOT ANALYSEN FORSTØRRET (STATENS KARTVERK, 2015)	57
FIGUR 4.7: SPREDNINGSPLOTT-MATRISSE OVER VARIABLENE SOM VISER FORHOLDET MELLOM AKKUMULERINGS- OG ERSTATNINGSVARIABLEN FORSTØRRET TIL HØYRE I FIGUREN.	59
FIGUR 4.8: ERSTATNINGSUTBETALINGER PÅ LEDNINGER AV GITTE ÅRSKATEGORIER.....	60
FIGUR 4.9: RESIDUALKART SOM VISER OVER- OG UNDERPREDIKSJONER, RØDE RUTER ER OVERPREDIKSJONER OG BLÅ RUTER ER UNDERPREDIKSJONER.	63
FIGUR 4.10: RESULTAT ETTER ROMLIG AUTOKORRELASJONSVERKTØY. INDIKERER AT DET IKKE FINNES STATISTISK SIGNIFIKANTE MØNSTER AV OVER- ELLER UNDERPREDIKSJONER.....	64
FIGUR 4.11: HISTOGRAM SOM VISER FORDELINGEN AV RESIDUALER	64
FIGUR 4.12: SPREDNINGSPLOTT AV RESIDUALER. INGEN INDIKASJON PÅ MØNSTER.	65
FIGUR 4.13: SPREDNINGSPLOTT-MATRISSE OVER VARIABLENE SOM VISER FORHOLDET MELLOM AKKUMULERINGS- OG ERSTATNINGSVARIABLEN I FORSTØRRET FORM TIL HØYRE.....	67
FIGUR 4.14: RESIDUALKART OVER OVER- OG UNDERPREDIKSJONENE I DET IDENTIFISERTE PROBLEMMOMRÅDE.	69
FIGUR 4.15: RESULTAT AV ROMLIG AUTOKORRELASJONSVERKTØY. MØNSTERET INDIKERER IKKE STATISTISK SIGNIFIKANTE OVER- ELLER UNDERPREDIKSJONER.	70
FIGUR 4.16: HISTOGRAM SOM VISER FORDELINGEN OVER RESIDUALER.....	70
FIGUR 4.17: SPREDNINGSPLOTT AV RESIDUALER. INGEN INDIKASJON PÅ MØNSTER.	71

Tabelliste

TABELL 1: VISER AGGREGERINGSFUNKSJON OG TRANSFORMERING BRUKT PÅ VARIABLENE.	49
TABELL 2: RESULTAT FOR HVER ENKELT AV VARIABLENE I OLS-REGRESJON FOR HELE KOMMUNEN.....	58
TABELL 3: GJENNOMSnittlig ALDER PÅ LEDNINGENE INNDELt I FEM KATEGORIER. ANTALL RUTER SOM REPRESENTERER DEN RESPEKTIVE KATEGORIEN AV GJENNOMSnittlig ANLEGGsÅR.	60
TABELL 4: OLS RESULTATER FOR HELE KOMMUNEN.	61
TABELL 5: RESULTAT FOR HVER ENKELT AV VARIABLENE I OLS-REGRESJON FOR UTSNITTET AV KOMMUNEN.....	66
TABELL 6: OLS RESULTATER FOR UTSNITTET AV KOMMUNEN	68
TABELL 7: TABELLEN ER HENTET FRA BREVIK M.FL (2015). EGNE MODIFIKASJONER ER GJORT I TABELLEN HVOR STAVANGER KOMMUNE HAR KODET ANNERLEDES.	101

Forkortelser

GIS – Geografiske informasjonssystemer

LiDAR – Light Detection and Ranging

GPS – Global Positioning System

IMU – Inertial Movement Unit

FKB – Felles Kartdatabase

OLS – Ordinary Least Squares (Minste kvadraters metode)

R^2 – Adjusted R-squared

FNO – Finans Norge

VIF – Varians inflasjonsfaktor (Variance Inflation Factor)

ESRI – Environmental Systems Research Institute

NVE – Norges Vassdrags- og Energidirektorat

NGI – Norges Geotekniske Institutt

IDW – Inverse Distance Weighting

1. Innledning

Klimaendringer er et aktuelt tema i verden, så vel som i Norge. Det bedrives mye forskning på fagfeltet og det finnes forskjellige ventede scenarier avhengig av regionens lokalisering. For Norges del er klimaet forventet å endre seg betydelig i dette hundreåret. Klimaframskrivningene er basert på rapportene fra FNs klimapanel (IPCC 2007) men er modifisert for å tilpasses regional skala. Årlig middeltemperatur for Norge er beregnet å stige med mellom 2,3 og 4,6 grader. Temperaturøkningen er forventet å være størst om vinteren og minst om sommeren. Videre er det store regionale forskjeller både innenfor temperaturøkning, men også forventet årsnedbør. Sistnevnte er ventet å øke mellom 5 og 30 prosent frem mot år 2100. Samtidig er det ventet høyere frekvens av intense nedbørshendelser (NOU 2010:10).

Flomtiltaksutvalget vurderte det dithen i NOU (1996:16) at usikkerhetene i retningen på regionale endringer var såpass betydelig at beslutninger knyttet til flomrisiko svært vanskelig kunne ta hensyn til klimaendringer. Her kan man spore både en kunnskaps- og holdningsendring til NOU-rapporten fra (2010:10) hvor selve *tittelen* på rapporten er 'Tilpasning til eit klima i endring' (Grann, 2011). Klimatilpasningen det snakkes om er erkjennelsen av at klimaet er i endring, og at tiden for handling er nå. FNs klimapanel anslo i 2007 at det var 90 prosent sikkert at klimaendringene var menneskeskapte. Den siste rapporten som kom i 2014, har oppjustert dette tallet til 95 prosent (IPCC 2007; IPCC 2013). Forståelsen for hvordan disse endringene påvirker samfunnet og å gjøre proaktive valg som reduserer negative konsekvenser samtidig som de positive blir utnyttet er NOU-rapportens budskap. Utvalget mener at kunnskapsgrunnlaget er tilfredsstillende nok til å konkludere at tilpasningsarbeidet bør settes i gang umiddelbart (NOU 2010:10).

Våren 2012 lyse organisasjonen Finans Norge ut på anbud et pilotprosjekt. Tildelingen ble gitt høsten 2012 til Vestlandsforskning og Geografisk Institutt ved NTNU. Formuleringen fra NOU (2010:10) dannet utgangspunktet for pilotprosjektet:

«For å styrkje arbeidet med tilpassinga til klimaendringane i forsikringsnæringa og forsikringsordningane si rolle i klimatilpassingsarbeidet tilrår utvalget: – Etablere ein database til offentleg bruk og forskning med eit samla, anonymisert datagrunnlag over

klimarelaterte skadar frå forsikringsselskapa og Naturskadepoolen.» (NOU 2010:10, s.22).

Pilotprosjektet har flere deltakende pilotkommuner, deriblant Stavanger Kommune. Denne oppgaven er posisjonert innenfor deres arbeid med prosjektet.

1.1 Problemstillinger

I denne oppgaven vil tilgang på forsikringsdata knyttet til vannskader være den primære kilden til å besvare mine forskningsspørsmål. For å underbygge disse dataene brukes det FKB-data og høyoppløselige høydemodeller. På bakgrunn av dette har jeg et overordnet forskningsspørsmål, med flere underliggende delspørsmål.

1.1.1 Hovedspørsmål

Kan vi, ved hjelp av tilgang på forsikringsselskapenes naturskadestatistikk, generere kunnskap om nye problemområder?

1.1.2 Delspørsmål

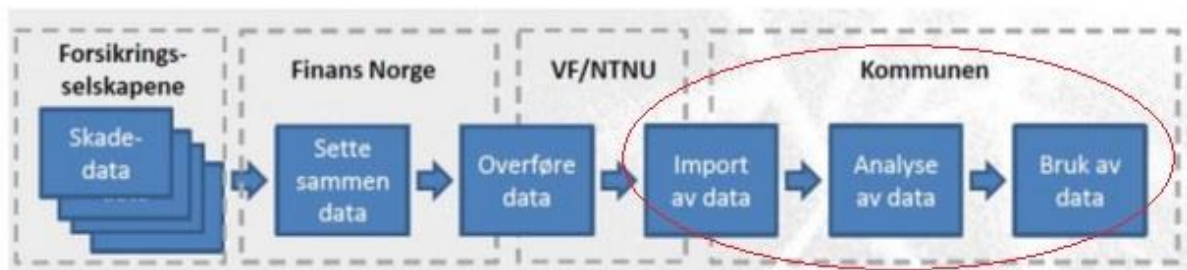
- 1. Vil kartlegging av skadedata fra forsikringsselskapene identifisere problemområder i kommunen og dermed gi kommunen et bedre kunnskapsgrunnlag for å forebygge klimarelaterte skader?*
- 2. Vil modellering av flomveier basert på høyoppløselige terrengmodeller gi et realistisk bilde av hvor flomveiene vil gå, og stemmer dette overens med problemområder identifisert ved kartlegging av skadedataene?*
- 3. Kan en kobling med forsikringsdata opp mot andre temadata gi en økt innsikt i hvorfor identifiserte problemområder er problematiske?*

1.2 Bakgrunn (FNO prosjekt)

Finans Norge (FNO) er hovedorganisasjonen for finansnæringen i Norge. Organisasjonen representerer 50 000 medlemmer og over 200 finansbedrifter. Nevnt blant sine viktigste virksomhetsområder, og som danner basis for datagrunnlaget i masteroppgaven, finnes skadeforsikring. Mer presist for mitt vedkommende, vannskadestatistikken kalt «VASK» (Finans Norge, udatert).

I mars 2013, startet FNO et pilotprosjekt i samarbeid med Klima- og miljøverndepartementet (ved Framtidens byer) for å se på hvordan kommunene kunne nyttiggjøre seg av tilgang til forsikringsselskaperens skadedata. Det overordnede formålet med prosjektet har vært å styrke forebyggingen av klimarelatert naturskade ved å teste ut potensialet og nytteverdien av tilgang til forsikringsselskaperens skadestatistikk. Flere kommuner har deltatt i prosjektet; Stavanger, Fredrikstad, Ringsaker, Trondheim, Tromsø, Løten, Bærum, Kongsvinger. Glåmdalen interkommunale vannverk har i tillegg koordinert aktiviteter for Kongsvinger, Grue og Nord-Odal. I tillegg har flere andre aktører bidratt i prosjektet; Norges vassdrags- og energidirektorat (NVE), Direktoratet for samfunnssikkerhet og beredskap (DSB), Norsk Vann og met.no (Brevik m.fl, 2015).

Nedenfor vises modellen for hvordan pilotforsøkene i prosjektet ble gjennomført.

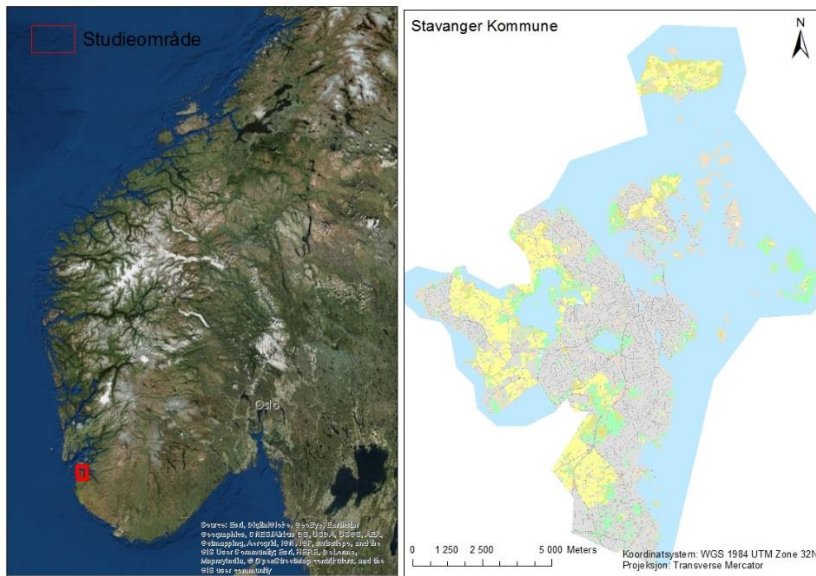


Figur 1.1: Arbeidsflyt i pilotprosjektene (Brevik m.fl, 2015)

Den røde sirkelen rundt de tre siste stegene er hvor denne oppgaven er konsentrert.

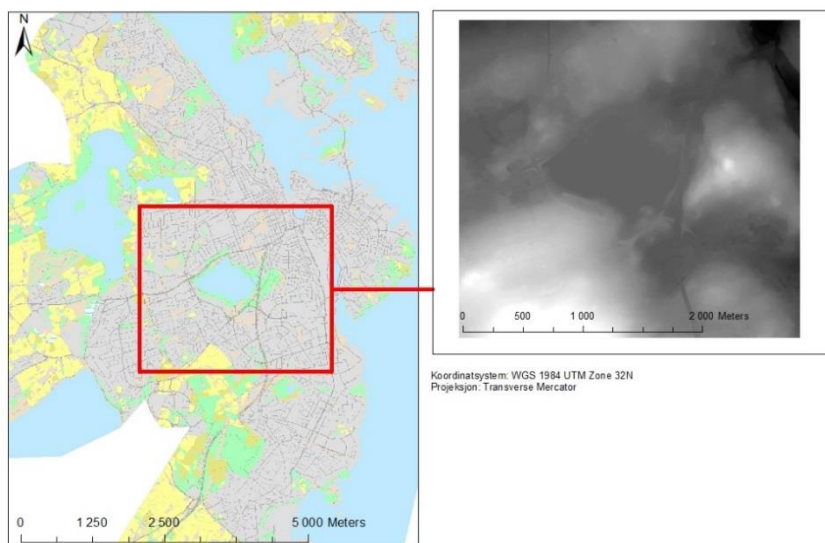
1.3 Studieområde

Stavanger Kommune ble med i pilotprosjektet i 2013. Kommunen har omtrent 130 000 innbyggere og er med det landets tettest befolkede kommune (Brevik m.fl, 2015). I oppgaven vil det være analyser på forskjellig skala. De generelle analysene vil bruke hele kommunen som studieområde.



Figur 1.2: Til venstre integrert bakgrunnskart i ArcGIS. Til høyre oversiktskart over kommunen (Statens Kartverk, 2015)

Mens mer detaljerte analyser vil bli foretatt på et utsnitt av kommunen som har vekket interesse igjennom de generelle analysene.



Figur 1.3: Utsnitt av kommunen brukt i detaljert analysenivå. (Statens Kartverk 2014; 2015)

De detaljerte analysene vil bli gjort innenfor det røde kvadratet i figur 1.3. Stavanger Kommune har siden pilotprosjektets start nyttiggjort seg av skadedataene som ble gjort tilgjengelige for dem via Finans Norge.

1.4 Tidligere arbeid for Stavanger Kommune

Stavanger Kommune fikk i første omgang levert skadedataene i et relativt uoversiktlig format. Den første forsendelsen var såpass rotete, at kommunen ba om å få rådataene isteden. Rådataene ble deretter tatt inn i et program kalt Feature Manipulation Engine (FME) hvor skadedataene ble organisert, systematisert og kodet i henhold til kodeboka (se Vedlegg i kapittel 8) (Brevik m.fl, 2015).

Arbeidsflyten i FME var å hente X og Y koordinatene, hvor adressene ble koblet opp mot matrikkelen. Deretter ble strengformateringen (tegnsetting, store bokstaver, feilstavelser ect.) vasket. Hvor adressenavnene var kjent, men ikke koordinatene, ble programvarer som ArcGIS og GISLine brukt til og geokode disse. En mengde av dataene ble ikke geokodet da vaskeprosessen ville blitt for krevende. Kommunen klarte og kartfeste ca. 12 000 av 17 000 skader. Dette utgjør ca. 70 prosent av hendelsene (Brevik m.fl, 2015).

For denne oppgaven er det klart å hente ut 6744 adressepunkter, altså omkring 39 prosent dekning av skadehendelsene i kommune. Skadedataene som ble tilgjengeliggjort for denne oppgaven ble levert ferdig vasket. Etter vaskingen, ble dataene kvalitetssikret og sammenstilt for klargjøring til analyser. Kommunen har pekt seg ut noen områder hvor skadedataene kan være nyttige i analyser:

- Forebyggende arbeid i planprosesser
 - Separeringsplanlegging for vann og avløp
 - Drift og kontroll av VA-nettet
 - Flomsoneplanlegging
 - **Flomveikartlegging**
 - Kommunale arealplaner
- Beredskapsplanlegging
- **Analyse mot andre temadata**
- **Visualisering av hendelsene**

(Brevik m.fl, 2015)

Punktene som er markert i fet skrift er delvis eller i stor grad berørt i denne oppgaven. Analysene som gjøres tar først og fremst utgangspunkt i skadedataene, men har samtidig annet datamaterialet som støtter, underbygger og brukes i kombinasjon med disse.

1.5 Datagrunnlag for skadedata

Datasettet over forsikringsdata inneholder mange forskjellige egenskaper tilknyttet adressepunktene. Deriblant årsak, installasjon, kilde til skade, forsikrings- og erstatningssum og skadedato. Tabellen som beskriver hvordan dataene er kodet er såpass omfattende at den er plassert i vedleggs-kapittelet på slutten av oppgaven (se kapittel 8).

1.5.1 Annet datamateriell

Skadedataene brukes i kombinasjon med andre data for å sette skadedataene i kontekst, se sammenhenger og visualisere hendelser.

To høydemodeller benyttes, en hentet fra Statens Kartverk som er generert fra høydekurver og høydepunkter fra FKB. Denne har en oppløsning på 10m og er tiltenkt de generelle analysene som dekker hele kommunens areal. Høydemodellen er tilgjengelig fra Statens Kartverk (2014). Den andre høydemodellen er derivert fra en LiDAR-punktsky. Bedriften COWI fikk i oppdrag fra Stavanger Kommune for å samle inn disse dataene. Flygingen ble startet 03.04.2014 og punktettheten er på omtrent 5 punkter per kvadratmeter.

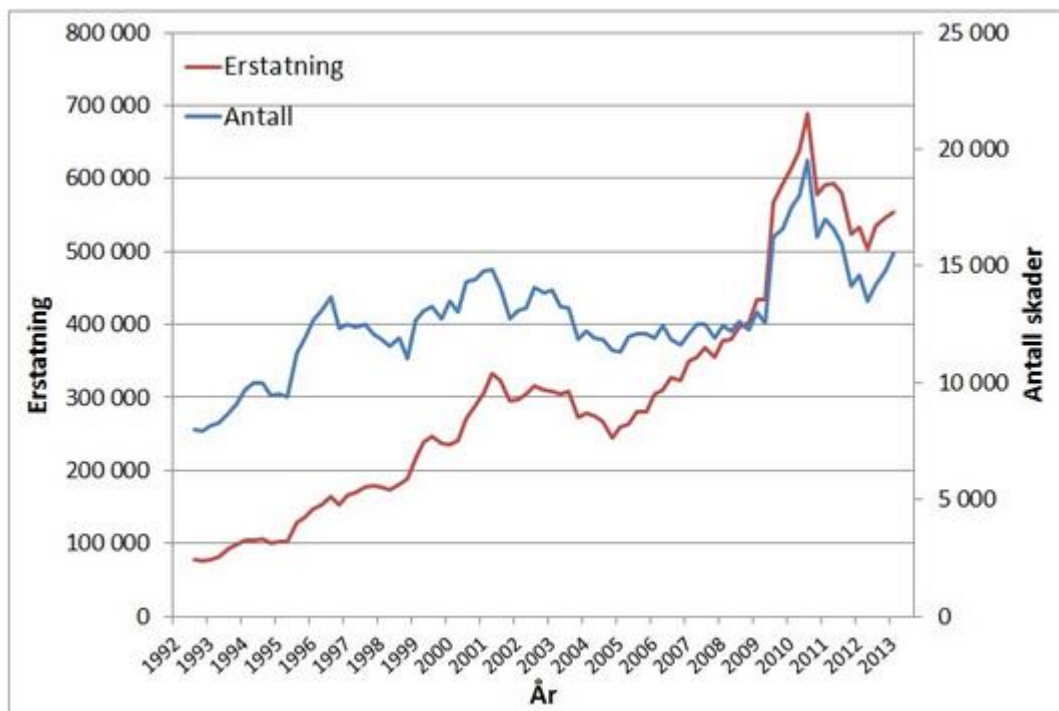
Videre er det benyttet FKB-data som er hentet fra geodatabasen til Stavanger Kommune. FKB-dataene er oppdelt i flere datasett: AR5, arealbruk, bygning, vann og vei. Datafangsten for de geografiske objektene i datasettene daterer seg tilbake til 1968 frem til 2014. Med unntak av AR5 og vei som følger SOSI-format 4.5, følger resterende datasett SOSI-format 4.02. Disse datasettene er også tilgjengelige fra Norge Digitalt (Statens Kartverk, 2015).

I tillegg til nevnte FKB-data, ble også et ledningsdatasett benyttet. Datasettet har oversikt over typen avløpssystem, ledningsmaterialet, alderen på ledningene osv. Ledningsdatasett er tilgjengelig fra Norge Digitalt (Statens Kartverk, 2015), men VA-avdelingen i Stavanger Kommune tilgjengliggjorde et mer detaljert og oppdatert datasett som ble foretrukket med utgangspunkt i SOSI versjon 4.

Alle datasettene brukt og presentert i oppgaven er projisert i WGS 1984 UTM Sone 32N såfremt ikke annet blir nevnt. Videre benytter punktskyen vertikaldatumet NN2000.

1.6 Oppgavens relevans og aktualitet

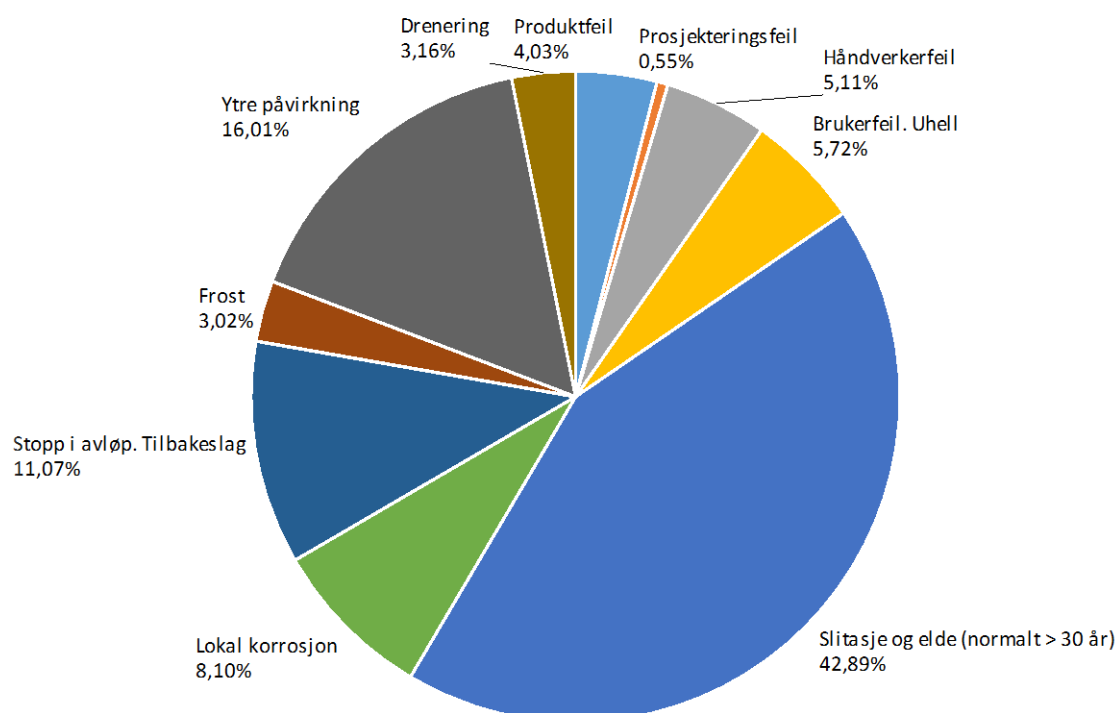
I de senere år har det vært en markant økning i urbane flommer med tilbakeslag fra ledningsnettene og påfølgende kjelleroversvømmelser. Forsikringsselskapenes utbetalinger for skader relatert til vann, har økt radikalt (Norsk Vann 2008:162). Ledningsnettene for overvann og fellesledninger for avløps- og overvann er dimensjonert ut ifra det eksisterende kunnskapsgrunnlaget om nedbørforhold ved tidspunktet for dimensjonering. Samtidig opplever kommuner stadige flom- og vannskader, samt tilbakeslag fra ledningene som følge av intense nedbørshendelser (NOU 2010:10). For de første 9 månedene i 2013, var erstatningsutbetalinger etter vannskader i private boliger og hytter på 1,7 milliarder kroner. Dette er en økning av erstatninger på 14 prosent, mens antall skader har gått opp med 17 prosent (Finans Norge, 2013).



Figur 1.4: Utvikling av antall skader og erstatninger for vannskader i perioden 1992-2013 (Finans Norge, 2013)

Mange av vannskadene skyldtes vanninntrenging utenfra på grunn av mye nedbør, jf. kodeboka (kapittel 8), årsakskode I og J (ytre påvirkning, drenering). Nettopp overvannsproblematikk er et av hovedproblemene Stavanger står ovenfor. Samtidig er dette en av de mest kompliserte utfordringene innenfor klimatilpassningstiltak, da den rådende politikken både statlig og lokalt er å satse på fortetting innenfor allerede bebygde områder (Saglie & Harvold, 2010). Stavanger er som nevnt Norges tettest befolkede kommune. I

praksis fører dette til økte forekomster av impermeable overflater, noe som er en utfordring i håndtering av overvannsproblematikk. De mest interessante skadehendelsene for min del er kodet med I, G og J. Henholdsvis ytre påvirkning, tilbakeslag fra avløpsnettet og drenering. Nedenfor vises et sektordiagram som viser fordelingen av årsaker til vannskader i Stavanger Kommune.



Figur 1.5: Fordeling av vannskader sortert på årsaker i Stavanger Kommune

Av 6744 observasjoner fantes det 3303 observasjoner med ukjent årsakskoding (kode 9). Disse er ikke inkludert i sektordiagrammet. Flere av variablene i sektordiagrammet er heller ikke relevante for mitt formål, da kommunen først og fremst er interessert i saker hvor de er ansvarlige. Saker hvor en vaskemaskin er gått i stykker er for eksempel ikke kommunens ansvarsområde. Det som *er* kommunens ansvarsområder er drift og vedlikehold av avløpsanlegg, jmfør forurensningsloven § 24 (Lovdata, 2015). Når skader oppstår som følge av mangelfullt vedlikehold eller utilstrekkelig kapasitet på avløpsanlegg (gjelder også overvannsanlegg), er kommunen ansvarlige forutsatt at de eier anlegget helt eller delvis, jmfør forurensningsloven § 24a (Lovdata, 2015). Etter en høyesterettsdom datert 24. mai

2012 (Lovdata, 2012b) er også veigrøfter, kummer, rister og stikkrenner i kategorien avløpsanlegg og derfor et av kommunens ansvarsområder (Miljøkommune, 2015). Tilsynelatende er ansvarsområdene klare. Rettspraksis tilsier imidlertid at forurensningsloven § 24a kan fravikes ved avtale, noe som ble klart i en sentral høyesterettsdom kalt ‘Stavangerdommen’ (se kapittel 1.6.2). Dommen har ført til at kommunene sitter på sterkere forhandlingskort i saker med forsikringsselskapene, ettersom de nå har muligheten til å sette inn vilkår i sine avtaler med innbyggerne. Videre er mesteparten av kommunikasjonen mellom forsikringsselskap og kommune gjennom såkalte regresskrav. Kombinert med til tider amper stemning i media har dette skapt grobunn for et konfliktnivå mellom partene (Brevik m.fl 2015, NOU 2010:10).

1.6.1 Regress

Både forsikringstakere og reguleringsmyndigheter har et selvstendig ansvar for å hindre skader, herunder forebygging og begrensnig. Dersom dette ansvaret ikke overholdes, har forsikringsselskapene mulighet til å komme med regresskrav eller nedsatt erstatning (avkorting) (Miljødirektoratet, 2012). Regresskrav blir fremmet ovenfor en tredjepart, når forsikringsselskapet vil ha dekket sine utlegg ovenfor sin forsikringstaker. Dette forutsetter at tredjeparten (som regel kommunen) er ansvarlig skadevolder. Hvis kommunen har unnlatt å følge opp sine ansvarsområder i forurensningsloven og plan- og bygningsloven, kan disse dermed bli gjort ansvarlige og erstatningspliktige (Miljødirektoratet 2012, Lovdata 2013). Dersom partene ikke kommer til enighet om skyldspørsmål, kan saken ende i rettssystemet.

1.6.2 Stavangerdommen

Sakens kjerne var erstatningskrav etter forurensningsloven § 24 a, etter et tilbakeslag fra det kommunale avløpsnett med kjelleroversvømmelse til følge.

Hovedspørsmålet var hvorvidt kommunen var ansvarlig, da vilkårene i tilknytning til vann- og avløpsnett gir kommunen en ansvarsfraskrivelse. § 10 punkt 3 annet ledd i reglementet bestemmer:

«Kommunen er også uten ansvar for skader og ulemper i en eiendom ved oversvømmelse som følge av ekstraordinære regnskylt eller på grunn av at avløpet hemmes av høyvann i fjord, elv eller bekk, uansett om vannet da trenger inn gjennom ledninger, vegger eller åpninger» (Norsk Vann 2013a, s.6)

Stavanger Kommune hadde altså en klausul i vilkårene sine at de var uten ansvar, dersom nedbøren oversteg forutsetningene som var lagt til grunn ved dimensjonering av avløpsledningene i utbyggingsperioden. Eiendommen det gjaldt hadde i lang tid vært tilknyttet det kommunale avløpsnett. Regnværet som kom sommeren 2003 var så kraftig at ledningen ikke hadde kapasitet til å ta unna vannet. Ledningen var dimensjonert for et 10 års gjentakintervall (NOU 2010:10, Norsk Vann 2013a). Forsikringsselskapet IF fremmet regresskrav ovenfor Stavanger Kommune, som deretter avviste kravet. Partene møttes derfor i Stavanger tingrett, hvor kommunen ble dømt som erstatningsansvarlig. Tingretten fant at forurensningsloven § 24a kunne fravikes ved avtale – men at ansvarsfraskrivelsen i dette tilfellet var klart urimelig, slik at den ble satt til side (NOU 2010:10, Norsk Vann 2013a).

Kommunen anket saken videre til Gulating lagmannsrett, hvor kommunen ble frifunnet. Lagmannsretten mente at ansvarsfraskrivelsen kunne gjøres gjeldende, ettersom nedbøren og medfølgende oversvømmelse oversteg forutsetningene som var lagt til grunn i dimensjoneringen av avløpsledningen (NOU 2010:10, Norsk Vann 2013a).

Deretter ble saken anket videre til Høyesterett, av IF skadeforsikring og deres forsikringstaker (Norsk Vann, 2013a). Høyesterett stadfestet lagmannsrettens dom og kom frem til at kommunen kunne fraskrive seg det objektive ansvaret for slike tilfeller, noe som innebærer at abonnenten må kunne påvise uaktsomhet eller skyld hos kommunen (Norsk Vann 2013a, Miljødirektoratet, 2012). I etterkant av dommen har flere kommuner i følge Dagens Næringsliv revidert sine vilkår i vann- og avløpsreglementet for å unngå ansvar ved oversvømmelse (sitert i Aftenposten 2012; Myrstad m.fl 2011).

1.6.3 Ansvarsfordeling

Stavangerdommen vitner om at det er behov for en avklaring av hvilket ansvar som ligger hos den enkelte part når ekstreme værhendelser inntreffer. Som nevnt, har kommunene adgang til å legge inn klausul om ansvarsfraskrivning under spesielle værforhold. Kommunene fritas ved såkalt «force majeure», et begrep som beskriver særlige omstendigheter som er utenfor menneskelig kontroll, og som ikke kan avverges (Lovdata 2012a, Miljødirektoratet 2012, NOU 2010:10). Andre eksempler på force majeure er krigsutbrudd eller større streik ved en vareleveranse (NVE, 1997:10).

Spørsmålet som oppstår i kjølvannet av Stavangerdommen er når force majeure kan brukes. Forsikringsselskapene kan fortsatt fremme regresskrav mot kommunene, men etter dommen i høyesterett er sjansene for å vinne frem redusert. At forurensningsloven ikke alltid gjelder,

gjør at hver enkelt huseier må sette seg inn i avtalen de har med kommunen. Samtidig svekkes forsikringsbransjen forhandlingskort med kommunene når sistnevnte har mindre incentiver enn tidligere til å vedlikeholde ledningene. Det er spekulert i hvor lenge forsikringen vil fortsette å betale, enkelte hevder at det maksimalt vil dekkes tre ganger på samme sted for deretter å kutte forsikringen. Dette er imidlertid ingen langsiktig løsning så lenge ledningssystemet forblir det samme (Miljøverndepartementet, 2010). I en nylig nyhetssak fra NRK (2015) henvises det til at Gjensidige kan komme til å nekte å forsikre eiendommer i utsatte områder. Videre finnes det ingen konkrete krav for hvordan overvanns- og avløpsrør skal dimensjoneres. Det er derfor opp til hver enkelt kommune hvor mye nedbør rørene skal dimensjoneres for, og dermed hvor mye risiko som er akseptabel (NOU, 2010:10). At rørene rustes opp for klimaendringene er nødvendig for å begrense fremtidig skadeomfang. Hvem som skal betale og hvor store skadene vil bli er imidlertid spørsmål som står åpne. Potensielle muligheter er høyere egenandeler og/eller forsikringspremier, det samme er økt vannavgift fra kommunene. Uavhengig av skisserte utfall må kostnadene i stor grad bæres av innbyggerne (Miljøverndepartementet, 2010).

1.7 Oppgavens oppbygning

Oppgaven består av seks kapitler. Det første kapitlet tar for seg det økende fokuset på klimaendringer, forskningsspørsmål, bakgrunnen for pilotprosjektet og oppgavens relevans og aktualitet. Andre kapittel introduserer teorier og begreper som er brukt i oppgaven. Det tredje kapitlet beskriver hvordan dataene jeg hadde til rådighet er blitt bearbeidet og gjort klar til analyse. I kapittel fire blir resultater presentert og analyser redegjort for. I det femte kapitlet blir metodisk tilnærming, analyser og resultater diskutert mer i detalj. Konklusjonen og forslag til videre arbeid finnes i det sjette, og siste kapitlet.

2. Teori og Metode

2.1 Geografiske informasjonssystemer

Mulighetene og bruksområdene innenfor et GIS er så brede og omfattende at en kort og presis definisjon av hva GIS er, vanskelig lar seg gjøre. Generelt sagt er det et system som muliggjør innsamling og lagring av stedfestede data med tilknyttede verdier, med formål om å analysere, manipulere og presentere data for å fremme forståelse og hjelpe beslutningstaking (DeMers 2005; Longley m.fl 2005; Kennedy 2006). Det finnes flere GIS programvarer i stand til å gjøre dette. For denne oppgaven er det brukt Environmental Service Research Institute (ESRI) sin programvarepakke ArcMap med Info lisens, versjon 10.2.2. I tillegg er det benyttet ekstralisenser kalt Spatial Analyst, 3D Analyst og Geostatistical Analyst.

2.1.1 Vektor og raster

I GIS er vektor og raster to forskjellige datastrukturer som representerer verden på hver sin måte. Begge har sine fordeler og ulemper avhengig av bruksområde, men kan i teorien brukes til det samme formålet. I praksis er det imidlertid en sterk tilknytning mellom vektor og *diskrete objekter*, og raster med *kontinuerlige data*. De diskrete objektene har klart gitte grenser, for eksempel veier, bygninger og elver. Disse blir representert ved punkter, linjer og flater (polygoner). For hvert objekt finnes det en korresponderende rad i egenskapstabellen som beskriver objektets attributter. Raster inneholder kolonner og rader av celler, hvor hver eneste celle kun kan holde på én verdi. Disse cellene opptrer i et grid, hvor alle cellene har lik form og størrelse. Raster representerer oftest kontinuerlige data hvor grensene ikke er klart definerte. Dette gjelder blant annet høydedata, temperatur, nedbørsdata og lufttrykk (Longley m.fl 2011; DeMers 2005)

2.2 Høydemodeller

En digital høydemodell (engelsk: digital elevation model) kan etableres både som vektormodell og rastermodell. Felles for dem begge er at de tar utgangspunkt i kjente innsamlede punkter som har X, Y og Z- koordinater. Ved vektormetode kan disse punktene

bli brukt til å triangulere en overflate med trekanter som har basis i de kjente punktene. Dette kalles en TIN (triangulated irregular network) (DeMers, 2005). En annen vektormetode er via høydekurver. Ved rastermetode blir punktene lagret i en datastruktur kalt grid (Gallant og Wilson 2000; Statens Kartverk 2014). Formatet består vanligvis av en rekke like store kvadratiske celler, hvor hver celle kun kan holde på én numerisk verdi (ESRI, udatert (a)). I denne oppgaven blir DEM og høydemodell brukt synonymt og refererer til en raster hvor hver celle har lagret en høydeverdi. Høydemodeller kan variere i kvalitet og nøyaktighet, ofte er dette betinget oppløsningen. I følge Way (1978, sitert i Liu 2008, s.40) er oppløsningen det minste detaljnivået eller det minste objektet som er identifiserbart i et flyfoto. Oppløsningen, eller mer presist den *romlige oppløsningen*, er området én celle dekker av et areal, altså distansen i X og Y retning som representeres av den individuelle cellen (ESRI, 2008). Dette blir ofte oppgitt i meter, for eksempel 10x10 meter oppløsning vil si at cellen strekker seg 10 meter i X og Y retning. Små geografiske objekter vil vanskelig la seg identifisere i en slik oppløsning.

En høyemodell i rasterformat tilbyr flere fordeler sammenlignet med alternativene TIN. Dens simple struktur og bredden i analysemuligheter ble avgjørende for valget. Ifølge Hengl og Evans (2009) er det også langt enklere å ekstrahere terrengparametere fra en raster fordi enklere algoritmer kan bli brukt. I oppgaven er det brukt to høydemodeller. En offentlig tilgjengelig DEM med 10x10 meter oppløsning fra Statens Kartverk, og en LiDAR-derivert DEM med 1x1 meter oppløsning. Førstnevnte dekker hele kommunen, mens sistnevnte dekker et mindre utsnitt. I utgangspunktet hadde det vært mest interessant å bruke en LiDAR-derivert høydemodell for hele kommunen, da denne tilbyr den høyeste oppløsningen. Dette er imidlertid tunge data som krever god maskinvare for å prosessere, noe som ikke var til rådighet.

2.2.1 LiDAR

Et av datasettene som skal brukes for å kjøre analyser i oppgaven er samlet inn ved hjelp av *Light Detection and Ranging* (LiDAR) teknologi. Dette er den mest effektive og pålitelige metoden for innsamling av terrengdata. En kortfattet og generell beskrivelse er at luftbåren LiDAR er en aktiv fjernanalytisk metode som sender laserpulser mot geografiske objekter som deretter reflekterer disse strålene tilbake til mottaker (Liu, 2008). Systemet består normalt sett av tre komponenter: en laserskanner, en GPS (Global Positioning System) mottaker og en IMU (Inertial Measurement Unit). Presisjonen til en LiDAR er prisdrevet nøyaktigheten til hvor godt den kjenner instrumentets egen posisjon (Liu 2008, NGI 2010).

For en luftbåren laser, må posisjonen til sensoren oppdateres for hver puls som sendes ut. Dette blir gjort ved hjelp av henholdsvis GPS og IMU som korrigerer og oppdaterer sensorens posisjon før man måler et punkt i tre dimensjoner (NGI, 2010). Med LiDAR data kan det bli generert svært detaljerte høydemodeller. For generering av disse er det bakkepunktene som er mest interessante for mitt formål. Rådata fra LiDAR har mye «støy», og må igjennom en filtreringsprosess før disse kan hentes ut (Liu, 2008). Støy i dette tilfellet refererer til uønskede returnerte pulser som har truffet objekter som ikke er av interesse. Dette kan være vegetasjon, skyer, bygninger og fugler. Flere filtreringsmetoder eksisterer for å fjerne disse fra dataene, vanligvis en kombinasjon av automatiserte prosesser og manuell editering (Stoker m.fl, 2006). LiDAR-dataene benyttet i oppgaven er både samlet inn og filtrert av COWI. En punktsky ble deretter levert til oppdragsgiveren Stavanger Kommune. Punktene følger et vektorformat, noe som betyr at modellen foreløpig ikke er en kontinuerlig flate. For å oppnå dette må datasettet interpoleres.

2.2.2 Interpolering

Fordi Z-verdiene fra punktskyen er et utvalg av kjente punkter, må interpolering benyttes for å skape en kontinuerlig overflate (DeMers, 2005). Interpolering tar utgangspunkt i algoritmer som gjør kvalifiserte og intelligente gjetninger om verdien til et ukjent punkt, basert på kjente punkter (Longley m.fl, 2005). Det finnes to hovedgrupper av interpoleringsmetoder; deterministiske og geostatistiske. Deterministiske metoder produserer overflater fra målte punkt, basert på graden av likhet mellom punkter (for eksempel Inverse Distance Weighting (IDW)). Geostatistisk interpolering bruker de statistiske egenskapene av de målte punktene ved å kvantifisere den romlige autokorrelasjonen mellom disse. Disse teknikkene kan i tillegg til å produsere interpolerte overflater, også produsere usikkerhetsgrad og feilkilder tilknyttet disse estimerte verdiene. Kriging er et eksempel på dette (Johnston m.fl, 2001). Det finnes imidlertid ingen universelt anerkjent interpoleringsmetode som er best egnet for alle formål, da det blant annet avhenger av fordelingen av punkter og hvilket fenomen som er fokus i studien (Childs, 2004).

Su & Bork (2006) fant i deres studie at IDW skapte en bedre interpolert overflate med mindre feil sammenlignet med kriging, med en gjennomsnittlig punkttetthet på 0.75 per kvadratmeter. Rapporten til Liu (2008), påpeker også at IDW er en god interpoleringsmetode dersom kildedataene har høy punkttetthet. Dette står i en kontrast til en studie av Lloyd og Atkinson

(2002) hvor kriging viste seg som den beste interpoleringsmetoden. Punkttettheten var imidlertid lavere for sistnevnte studie, noe som kan forklares med at IDW fungerer best når punktettheten er høy mens kriging fungerer bedre når den er lav. Lloyd og Atkinson (2002) konkluderer forøvrig at enklere interpoleringsalgoritmer er å foretrekke når punktavstanden er veldig lav. Punktskyen som er brukt i oppgaven har en høy punkttetthet, og IDW ble derfor benyttet som interpoleringsmetode med bakgrunn i studiene som nylig ble omtalt.

IDW bruker ulik vektning av verdier for å estimere verdien til ukjente punkt. De nærmeste kjente punktene har størst betydning i estimeringen, mens punkt lengre unna har mindre betydning (Longley m.fl, 2005). Denne interpoleringsmetoden bygger på Tobler's første lov om geografi, videre omtalt i kapittel 2.2.3. For å beregne det ukjente punktet x , brukes de kjente punktene x_i , hvor i er mellom 1 til n , der n kan være brukerbestemt. Ved å beregne den ukjente høydeverdien til punktet $z(x)$ og de kjente $z(i)$, vil punktene få en vekt d_i som vil bli evaluert basert på avstanden fra x til x_i (Longley m.fl, 2005).

Dette kan uttrykkes ved formelen (Longley m.fl, 2005):

$$z(x) = \frac{\sum_i w_i z_i}{\sum_i w_i}$$

Det vil si at den interpolerte verdien er et gjennomsnitt av de kjente verdiene, vektet ved w . Beregningen av vektningen kan gjøres på forskjellige måter, men den mest brukte er denne (Longley m.fl, 2005):

$$w_i = 1/d_i^2$$

Dette betyr at vektningen til et gitt punkt reduseres med en faktor på fire når avstanden til punktet doubles (eller med en faktor på ni når avstanden tredobles). Dersom det er brukt en kjent z_i hvor z allerede er kjent, vil den kjente verdien bli returnert, da vekten på et punkt med null i avstand er uendelig, dette er grunnen til at IDW faller under den deterministiske kategorien av interpoleringsmetoder. Det vektete gjennomsnittet returnerer alltid en verdi som er innenfor verdiintervallet av de målte verdiene – noe som er en av IDW metodens svakheter. Det vil si at ingen punkter kan ha en interpolert z -verdi som er høyere enn den høyeste kjente z -verdien, eller mindre enn den minste kjente z -verdien (Longley m.fl, 2005). Et annet hensyn som må tas i bruk av IDW ble kommentert i en studie av Bater og Coops (2009) som fant at hvilke parametere som benyttes i IDW-interpoleringen har stor betydning for feilkildene i den resulterende høydemodellen. Det ble ikke foretatt noen ytterligere

justering på parameterne tilgjengelig i verktøyet og standardverdiene ArcGIS foreslo ble benyttet.

2.2.3 Romlig autokorrelasjon

Romlig autokorrelasjon (engelsk: *spatial autocorrelation*) kan defineres ved hvilken grad et sett av romlige objekter og dets tilknyttede verdier heller mot å være geografisk konsentrert i rom (positiv romlig autokorrelasjon) eller spredt (negativ romlig autokorrelasjon) sammenlignet med forventet tilfeldig fordeling (Legendre 1993, ESRI udatert (b)). Et sentralt sitat i sammenheng med romlig autokorrelasjon er Tobler (1970, sitert i Pike m.fl 2008, s.24) “Everything is related to everything else, but near things are more related than distant things”

Oversatt til norsk vil det si at alle objekter er beslektet, men nære objekter er mer beslektet med hverandre enn fjerne objekter. For eksempel er det mer sannsynlig at temperaturen i nabokommunene Stavanger og Randaberg er like, enn temperaturen mellom Stavanger og Barcelona. Romlig autokorrelasjon er et sentralt begrep som er berørt ved flere anledninger i oppgaven. Først i interpoleringen fra punktsky til høydemodell, videre i OLS-regresjonene for vurdering av residualer samt definering av terskelverdien for avdekking a romlige mønster i Hot Spot analysene (kapittel 3.1.1).

2.3 Romlige mønster

Romlige mønster kan identifiseres ved å se på hvordan geografiske objekter er fordelt utover et areal. Det kan være både positiv og negativ romlig autokorrelasjon, som viser konsentrasjon eller spredning av objektene. Samtidig kan romlige mønstre også identifiseres ved å studere objektene tilknyttede verdier. Adressepunktene i oppgaven, har alle en sum for erstatningsutbetalinger. Disse verdiene er ikke nødvendigvis spredd jevnt over hele kommunen. I noen områder kan det være geografiske konsentrasjoner av høye erstatningsutbetalinger mens denne tendensen ikke kan spores i andre områder. En metodisk tilnærming til å identifisere romlige mønstre er kalt Hot Spot analyse. En slik analyse vil kunne besvare spørsmål om hvorvidt objektene eller verdiene tilknyttet disse er geografisk konsentrert (ESRI, 2014a).

Hot spot analyser kan brukes på flere områder, blant annet er det populært på analyser av kriminelle hendelser (Eck m.fl, 2005). For mitt formål, vil analysene fokusere på henholdsvis erstatningsutbetalinger og intensitet av hendelser. Et hot spot vil i dette tilfelle bety at det finnes geografiske konsentrasjoner hvor høye erstatningsutbetalinger opptrer ved siden av hverandre. Implisitt i navnet på analysen vil dette også bety at dersom det motsatte mønsteret

identifiseres, at lave erstatningsutbetalinger er geografisk konsentrert vil det være et cold spot. Analysen kan foretas på flere skalanivåer, alt fra gater til store byer (Eck m.fl, 2005). Hot spot analysene presentert i oppgaven vil fokusere på hele Stavanger Kommune.

Et adressepunkt med en høy erstatningsutbetaling vil i og for seg være interessant i seg selv. Men for å bli klassifisert som et hot spot, må det gitte adressepunktet være omringet av naboer som også har høye erstatningsutbetalinger. Videre må det være statistisk signifikant, noe som betyr at det med høy sannsynlighet ikke skyldes tilfeldigheter. På den andre siden av skalaen gjelder det samme, et adressepunkt med lav erstatningsutbetaling klassifiseres ikke som et cold spot uten at det er omringet av naboer med samme karakteristikk (ESRI, 2012b). ArcGIS bruker verktøyet kalt Getis-Ord G_i^* for å beregne dette, som har utgangspunkt i metoden oppfunnet av Getis og Ord (1992). Prosessen for å identifisere de romlige mønstrene gjøres via et analysevindu som 'skanner' kartet fra hjørne til hjørne. Hver gang vinduet stopper, beregnes den lokale summen innenfor analysevinduet. Med utgangspunkt i adressepunktene erstatningsutbetalinger, vil den lokale summen av variabelen bli beregnet for deretter å bli sammenlignet proporsjonalt mot summen av alle adressepunktene i hele kartet. Dersom den lokale summen er veldig annerledes fra den predikerte lokale summen, og forskjellen er så stor at det ikke kan skyldes tilfeldigheter, vil dette bli markert ved en statistisk signifikant z-verdi (ESRI, 2012b).

Z-verdien kan leses av direkte i egenskapstabellen som produseres i kjøringen av verktøyet. Videre er det også assosierte p-verdier for hver z-verdi som er beregnet, som indikerer sannsynligheten for hvorvidt det observerte mønsteret er tilfeldig eller ikke. Desto mer intens hot spotten er, desto høyere z-verdi. Vice versa gjelder for cold spots, desto mer intens cold spot, desto lavere z-verdi (ESRI, 2012b). Klassifiseringen av verdiene følger standardavviksinndelingen med fem klasser. Verdier lavere enn 1.96 og større enn -1.96 er ikke signifikante, mens verdier høyere enn 1.96 og lavere enn -1.96 vil være henholdsvis signifikante hot og cold spots (ESRI, 2007).

Hensikten med hot spot analysene er først og fremst og identifisere problemområder, men analysen kan kun svare på spørsmålet om hvor, den sier imidlertid ingenting om hvorfor. Det neste logiske steget er en OLS-regresjon (ESRI, 2013f) med formål å forklare hvorfor problemområdene er problematiske, og hvilke underliggende årsaker som ligger bak. Potensielle årsaker for dette blir presentert i påfølgende delkapitler mens resultatene presenteres i kapittel 4.3 og 4.4

2.4 Flom og overflateavrenning

Den mest klassiske og kjente definisjonen på hvordan overflateavrenning fungerer er sannsynligvis Horton (1933) som beskriver hvordan jordoverflater har en maksimal grense for hvor mye nedbør som kan absorberes. Denne grensen er priggitt infiltrasjonskapasiteten til jordoverflaten. Når denne overstiges som følge av intens nedbør, vil vannet akkumulere i overflaten og i groper som fungerer som magasiner. Disse gropmagasinene er summen av vannet som fanges i groper og sprekker på overflaten. Tre mulige scenarioer oppstår: infiltrasjon, evaporasjon eller at kapasiteten overstiges og vannet går med til overflateavrenning (Horton 1933, NVE 1997:10). Infiltrasjonen bestemmes i betydelig grad av egenskapene til grunnforholdene, vegetasjon og terrengutforming. Stavanger Kommune har som den tettest befolkede kommunen i Norge mye bebygde områder og asfalt. Tilsynelatende har også tette flater en viss grad av infiltrasjonskapasitet. Grunnet dens lave infiltrasjonskapasitet vil flaten reagere raskt, i motsetning til mer permeable flater som har en lenge reaksjonstid grunnet fordrøyning i overflaten (NVE, 1997:10).

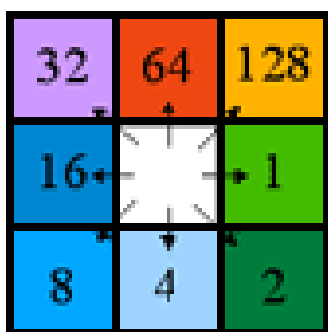
Infiltrasjonskapasitet i kombinasjon med meteorologiske forhold som temperatur og intensitet av nedbør har stor betydning for størrelsen på flom (NOU 1996:16). Samtidig er nedbørfeltets størrelse av betydning. Når det gjelder urbanhydrologiske hendelser er disse ofte knyttet til små felt. Et lite felt er ikke nødvendigvis bare karakterisert av dets areal, men også av dets høye sensitivitet overfor intens nedbør og korte varighet. I slike situasjoner vil hele feltets areal motta mye nedbør på kort tid og lede til en meget rask flom med høy intensitet, noe som gjør det vanskelig å varsle (NVE, 1997:10).

Det finnes forskjellige typer flommer. For noen av pilotkommunene i prosjektet er hovedutfordringen flom langs vassdrag. For Stavanger Kommune er skadeflommene av urban karakter. Begrepet skadeflom innebærer både skader på mennesker og menneskeskapte verdier (NOU 1996:16). Fokuset i denne oppgaven er utelukkende de økonomiske kostnadene skadeflommene forårsaker på bygninger.

2.4.1 Dreneringsretning

Vannets bevegelse er primært drevet av tyngdekraften og i mindre grad av egenskapene til materialet vannet flyter over eller igjennom. Premisset for rasteroperasjonene i ArcGIS knyttet til hydrologi er at vannet renner nedover, og på den måten følger stien med den bratteste gradienten (Olivera m.fl, 2002). Hvert nedbørsfelt har et dreneringspunkt, eller utløp. Innenfor nedbørsfeltet vil alt vannet til slutt drenere til dette punktet (Gruber og Peckham 2009, Rød 2015). Dreneringsretningen (engelsk: flow direction) er den mest grunnleggende hydrologiparameteren og danner grunnlaget for de resterende stegene i metodikken. Cellen som dreneringsretning beregnes fra (se den hvite ruten i figur 2.1), kalles den *prosesserende cellen*. Algoritmen ArcGIS benytter seg av for å bestemme dreneringsretning kalles D8 (O’Callaghan og Mark, 1984), hvor tallet åtte er antall mulige dreneringsretninger (Gruber og Peckham 2009; Rød 2015)

Dreneringsretningen blir gitt til den av de åtte naboene som har brattest gradient fra den prosesserende cellen. Hver eneste celle i rasteren har én høydeverdi tilknyttet seg som er plassert i midten av cellen (ESRI 2012a, Gruber og Peckham 2009, Rød 2015). Hvor fallet i høyde (z-verdi) er størst mellom de åtte naboene og den prosesserende cellen, vil altså bli kodet som dreneringsretningen (Rød, 2015). Dreneringsretningen er enklest forstått ved en illustrasjon:



Figur 2.1: Dreneringsretninger i algoritmen D8 (ESRI, 2012a)

Øst og sør-øst kodes 1 og 2. Sør og sør-vest kodes 4 og 8. Vest og nord-vest kodes 16 og 32. Nord og nord-øst kodes 64 og 128. Den hvite ruten i midten er den prosesserende cellen. (ESRI, 2012a).

2.4.2 Forsenkninger

Forsenkninger (engelsk: sinks) er områder som avbryter dreneringsretningen og stanser modellen. Vann kan ikke flyte forbi forsenkningen fordi ingen av nabocellene har en lavere høydeverdi (Jones, 2002). Dersom alle naboene har en høyere Z-verdi, klassifiserer ArcGIS dette som en forsenkning med udefinert dreneringsretning. Celler med udefinert dreneringsretning kan identifiseres ved funksjonen kalt *Sink*. For å generere en nøyaktig representasjon av terrengmodellen, bør gropen ¹fylles (ESRI 2011b, Rød 2015). Dette øker høyden på cellene i gropen til minimumshøyden av nabocellene slik at dreneringsretningen ikke avbrytes (Olivera m.fl, 2002). Fyllingen av groper er en filtreringsprosess som fjerner forsenkninger fra høydemodellen. Metodikken blir rettferdiggjort ved å anta at forsenkningene er feil i høydemodellen (Martz og Garbrecht 1998).

I flate områder, vil det ofte være tilfeller der nabocellene har den samme høyden som den prosesserende cellen, naboskapet blir da utvidet til cellen med brattest gradient er funnet (Olivera m.fl 2002; Rød 2015). Dersom dreneringsretningen fremdeles ikke er bestemt etter utvidelsen av naboskapet vil cellen få verdien som er summen av to retninger med lik gradient. For eksempel hvis gradienten er lik i sør-østlig og sør-vestlig retning, vil cellen få verdien 10 ($8 + 2 = 10$) (Rød, 2015).

Et forbehold er at høydemodellen må ha høy nok oppløsning for å få korrekte dreneringsretninger. Store områder av flate områder kan i tilfeller produsere unaturlige dreneringsmønstre. I slike tilfeller, bør en bedre oppløst høydemodell anskaffes for å få tilfredsstillende resultater (Olivera m.fl, 2002). Groper skyldes som regel feil i datasettet, eller naturlige forsenkninger i terrenget. Sistnevnte er sjelden å finne i høydedata dersom cellerestørrelsen er 10 meter eller større, med unntak av bre- og kalksteinlandskap (Rød 2015, Mark 1988). Det kan imidlertid finnes flere naturlige forsenkninger i dagens høydemodeller som er i stand til å representere også små terrengvariasjoner som følge av bedre målemetoder og oppløsning (Rød, 2015).

¹ Forsenkninger og groper blir brukt synonymt om samme fenomen i oppgaven

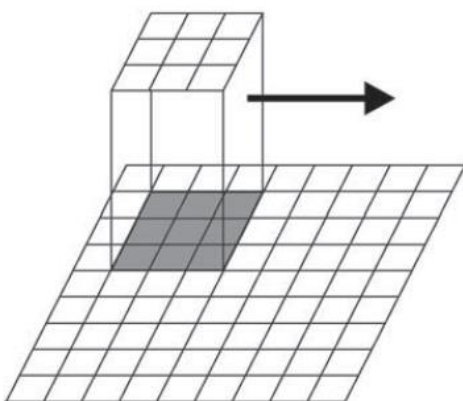
2.4.3 Akkumulering

Akkumulering av overflatestrømming (engelsk: flow accumulation) er kalkulert fra dreneringsretningen. Akkumuleringsverktøyet teller antall celler som drenerer til en individuell celle i gridet. Den prosesserende cellen teller ikke med i beregningen. Akkumuleringen kan sees på som et dreneringsområde målt i antall gridceller (Olivera m.fl, 2002).

Hver celle bidrar med verdien én til akkumulert overflatestrømming såfremt ikke et vektraster er benyttet. Det er det ikke i denne oppgaven. Et vektraster er vanligvis et nedbørsraster hvor verdiene er interpolert for en jevn overflate. Resultatene uten vektraster gir svar på hvor det sannsynlig finnes områder med mye overflatevann og områder det forventes å finne lite overflatevann. Man får ikke rede på hvor mye volum overflatevann som forventes i akkumuleringsrasteret uten et vektraster (Rød, 2015).

2.5 Terrengparametere

En terrengparameter, er et deskriptivt mål av overflaten derivert fra en høydemodell (Pike m.fl, 2008). En kategori av terrengparametere er geometriske parametere som er basert på de geometriske egenskapene til jordoverflaten. Dette inkluderer blant annet skråning, aspekt og kurvatur (Olaya, 2009). Disse er vanligvis i et rasterformat med samme omriss som høydemodellen de er derivert fra. Ekstrahering av de nevnte parameterne beregnes lokalt av et 3x3 analysevindu, et såkalt sub-grid. Dette vinduet skanner hele kartet fra øverst til venstre til nederst til høyre og gjentar den matematiske beregningen hver gang 3x3 vinduet flytter på seg (Pike m.fl, 2008)



Figur 2.2: 3x3 vindu som beveger seg over høydemodellen (Olaya, 2009)

I figuren vil det store rutenettet representere høydemodellen, mens 3x3 vinduet beregner en gitt terrengparameter i flere omganger.

2.5.1 Skråning

Det finnes flere alternativer for hvordan skråningen blir målt, avhengig av programvarepakke og brukerbestemte løsninger. I ArcGIS kan skråning bli målt i grader, som varierer mellom 0 og 90° eller som prosent (Longley m.fl, 2005).

ArcGIS beregner skråningen direkte fra DEM med 3x3 vinduet, og sammenligner høyden på den prosesserende cellen med de åtte naboene (Longley m.fl, 2005). For hver celle blir dermed skråningen beregnet som den maksimale endringen av z-verdi fra den prosesserende cellen til sine naboer. Desto høyere skråningsverdi, desto brattere er terrenget. Vice versa, desto lavere skråningsverdi, desto slakere er terrenget (ESRI, 2011c).

Skråning er en viktig bestanddel i mange modelleringer gjort via GIS. Blant annet er det viktig for hydrologiske analyser. Overflatens skråning påvirker avrenning på to måter, for det første ved at vannhastigheten endres med en endring i skråning. For det andre påvirker skråningen magasinvolument på overflaten (NVE, 1997:10). I en studie av Thompson m.fl (2001) ble det rapportert at nøyaktigheten av skråning sank ved lavere oppløsning i høydemodellen. Studien fant også at feilkildene var størst i bratte skråninger. Det kan derfor sies at skråning er en funksjon av høydemodellen, hvor nøyaktigheten av skråningsrasteren er priggitt oppløsningen i DEM (Longley m.fl, 2005).

2.5.2 Kurvatur

Mens skråning er relativt enkel å forstå, er kurvatur mer krevende. I ArcGIS kan det produseres tre ulike raster med kurvatur: profilkurvatur, plankurvatur og en kombinasjon av disse to. Sistnevnte er standardresultatet i kurvaturverktøyet, mens de to førstnevnte er alternative valg. Profil- og plankurvatur er i samme kategori primære terrengparametere som aspekt og skråning, som direkte ekstraheres fra høydemodellen. En sekundær parameter er sammensatt av primære parametere, og kombinasjonen av profil- og plankurvatur er nettopp det (Gallant og Wilson, 2000).

Profilkurvaturen, også kalt vertikal kurvatur (Olaya, 2009), gir informasjon om landformenes form og hvorvidt disse har en konveks, konkav eller horisontal karakter. Dette har betydning for vannets hastighet nedover skråningen (Blaga, 2012). Konveks profilkurvatur indikerer en

akselerering av hastigheten, mens konkav profilkurvatur indikerer en bremsing av hastigheten (Olaya, 2009). I ArcGIS vil en positiv verdi for profilkurvatur være konkav, mens negativ verdi er konveks. Flate områder har null i verdi (ESRI, 2010).

Plankurvatur, også kalt horisontal kurvatur (Olaya, 2009) gir informasjon om konvergente eller divergente strømminger. Ifølge Olaya (2009) vil strømminger i celler med konkav kurvatur konsentreres (konvergere) mens strømminger i celler med konveks kurvatur vil spre seg (divergere). Her ser man en endring i fortegn mellom profil- og plankurvatur. Positive verdier indikerer at cellen er sideveis konveks, mens negative verdier indikerer at cellen er sideveis konkav (ESRI, 2010).

Kurvaturen som brukes i videre analyser i denne oppgaven er den sammensatte kurvaturen av profil- og plankurvatur. Dette inkluderer dermed både effekten av konkave og konvekse formers betydning for vannhastighet fra profilkurvaturen, og plankurvatures betydning for konvergens og divergens av strømminger. De positive verdiene i den sammensatte kurvaturen er konvekse overflater. Dette vil gjøre det mulig å identifisere celler som både har divergente strømminger og høyere vannhastighet (Olaya 2009; ESRI 2010).

2.5.3 Aspekt

Mens skråningsverdien er bestemt av maksimal endring i z-verdi, identifiserer aspekt retningen på skråningen (ESRI, 2011a). Skråning og aspekt er nært beslektet, både fysisk og analytisk (DeMers, 2005). Vann og andre materialer som flyter nedover en skråning vil følge retningen som er bestemt av aspektet. Dette er grunnlaget for de fleste dreneringsretning-algoritmer, blant annet D8 metoden brukt i denne oppgaven. (Olaya, 2009).

Resultatet etter kjøring av aspektverktøyet er en raster inndelt i kompassretninger. Fra 0° til 360°, hvor direkte nord både er fra 0-22.5° og fra 337.5-360°, altså til sammen 45°. Resterende kompassretninger (NV, V, SV, S, SØ, Ø, NØ) er også delt inn i et intervall på 45 grader hver. Celler i inngangsdataene som er flate, får tilegnet verdien -1 for aspekt. Dersom celler har NoData, får disse tilegnet samme verdi som den prosesserende cellen innenfor 3x3 vinduet. Hvor den prosesserende cellen har NoData vil resultatet være en ikke definerbar retning (ESRI, 2011a). Hvis for eksempel Z-verdien er størst og lavest i henholdsvis øvre venstre og nedre høyre hjørne i analysevinduet vil dette være et sørøstlig aspekt (DeMers, 2005).

På nordlige breddegrader pleier områder med sydvendt aspekt å være varmere og tørre enn skråninger som har nordvendt aspekt. Dette har blant annet med solstråling å gjøre, da sydvendte aspekt (for eksempel 180°) får mer innstråling enn de som vender mot nord (ESRI udatert(c); Olaya 2009). I tillegg til dreneringsretning og innstråling har aspekt også betydning for evapotranspirasjon (summen av fordampning og transpirasjon fra planter) (Gallant og Wilson, 2000).

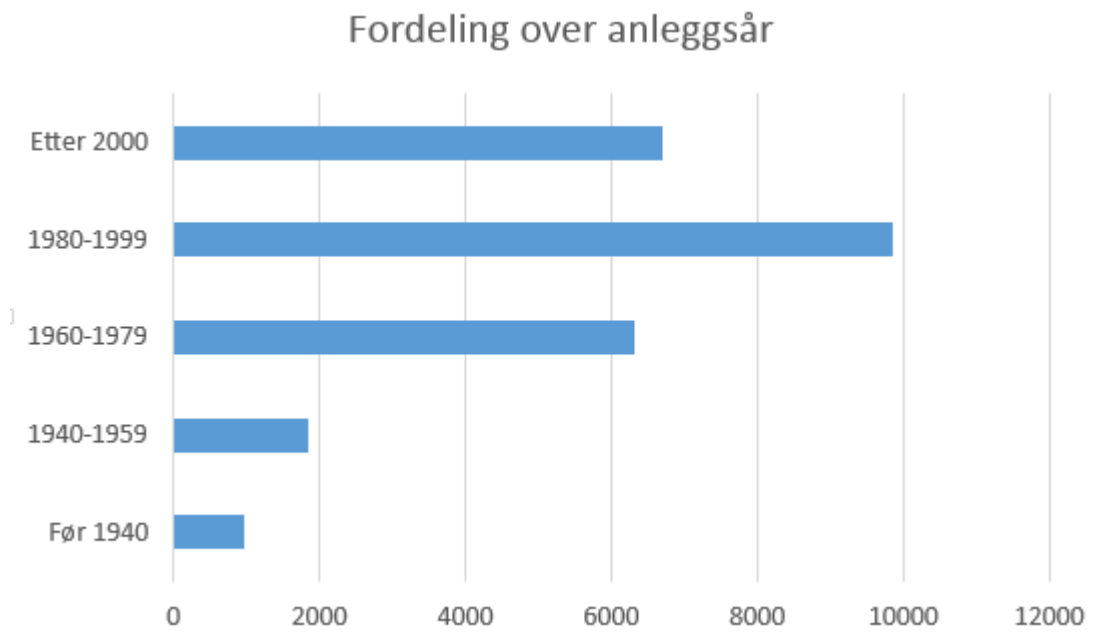
Alle de nevnte terrengparameterne samt akkumuleringen vil bli benyttet som inngangsdata i regresjonen. I tillegg til dette, vil også egenskaper tilknyttet avløpsnettets bli inkludert.

2.6 Avløpsledninger

Det finnes tre typer ledninger som kategoriseres som avløpsledninger. Overvannsledninger, spillvannsledninger og felles avløpsledninger. De to førstnevnte blir gjerne samlet sett kalt separat system. Mens sistnevnte kalles felles system. Felles system er naturlig nok hvor spillvann og overflatevann går i samme rør. Disse er gjerne av eldre årgang. Nedbøren ledes direkte ned i samme rør hvor spillvann og overflatevannet går og sendes til renseanlegg. Overvannsledninger frakter overflatevannet, altså alt regnvann som samles opp av sluk i veier og diker. Deretter fraktes overflatevannet videre til resipient. Dette er gjerne større bekker, elver eller naturlige innsjøer. Spillvannsledninger er avløpsvann fra industri og næringsvirksomhet og husholdninger. Dette innebærer vann fra toaletter, oppvaskvann, dusj og annet fettholdig eller spesielle avløp fra industrien. Resipienten her er et renseanlegg (Fredrikstad Kommune, 2007).

I møte med et stadig tøffere nedbørsregime og økende andel impermeable flater, påpeker rapporten fra Norsk Vann (2013b) at avløpsledningene er under stort press, hvor det er mangel på kapasitet på særlig felles- og overvannsledningene. Foruten utfordringer knyttet til ledningsbrudd og lekkasjer, er også tilbakeslag et stort problem. De avløpsspesifikke variablene er hentet ut fra ledningsdatasettet gjort tilgjengelig av Stavanger Kommune. Datasettet inneholder informasjon om blant annet materialtype, separat eller felles system og anleggsår for ledningene. Sammenkobling av disse egenskapsdataene og datasettet med skadedata vil være interessant for analyseformål.

Alderen på ledningene varierer fra 1866-2014, og er i diagrammet under delt inn i fem kategorier.



Figur 2.3: Antall ledninger innenfor gitte årstalls kategorier i Stavanger Kommune

Eldre ledninger trenger ikke nødvendigvis være ensbetydende med en høyere skadefrekvens enn ledninger av nyere dato. Blant annet er anleggsmetode og rørkvalitet viktige faktorer for ledningens tilstand (Norsk Vann, 2013b). Særlig er etterkrigstiden preget av ledninger med dårlig byggemateriale og kvalitet, mens det på slutten av 1970-tallet bedret seg kraftig (Eigersund Kommune, 2014).

2.7 Regresjonsanalyse

Regresjonsanalyse er en statistikkmetode som har som hensikt å evaluere forholdet mellom to, eller flere variabler. For å forklare variasjonen i den avhengige variablelen forutsettes det at man bruker fornuftige variabler som er effektive i modelleringen. Disse variablene bør aller helst ha støtte i etablert teori, alternativt basert på sunn fornuft eller at variabelen som undersøkes er fokus for oppgaven. En av de mest kjente regresjonsanalyser brukt på romlige forhold er minste kvadraters metode (ESRI, 2013f).

2.7.1 Minste kvadraters metode (MKM)

Minste kvadraters metode (engelsk: Ordinary Least Squares), heretter omtalt som OLS, er en metode brukt for å gi pålitelige resultater i beregning av lineære forhold (ESRI, 2013f). Dette oppnås ved å bruke forklarende variabler, potensielle predikatorer x , til å forklare variasjonen i den avhengige variabelen y (Legendre, 1993). De lineære forholdene mellom x og y kan være både positive og negative. For eksempel, dersom man observerer en økning i erstatningsutbetalinger samtidig som akkumuleringsverdiene går opp, kan man si at forholdet er positivt. Alternativt kan det også være et positivt forhold mellom de nevnte, ved å observere at erstatningsutbetalingene går ned, samtidig som akkumuleringsverdiene går ned. Negative forhold oppstår for eksempel hvis erstatningsutbetalingene går opp mens akkumuleringsverdiene går ned. I tillegg til positive og negative forhold, kan det også oppstå situasjoner hvor det ikke finnes noe forhold mellom x og y (ESRI, 2013f).

Forholdene kan beskrives matematisk ved hjelp av den lineære funksjonen:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon \quad (\text{ESRI, 2013f})$$

Hvor y representerer den avhengige variabelen. I mine regresjoner er dette en logaritmisk transformasjon av variabelen erstatningsutbetalinger. Variasjonen i den avhengige variabelen skal forklares ved hjelp av de forklarende variablene (x) som er en funksjon av y .

På høyre side av likhetstegnet vises en β . Dette er regresjonskoeffisientene som blir beregnet av OLS-verktøyet. Koeffisientene finnes for hver eneste forklarende variabel og representerer både styrken, og type forhold disse har til den avhengige variabel. Når forholdene er negative, vil dette bli symbolisert ved det negative fortegnet minus (-). Dersom koeffisientene er svært nær null, indikerer dette at det ikke finnes noe forhold mellom variablene. β_0 representerer regresjonens skjæringspunkt (engelsk: *intercept*), det vil si den forventede verdien for den avhengige variabelen dersom alle forklarende variabler var lik 0 (ESRI, 2013f).

Helt på høyre side i formelen viser symbolet ε . Dette er differansen mellom observerte y -verdier og predikerte y -verdier, og kalles residualer. Dette er den uforklarte porsjonen av den avhengige variabelen. Inngangsdataene som mates inn i regresjonsverktøyet er de observerte x - og y -verdiene, regresjonen bygger deretter den funksjonen som predikerer de kjente y -verdiene på best mulig måte. Det skjer allikevel sjelden at de predikerte og observerte y -verdiene sammenfaller perfekt. Residualstørrelsen blir i det tilfelle et mål på modellens

forklaringsevne. Dersom det finnes store avvik mellom predikert og observert verdi indikerer dette at modellen trenger mer kalibrering (ESRI, 2013f)

2.7.1.1 Begreper og tester

F-statistikk:

I multivariat regresjon med p forklarende variabler, er spørsmålet om alle regresjonskoeffisienter er null. For eksempel: $\beta_1 = \beta_2 = \dots = \beta_p = 0$

Ved å teste nullhypotesen kan vi besvare dette spørsmålet ved: $H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$

Denne hypotesen blir testet ved F-statistikk, som er innebygget i OLS regresjonen som skal kjøres i ArcGIS. Når det ikke finnes noe forhold, forventes en F-verdi på rundt 1. Hvis nullhypotesen kan forkastes, er forventningen en F-verdi større enn 1. Hvorvidt vi kan forkaste nullhypotesen avhenger av antall observasjoner n og antall forklarende variabler p . Dersom n er stor, kan en F-verdi like over 1 allikevel indikere bevis mot nullhypotesen. Til kontrast er en stor F-verdi nødvendig for å forkaste nullhypotesen hvis n er liten. For hvilken som helst verdi av n og p vil den statistiske programvaren beregne p-verdien assosiert med F-statistikken. Basert på den gitte p-verdien kan det deretter bestemmes om nullhypotesen forkastes eller ikke (James m.fl, 2013)

Dummy-variabel:

Dummy variabler er forklarende variabler som enten har verdien 0 eller 1. Dette er en nyttig måte når man ønsker å inkludere kvalitative data i regresjonsmodellen. En av de mest klassiske dummy variablene brukt i GIS og regresjon er rurale og urbane områder (ESRI, 2013g). Verdiene 0 og 1 kan sees på som enten tilstedeværelse eller fravær av en gitt variabel. Observasjonene som blir kodet 0 i dummy variabelen vil ikke ha noen påvirkning på koeffisienten, og slikt sett ingen påvirkningskraft i responsen av avhengig variabel. Mens når observasjonene er kodet 1 vil koeffisienten påvirke modellens skjæringspunkt (den forventede verdien for den avhengige variabelen dersom alle forklarende variabler var lik 0). Et relevant eksempel fra oppgaven er variabelen avløpssystem. Dette er en kvalitativ variabel som beskriver hvorvidt rørene er i et felles eller separat system. Felles og separate system er derfor de kvalitative dataene i variabelen. Når det separate systemet blir gitt verdien 1, og det fellessystemet får verdien 0 vil skjæringspunktet i modellen være en konstant for fellessystemet (koeffisienten påvirker ikke), mens det vil være en konstant pluss koeffisienten for det separate systemet (Caravaglia og Sharma, 2000).

3. Databearbeiding

3.1 Skadedata

Skadedataene ble rensket og prosessert først av Stavanger Kommune, før det ble hentet ut 6744 adressepunkter til eget bruk. Etter samtale med kommunen ble det klart at tre seleksjoner var relevante for videre analytiske formål:

Først ble det selektert ut et lag bestående kun av vannskader som var kodet med årsaksbokstavene I, G og J. Henholdsvis ytre påvirkning, stopp i avløp/tilbakeslag og drenering.

Etter den første seleksjonen var de rette årsakene i datasettet, men det fantes fremdeles flere installasjoner som ikke var relevante. Seleksjonen som ble utført hentet ut adressepunktene som hadde skader på installasjonsbokstavene G, H og I. Henholdsvis utvendig vann- og avløpsanlegg, vanninntrenging utenfra gjennom grunn og vanninntrenging utenfra over grunn.

Til slutt ble den siste seleksjonen gjort, som ferdigstilte datasettet til analysene som skulle bli foretatt. Variabelen kalt kilde betyr hvordan skaden på grunn av vannlekkasje er oppstått. Den eneste kildebokstaven som var av interesse for mitt arbeid var kodet I. Nedbør, smeltevann og grunnvann. Etter alle seleksjonene var fullført fantes det 497 adressepunkter med disse kombinasjonene, ca 7.4 prosent av det totale antallet som ble hentet ut.

Se forøvrig vedlegget i kapittel 8 for en mer utfyllende beskrivelse av de ulike variablene innenfor årsak, installasjon og kilde.

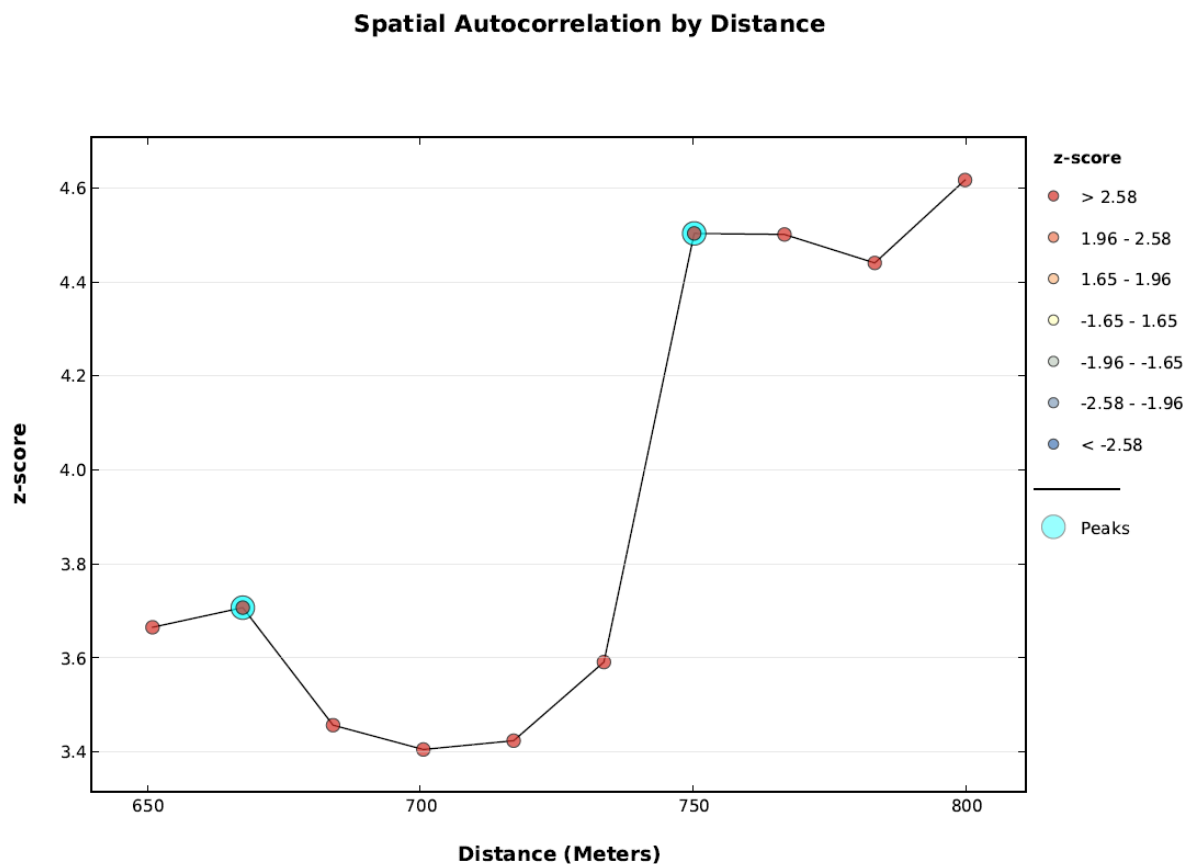
3.1.1 Identifisering av romlige mønstre i skadedata

I analysevinduet til Hot Spot verktøyet er det en mulighet å bestemme hvilken avstand som skal ligge til grunn for analysen.

Geografiske konsentrasjoner i landskapet er et resultat av underliggende romlige prosesser. Ved å finne ut på hvilken skala disse prosessene er mest fremtredende, kan det hjelpe å definere nabolagsstørrelsen som skal brukes i hot spot analysene. For å finne ut denne avstanden ble det kjørt en inkrementell versjon av et romlig autokorrelasjonsverktøy. Inkrementell vil si at verktøyet Global Moran's I autokorrelasjon, kjøres flere ganger på et intervall av avstander. Denne avstanden blir gradvis økt til den brukerdefinerte grensen. Resultatfilen er en graf som viser hvor intensiteten av romlig autokorrelasjon er sterkest. Toppunktene i grafen illustrerer på hvilken avstand dette er, markert med en lyseblå sirkel,

som viser hvor romlig autokorrelasjon er mest fremtredende (se figur 3.1). Flere toppunkter kan oppstå i samme graf, og det blir da et vurderingsspørsmål hvilken statistisk signifikant avstand som er mest hensiktsmessig å bruke i hot spot analysen. En strategi for å velge avstanden er den som best reflekterer skalaen av oppgaven, som ESRI (2013c) påpeker ofte er det første toppunktet (ESRI 2013c, ESRI 2015).

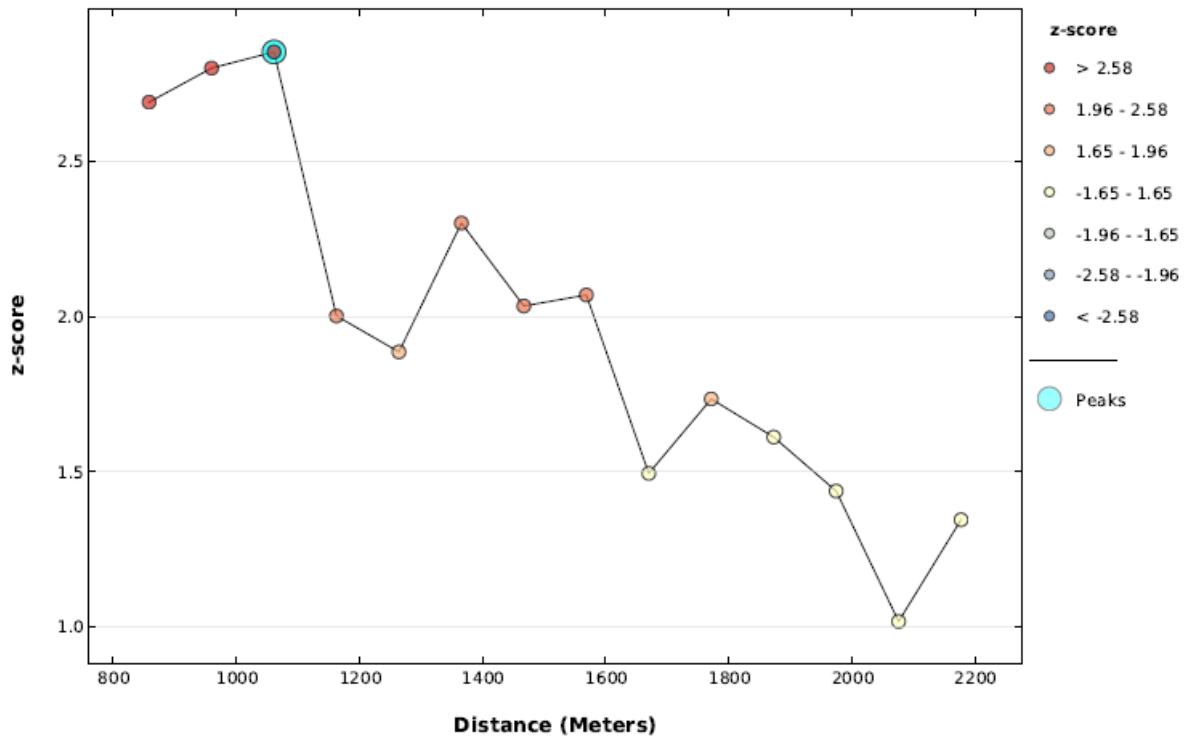
Som vist i grafen under finnes det to toppunkter, på henholdsvis 667 og 750 meter. Anbefalingen fra ESRI ble derfor tatt til etterretning, og valget for den første Hot Spot analysen ble derfor 667 meter.



Figur 3.1: Toppunkter hvor romlig autokorrelasjon er størst. Den første på 667m ble valgt

Denne avstanden ble imidlertid kun brukt på den første Hot Spot analysen for erstatninger som inkluderte alle de 6744 adressepunktene. Det inkrementelle autokorrelasjonsverktøyet ble kjørt på nytt når antall observasjoner gikk ned til 497.

Spatial Autocorrelation by Distance



Figur 3.2: Toppunkt hvor romlig autokorrelasjon er størst. Den eneste på 1061m ble valgt

I noen tilfeller, vil det kun være ett statistisk signifikant toppunkt, fordi det første og største toppunktet er funnet på samme avstand (ESRI, 2013c). Dette var tilfelle her, da det eneste toppunktet ble markert på 1061 meter, som da naturlig nok ble brukt som avstandsalternativet i den påfølgende hot spot analysen.

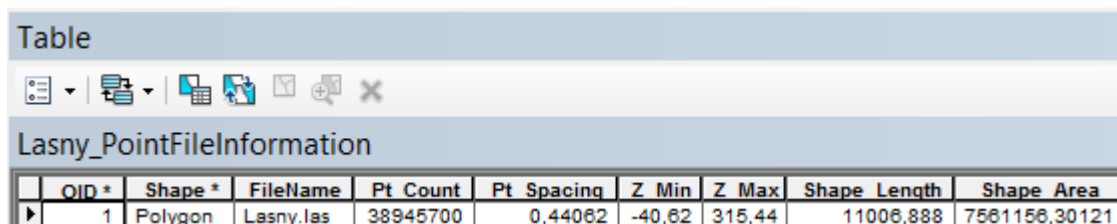
Merk at figur 3.1 og 3.2 kun gjelder hvor erstatningsutbetalingene er analysefeltet. For intensitet av hendelser fantes det ikke noen analysefelt som kunne benyttes. Målet var å ta utgangspunkt i adressepunktene, og ikke verdiene tilknyttet dem. For å få til en hot spot analyse med intensitet av hendelser måtte datasettet bearbeides på en annen måte. Et aggregeringsverktøy kalt *integrate* ble brukt for å vekte adressepunkter som lå innenfor en spesifisert distanse av hverandre (ESRI, 2013d). Distansen kalt XY toleranse er brukerdefinert og ble satt til 15 meter for hot spot analysen som gjaldt alle adressepunktene. Bygninger som er innenfor XY toleransen på 15 meter blir dermed vektet i henhold til antallet bygninger. Et verktøy kalt *collect events* (ESRI, 2009) ble deretter benyttet, hvor inngangsdataene var de aggregerte adressepunktene fra forrige operasjon. På bakgrunn av vektingen som ble tildelt, henter verktøyet ut antallet og oppretter et egenskapsfelt som holder

oversikten over dette. Dette feltet ble således benyttet som inngangsdata i hot spot verktøyet som da kan måle intensiteten av hendelsene.

Samme fremgangsmåte ble benyttet når hot spot analysen skulle beregne de selekterte adressepunktene med relevante årsaker for installasjoner, kilde og årsaker i henhold til kapittel 3.1. Ettersom dette ga langt færre adressepunkter måtte XY toleransen utvides, slik at adressepunktene hadde naboer. Toleransen ble utvidet til 75 meter, mer om dette i kapittel 5.1.2

3.2 Bearbeiding av punktsky

Punktskyen fra Stavanger Kommune ble levert i et LAS-format. Det fantes ikke metadata tilknyttet dataene, noe som gjør det vanskelig å vurdere filtreringsmetoden som er brukt for å sortere ut bakkepunktene. Bedriften COWI sto for både innmålingen og datavask før punktskyen ble levert til Stavanger Kommune. Selv om metadata manglet, finnes det verktøy i ArcGIS som kan benyttes for å hente ut statistikk fra punktskyen. Dette gir blant annet informasjon om antall punkter, areal på området og punktavstand.



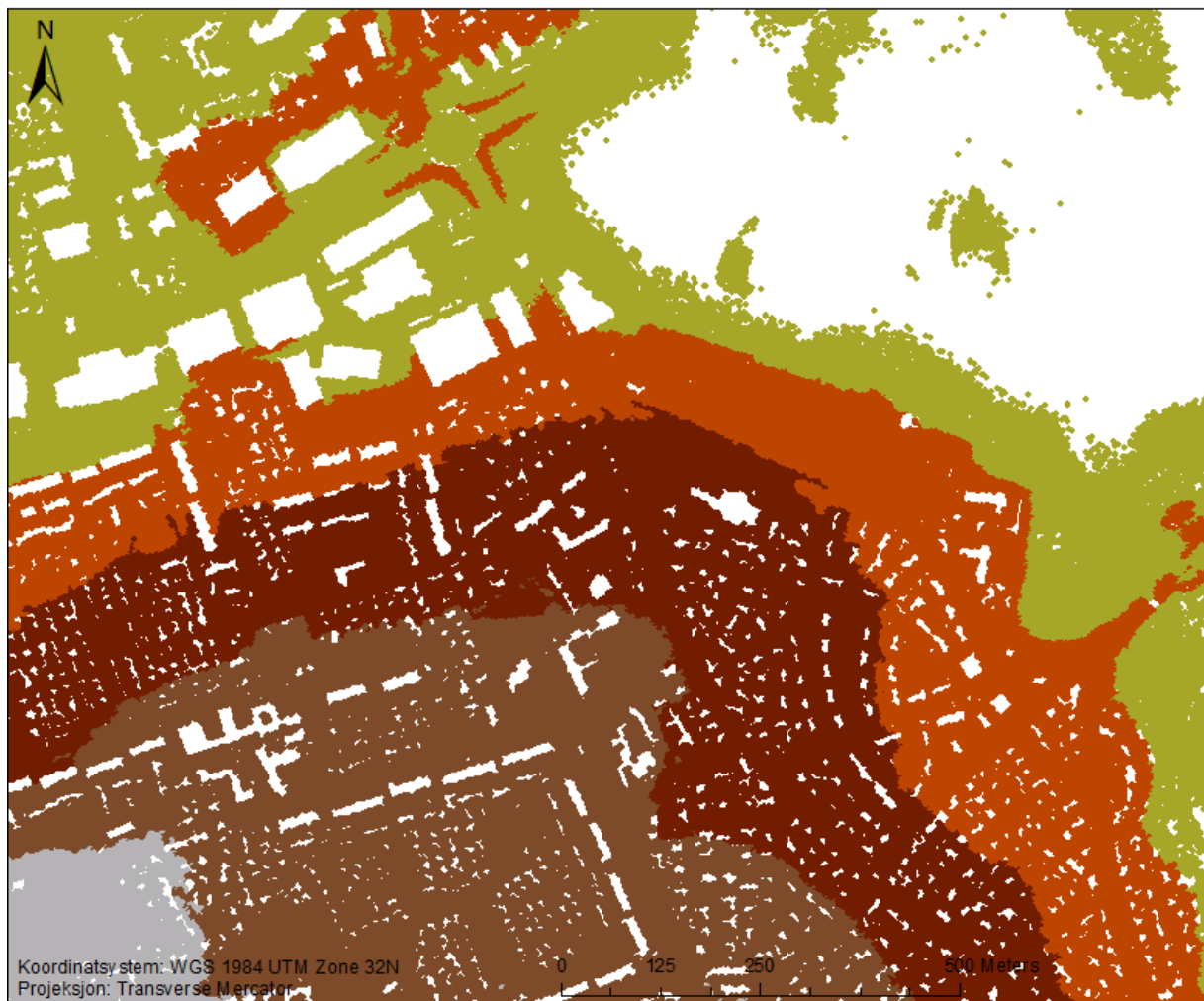
OID *	Shape *	FileName	Pt Count	Pt Spacing	Z Min	Z Max	Shape Length	Shape Area
1	Polygon	Lasny.las	38945700	0,44062	-40,62	315,44	11008,888	7581156,30121

Figur 3.3: Skjermdump fra ArcGIS med resultat av punktfilinformasjon-resultatene

Som det går frem av figuren over ser man at det finnes 38 945 700 punkter. Ettersom punktskyen er filtrert på bakkepunkter, er dette antall punkter som representerer den nakne jordoverflaten i utsnittet. Punktavstanden (engelsk: point spacing) er på 44 centimeter. Punktavstand må ikke forveksles med punktettheten. Punktavstand (PA) er definert som lineære enheter per punkt, mens punktetthet (PT) er punktene per kvadratenhet, i dette tilfellet kvadratmeter. Man kan beregne punktetthet direkte fra punktavstand med følgende formel (ESRI, 2014b):

$$\text{Punkttetthet} = \frac{1}{(\text{Punktavstand})^2}$$

Altså er punktettheten $1/(0,44)^2 = 5.15$ punkter per m^2 (ESRI, 2014b). Merk at dette tallet ikke er noen fasit for punktetthet, da punktettheten ikke er jevn over hele datasettet. Den gjennomsnittlige punktavstanden er 0.44m, men ettersom dataene er filtrert på bakkepunkter vil det mangle målinger hvor bygninger og vann ligger. Figur 3.4 er en visuell fremstilling hvor dette er tydelig. Dette vil naturligvis bety at punktettheten er mindre her, enn i resterende områder. Når punktskyen skal konverteres til en høydemodell vil den gjennomsnittlige punktavstanden og punktettheten fungere som ledetråder på hvilken celledørrelse som er fornuftig å bruke.



Figur 3.4: Forstørret område av punktskyen for å vise bygninger og vann som er filtrert bort (hvite områder)

Punktskyen ble projisert i WGS 1984 Sone 32N og vertikaldatumet NN2000. Når punktskyen skal konverteres til høydemodell må det brukes en interpoleringsalgoritme for å estimere høydeverdier hvor målingene er ukjent.

3.2.1 Konvertering og interpolering/cellestørrelse

Hvilken interpoleringsalgoritme som brukes i konvertering fra LAS til en DEM er styrt av brukeren, og er et viktig steg for å få den best mulige høydemodellen. Generelt sett kan man si at desto høyere punkttettheten LiDAR-dataene har, jo bedre vil den konverterte høydemodellen bli. Men dette er en overforenkling da flere punkter ikke nødvendigvis gir høyere nøyaktighet og kvalitet i resultatet. Mange punkter øker dessuten prosesseringstid og lagringsplass (DeMers 2005; Liu m.fl 2007). Dersom man har lav punkttetthet vil det være lite hensiktsmessig og velge en liten cellestørrelse for å oppnå en høyoppløselig DEM. Vice versa gjelder hvis inngangsdataene har høy punkttetthet og man velger en DEM med stor cellestørrelse (Liu m.fl, 2007). Valg av oppløsning er derfor dels avhengig av punkttettheten til inngangsdataen. Dessuten vil en liten cellestørrelse gi økt detaljeringsgrad i DEM, og i forlengning hvor nøyaktige de ekstraherte terrengparameterne blir (Liu m.fl 2007; Hengl 2006; Liu 2008).

Valg av cellestørrelse er en viktig del av konverteringen fra punktsky til høydemodell. Hva som er den fornuftige cellestørrelsen er det delte meninger om. I følge ESRI (2013a) er en fornuftig cellestørrelse fire ganger punktavstanden, noe som ville blitt 1,75meter i dette tilfelle. Formelen til McCullagh (1988, sitert i Liu 2008, s.41) foreslår at en fornuftig oppløsning i DEM kan bestemmes av:

$$s = \sqrt{\frac{A}{n}}$$

Hvor størrelsen på cellen (s) blir beregnet av arealet (A) og antall punkter (n). For mine data ville dette resultert i en cellestørrelse på ca. 0,5 meter. Hengl (2006) påpeker at det ikke finnes noen ideell cellestørrelse, men at dette avhenger av formålet med analysen og terrengformasjonen i den virkelige verden. Flomveirapporten utført for Trondheim Kommune av Meiforth (2013) brukte 1 meter cellestørrelse for en punktsky med 0,96m punktavstand, mens en studie for samme område med punktavstand på 0.60m brukte også 1 meter cellestørrelse samt IDW interpolering (Roalkvam, 2014).

For denne oppgaven ble det valgt en oppløsning på 1x1meter. En kompromissløsning mellom forslaget til ESRI og McCullagh ble benyttet, da dette både gir høy detaljeringsgrad av terrenget samtidig som det ikke blir for tungt å prosessere. Etter konverteringen var fullført

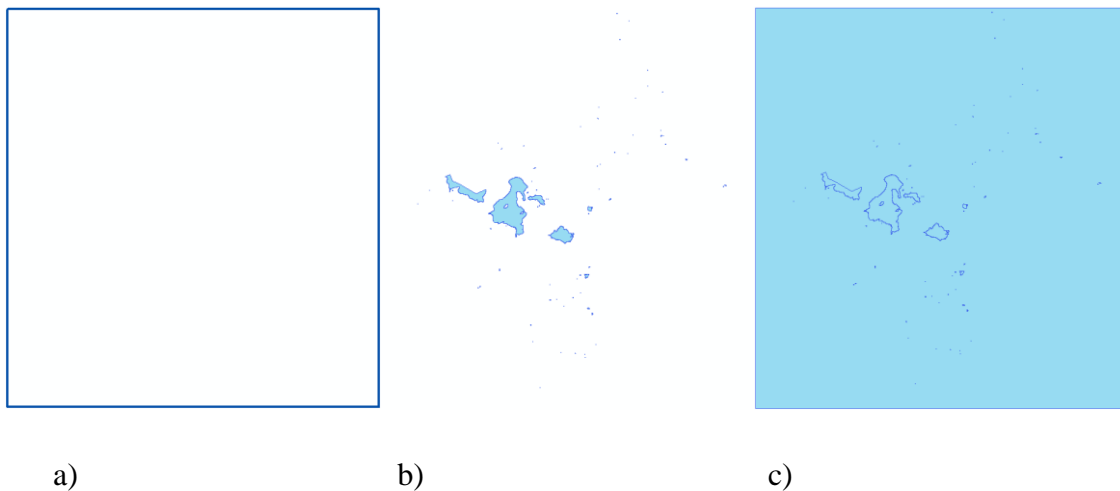
var resultatet en kontinuerlig høydemodell, men videre bearbeiding måtte til før datasettet kunne tas i bruk i hydrologiske analyser.

3.3 Bygging av hydrologisk modell

I byggingen av en hydrologisk modell, er det viktig at høydemodellen blir modifisert slik at den er hydrologisk korrekt. Dette betyr at modellen blir endret slik at de beregnede flomveiene følger de faktiske strømmingene i terrenget (Meiforth, 2013). Samme fremgangsmåte ble fulgt for å gjøre begge høydemodellene så hydrologisk korrekt som mulig, og på bakgrunn av det vil det kun bli gjengitt hvordan høydemodellen med 10m oppløsning ble bearbeidet. Det første som ble gjort var å bestemme dreneringsretninger, dette identifiserer hvor det ikke finnes definerte dreneringsretninger. Resultatet av kjøringen vil tjene som inngangsdatasett i neste steg.

3.3.1 Forsenkninger

Forsenkninger vil avbryte dreneringsretning og produsere resultater som ikke kan brukes. Disse må bli identifisert og fylt igjen (med verktøyet 'sink') eller modifisert på andre måter. Ettersom vannet i utsnittet ikke har jevn høyde fordelt over sitt vannareal må dette justeres på. Ifølge en studie av Romano (2004) blir ikke hydrografiske objekter nøyaktige i innsamlingsfasen med LiDAR, grunnet instrumentets bølgelengde. Det blir videre anbefalt at disse erstattes med en konstant høyde. Høyden var heller ikke konstant i 10x10 meter modellen. Dette ble håndtert ved å lage et nytt tomt polygon, med samme omriss som høydemodellen. Rammen måtte manuelt editeres for å inneha samme areal. Fra FKB-datasettet ble det selektert ut et eget datasett som kun inneholdt vanntemaet og høydeverdien til dette. Deretter ble rammen kombinert sammen med vandatasettet via *union* verktøyet. De to datasettene ble slått sammen, ved at union finner ut hvor vannene ligger i rammen. Resultatet blir et nytt datasett hvor vannpolygonene og høydene blir beholdt, og resterende landområder blir kodet 0.



Figur 3.5: Rammen (a) viser omrisset av kommunen, mens (b) viser seleksjon av vanntemadata og (c) hvor disse datasettene er kombinert i verktøyet union. Skjermdump fra ArcGIS (Statens Kartverk, 2015)

Etter vanntemaet var laget, som vist på figuren over, ble det konvertert fra vektor til raster. Hvor det fantes celler som inneholdt både vann- og landområder, ble cellen kodet til verdien som hadde størst areal innenfor cellen (engelsk: *maximum area*). Det vil si at hvis det var mer areal av cellen som inneholdt vann, enn land, vil hele cellen bli kodet til vann. For at konverteringen skulle passe med høydemodellen, ble alternativet *snap raster* valgt for å unngå overlappinger av celler. Dette gjør at cellene faller nøyaktig over hverandre. Samme cellestørrelse og omriss som høydemodellen ble også valgt. For at konverteringen til raster skal tildele cellene riktig verdi, er høyde brukt som verdifelt. Status på dette punktet er at det finnes to rastere, en høydemodell uten jevn høydeverdi på vann, og en raster med konstant høydeverdi på vann, men alle landområder kodet til 0. Neste steg var å slå sammen disse via rasterkalkulatoren og *Con* funksjonen. Rasterkalkulatoren gjør at man kan bygge kartalgebrafunksjoner ved hjelp av programmeringsspråk til å produsere en ny raster (ESRI, 2011d). *Con* (engelsk: *Conditional*) gjør at man kan kombinere to separate rastere ved hjelp av et uttrykk man selv setter opp. I dette tilfellet var målet å hente de konstante vannhøydene fra FKB-data, og erstatte dataene som fantes i høydemodellen. Uttrykket som ble laget for dette formålet:

```
Con("PolyRaster" == 0,"demextract", "PolyRaster")
```

Kort fortalt så betyr dette at, dersom PolyRaster datasettet har høyden 0 (altså alle landområdene), så henter den høydeverdier fra den opprinnelige høydemodellen, men dersom høyden *ikke* er 0, så henter den data fra vanndatasettet. Resultatet blir en ny raster med

konstant høyde på vanntemaet, samtidig som den bevarer høydeverdiene fra høydemodellen hvor det var landområder.

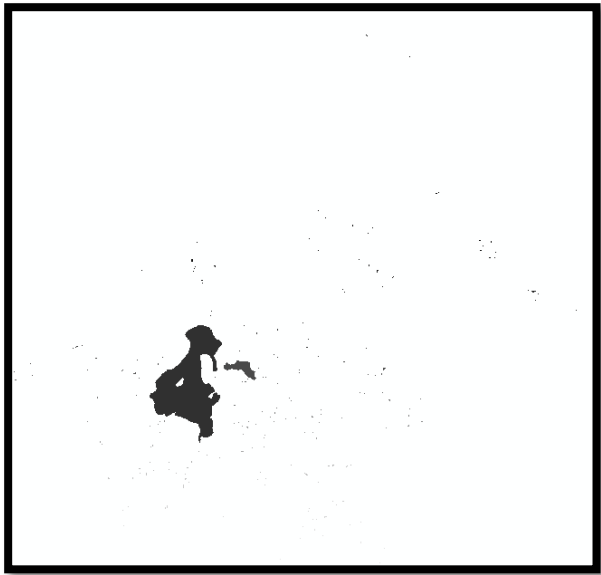
I tillegg til forsøkninger og vanntema, finnes det også andre elementer i høydemodellen som klassifiseres som forsøkninger. For en mest mulig nøyaktig modellering av overflatestrømming, bør disse forsøkningene fylles. Før man fyller igjen disse forsøkningene, må man identifisere hvor de finnes, og hvor dype disse er. Dette krever bruk av flere verktøy som kan knyttes sammen via geoprosessering i et Python skript (Rød, 2015).

```
1 # Henter inn systemmodulen ArcPy
2 import arcpy
3 from arcpy import env
4 # Definer hvor det finnes inngangsdatasett for prosessen
5 env.workspace = "C:\Users\wurdin\Desktop\GEOG3900\Datagrunnlag\Slett"
6 # Sjekk at Spatial Analyst extension er skrudd på
7 arcpy.CheckOutExtension("spatial")
8 arcpy.gp.Sink_sa("FlowDirSjo", "sinks_s")
9 arcpy.gp.Watershed_sa("FlowDirSjo", "sinks_s", "sink_areas_s", "VALUE")
10 arcpy.gp.ZonalFill_sa("sink_areas_s", "TestPolyisjo", "sink_max_s")
11 arcpy.gp.ZonalStatistics_sa("sink_areas_s", "VALUE", "Testpolyisjo", "sink_min_s",
"MINIMUM", "DATA")
12 arcpy.gp.Minus_sa("sink_max_s", "sink_min_s", "sink_depth_s")
13 # Fortell når prosessen er ferdig, og filene er klare til bruk
14 print("Ferdig")
```

Figur 3.6: Geoprosessering i form av et Python skript. Modifisert etter Rød (2015)

Geoprosesseringen produserer flere resultater, men det mest interessante er å finne terskelverdien som kan brukes for å fjerne forsøkningene. Koden er laget i Python IDLE, versjon 2.7.3 og er kompatibel med ArcGIS plattformen. Det første som gjøres er å hente inn arcpy, som gjør at Python forstår hvilken systemmodul som skal brukes. Videre defineres arbeidsfolderen, hvor datasettene som det skal kjøres geoprosessering på, ligger. Dette er henholdsvis dreneringsrasteren FlowDirSjo, og høydemodellen TestPolyisjo. Førstnevnte finnes på linje 8 og 9, mens sistnevnte finnes på linje 10 og 11.

De fem verktøyene som brukes krever at lisens for tilleggspakken Spatial Analyst er skrudd på. Dette blir kontrollert av linje 7 i skriptet. Linje 8 henter inn dreneringsrasteren og skanner alle cellene hvor det ikke finnes en definert dreneringsretning, dette blir kodet som en forsøkning.



Figur 3.7: Skjermdump fra ArcGIS. Resultat av verkøyet Sink. Grå/sorte områder indikerer områder uten definert dreneringsretning. (Statens Kartverk, 2014)

Forsenkningene er kodet med heltallsverdier (engelsk: *integer*). Hver forsenkning har sin egen unike verdi, mens naboceller som utgjør en forsenkning har samme verdi. Definerte dreneringsretninger blir kodet som NoData og er de hvite områdene som omslutter de grå områdene på figuren ovenfor. Neste steg er verktøyet kalt *watershed* som benytter forsenkningsrasteren vist i figur 3.7 som inngangsdata. Resultatet blir regioner og områder med piksler som drenerer til sine respektive forsenkninger (Rød, 2015).



Figur 3.8: Skjermdump fra ArcGIS. Viser områder som drenerer til hver sin forsenkning. (Statens Kartverk, 2014)

Videre brukes dette resultatrasteret i neste steg, som er en sonisk kartalgebrafunksjon kalt *Zonal Fill*. Verktøyet henter inn høydemodellen som inngangsdata, deretter fyller den igjen forsenkningene til minimumsverdien utenfor forsengkningsgrensen. Det vil si at dersom en forsenkning har høydeverdi på 0, og minste høydeverdi utenfor forsenkningen er 0.1, vil sistnevnte bli skrevet til resultatrasteret (*sink_max_s*). (Rød 2015, ESRI 2011b). Mens denne soniske kartalgebrafunksjonen identifiserer minste høyde utenfor forsenkning, vil neste verktøy *Zonal Statistics*, som også er en sonisk kartalgebrafunksjon, finne minste høyde som faller innenfor forsenkningen. Differansen mellom høydeverdien for kanten av forsenkningen fra *sink_max_s* og forsenkningens minste høydeverdi *sink_min_s* gir svaret på hvor mye hver forsenkning må justeres. Differansen beregnes av den lokale kartalgebrafunksjonen *Minus* som trekker fra verdien av *sink_min_s* fra den første rasteren *sink_max_s*. Resultatrasteren *sink_depth_s* angir hvordan høydemodellen bør modifiseres for å produsere mer nøyaktige overflatestrømminger og bli hydrologisk korrekt (Rød 2015).

Deretter ble verktøyet *Fill* kjørt for å fylle gropene, resultatverdien jeg fikk fra *Minus* (5,4) ble justert opp til 5,5 for å sørge for at forsengkninger også på 5,4meter ble fylt igjen. Dette fyller igjen alle groper med denne terskelverdien. Resultatrasteret inneholdt også en negativ høydeverdi på -0,7 meter. Denne høydeverdien befant seg under havnivå. I rasterkalkulatoren ble denne den negative høyden erstattet med 0 ved følgende uttrykk:

```
Con("TestPolyFill" < 0, 0, "TestPolyFill")
```

Verktøyet gjør at dersom høydemodellen har høydeverdier som er mindre enn 0, vil disse negative høydeverdiene bli tilegnet verdien 0, mens dersom den er 0 eller over 0, beholder den opprinnelige høydeverdier.

3.3.2 Heving av bygninger

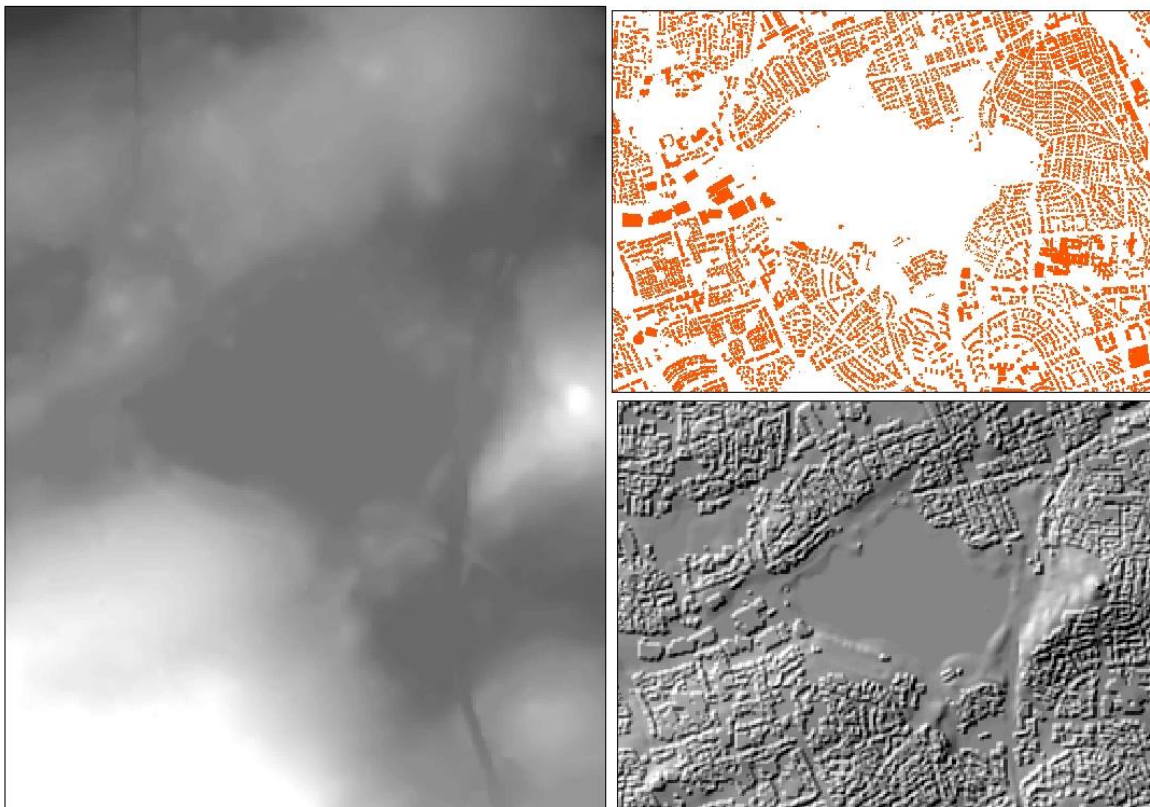
Omtrent samme metodiske tilnærming ble brukt for å heve bygninger. Etter å ha hentet inn bygningsdatasettet fra FKB ble det laget et nytt egenskapsfelt i tabellen. Det nye feltet kalt høyde, ble kalkulert slik at alle bygninger ble registrert med 5 meter høyde. Denne verdien er brukerdefinert og gjenspeiler ikke virkeligheten nøyaktig, men hensikten blir oppnådd fordi vannet blir tvunget rundt bygningene (Meiforth, 2013).

Deretter ble datasettet som var i vektorformat, konvertert til raster ved å velge det nye egenskapsfeltet som inngangsdata. Resultatet var en raster med hvite NoData områder eller

bygninger som var 5 meter høye. For å slå sammen dette med høydemodellen ble rasterkalkulatoren igjen tatt i bruk. Ved uttrykket:

```
Con(IsNull(«Bygning_Flate_Raster»), «demferdig», «demferdig» +5)
```

IsNull identifiserer hver eneste celle som er NoData, ved å returnere 1, i kombinasjon med Con blir NoData cellene gitt en verdi. Uttrykket bruker dermed høydemodellen når bygningsrasteren ikke er 5, mens hvor bygningsrasteren har verdi blir det lagt til 5 meter på høydemodellen. En skjermdump som illustrerer denne prosessen vises i figuren.

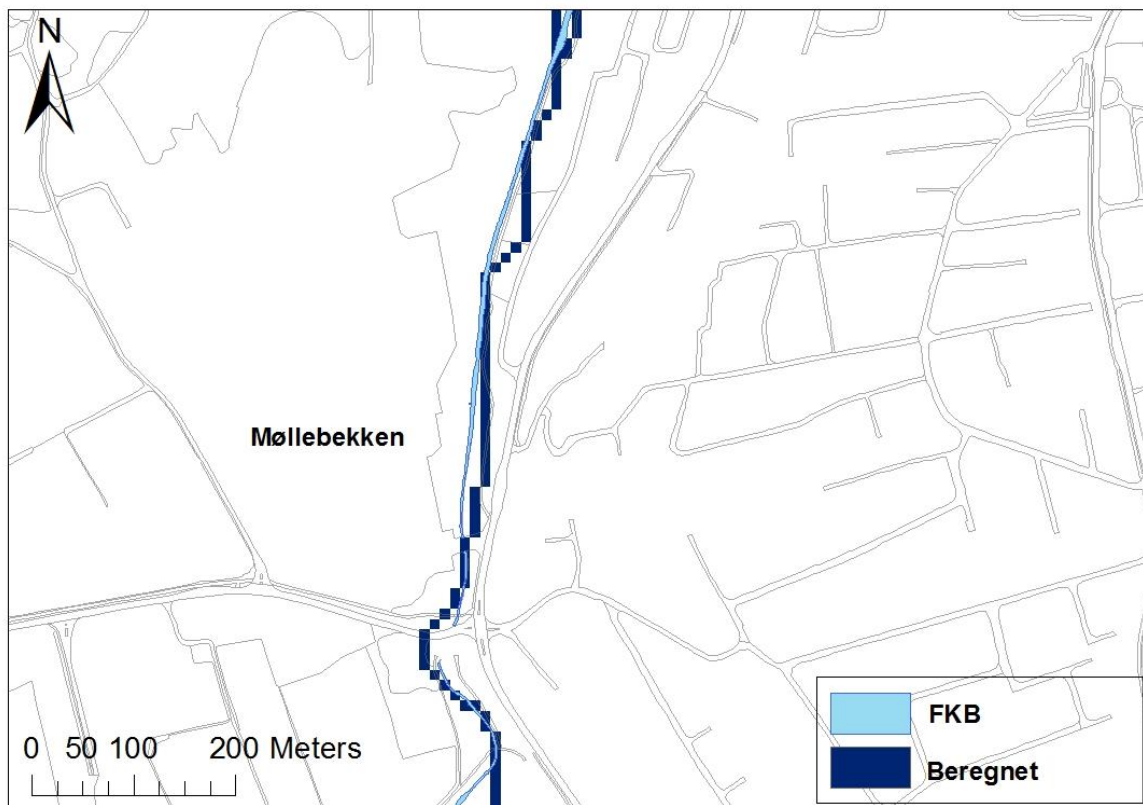


Figur 3.9: Skjermdump fra ArcGIS som viser høydemodellen til venstre, bygningsrasteren øverst til høyre og de to kombinert til en høydemodell med hevede bygninger nede til høyre. (Statens Kartverk 2014;2015)

Den kombinerte rasteren er vist som skyggekart for å gi et bedre inntrykk av relieffet. Heving av bygningene vil skape mer nøyaktige flomveiberegninger når dreneringsretningen skal bestemmes på ny.

3.3.3 Akkumulering

Etter at forsenkninger var fylt og bygningene hevet ble dreneringsretningen bestemt på ny. Deretter ble dette datasettet brukt som inngangsdata i beregning av akkumuleringsrasteren. Nedenfor vises et sammenligningsgrunnlag mellom den beregnede akkumuleringen av D8-algoritmen og data hentet fra FKB som viser Møllebekken. Som det går frem av figuren ser det ut til at det er relativt likt, samtidig bærer det beregnede resultatet preg av å være mer grovinndelt og rettvinklet, dette kan skyldes den såpass lave oppløsningen den er beregnet fra samt at den ikke er konvertert til vektor og at D8-metoden er beregnet for overflateavrenning og ikke etablerte bekker. Som man ser av figuren er FKB registreringen av Møllebekken mer naturlig buet.

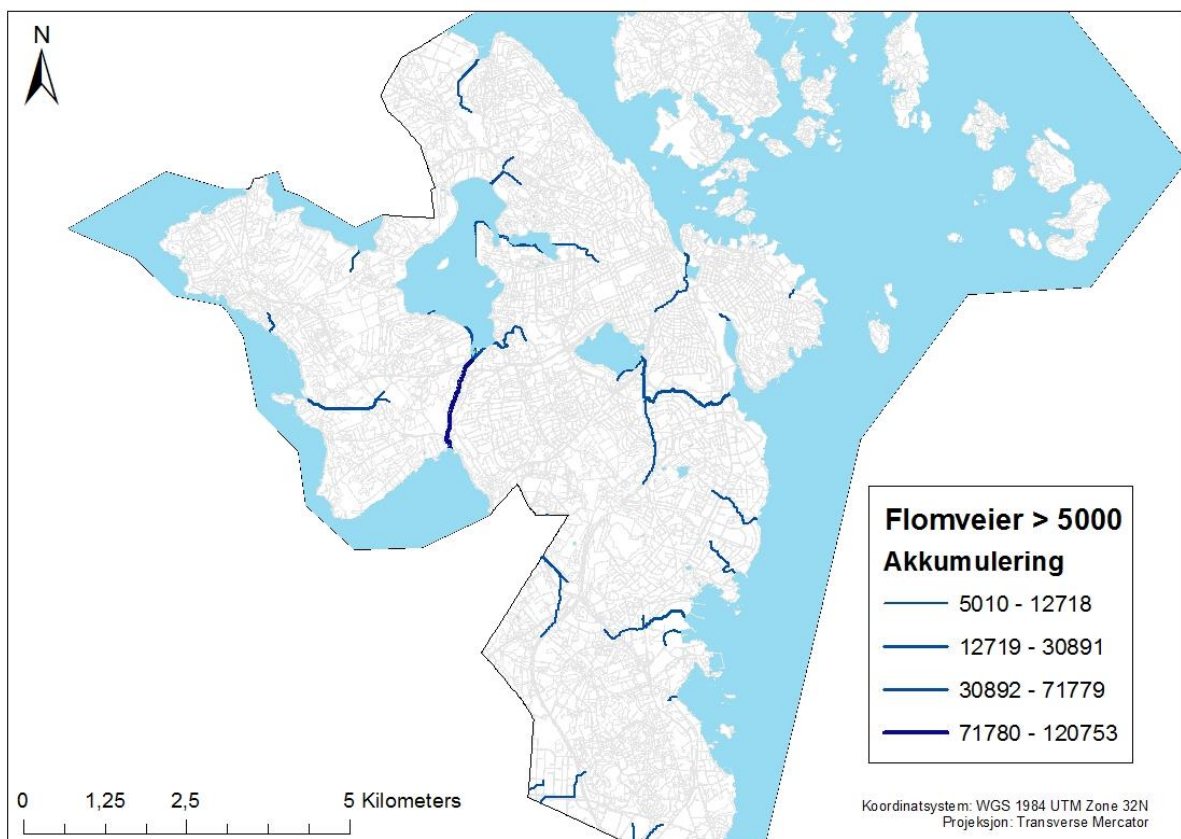


Figur 3.10: Beregnet flomvei i ArcGIS og Møllebekken fra FKB-data. (Statens Kartverk, 2015)

Verdien for akkumulering som cellene blir tildelt er antall celler som flyter inn i hver celle. Ifølge Olivera m.fl (2002) bruker den nasjonale høydedatabasen i USA (engelsk: National Elevation Dataset) en verdi på 5000 celler for å klassifisere bekker. Ved hjelp av Con verktøyet og uttrykket $VALUE > 5000$, blir alle celler med høyere akkumuleringsverdi enn 5000 klassifisert som dette. Resterende verdier blir utelatt fra denne operasjonen, og kodet til

NODATA. Denne terskelverdien er brukerdefinert og dermed kan den skreddersys til hvilke type strømninger som er ønskelig å fremstille. I figur 3.10 er det gjort nettopp dette, ved at kun celler med akkumuleringsverdi over 5000 blir presentert (Olivera m.fl, 2002). Dette ble gjort for å ha et sammenligningsgrunnlag med FKB data uten at mindre flomveier ble inkludert i kartet og lagde unødig støy i bildet.

For akkumuleringskartet over hele Stavanger Kommune ble 5000 igjen brukt som terskelverdi. Flomveier med akkumuleringsverdi under 5000 blir utelatt. Disse flomveiene vil være priggitt oppløsningen i høydemodellen og hvor hydrologisk korrekt den er bearbeidet. Flomveiene er i denne figuren konvertert til vektorformat.



Figur 3.11: Kart over flomveier med en akkumuleringsverdi på over 5000 (Statens Kartverk, 2015)

I figuren er bakgrunnskartet gjort transparent for å gjøre de beregnede flomveiene mer visualiseringsvennlig. Resterende kart over flomveier i denne oppgaven er med utgangspunkt i akkumuleringsverdier med 500 for høydemodellen med 10m oppløsning og 50 000 for 1m oppløsning.

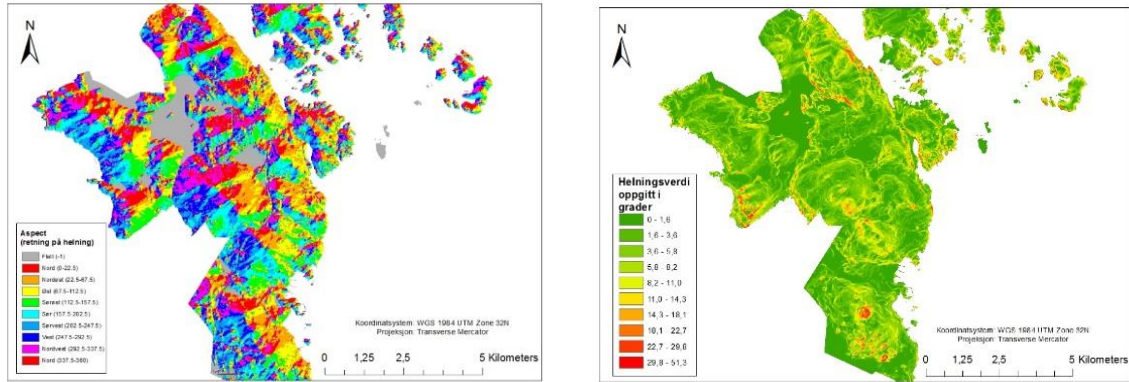
For høydemodellen med 10m blir formelen $500 \times 10 \times 10 = 50\,000 \text{ m}^2$, som skaper et sammenligningsgrunnlag med høydemodellen med 1m hvor formelen er $50\,000 \times 1 \times 1 = 50\,000 \text{ m}^2$ (Olivera m.fl, 2002)

3.4 OLS

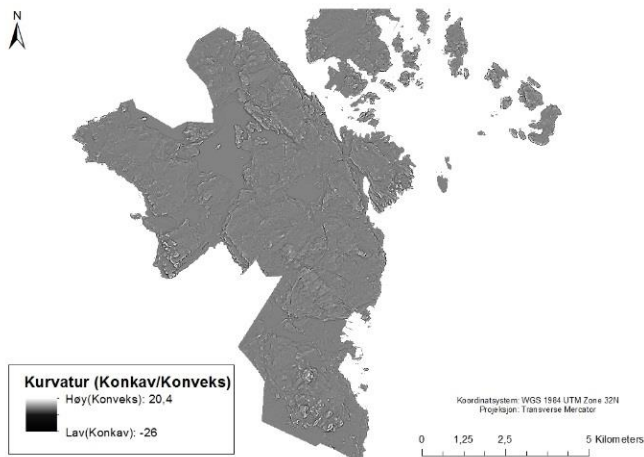
Bygging av en regresjonsmodell er en møysommelig prosess som går frem og tilbake. Det involverer mye bearbeiding for å tilpasse variablene til regresjonen. I de påfølgende delkapitlene 3.4.1 og 3.4.2 blir fremgangsmåten som er brukt beskrevet.

3.4.1 Enhet for kobling av statistiske data

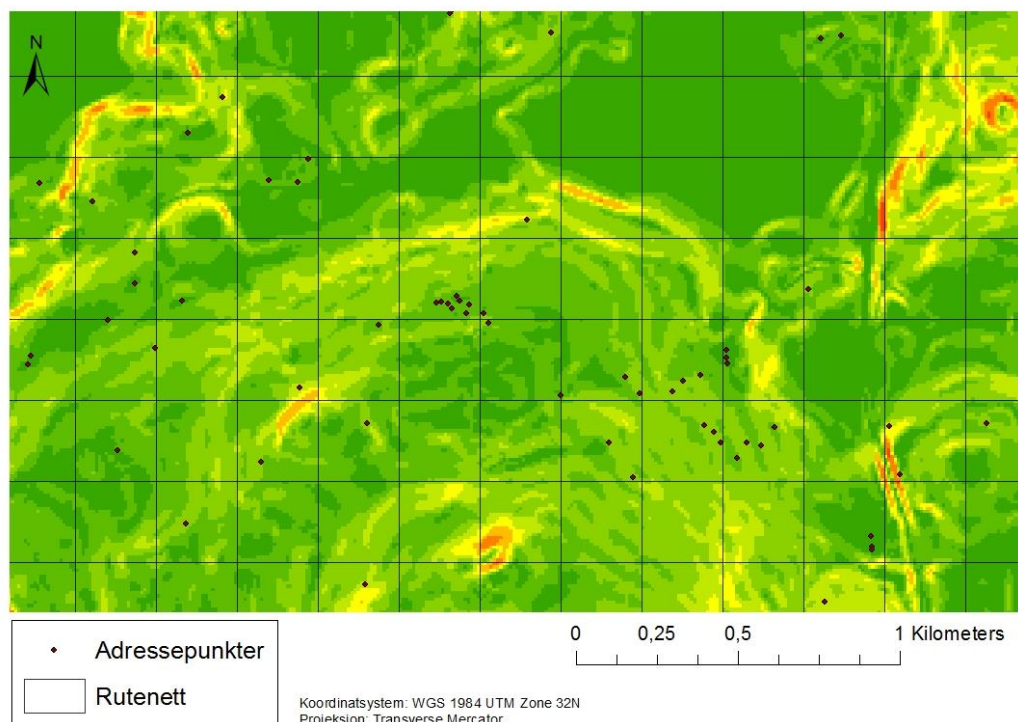
For å kombinere data fra forskjellige kilder trengs en enhet for analysen. Et rutenett (engelsk: fishnet) gjør dette mulig. Det muliggjør en kobling av både avløpsspesifikke data, terrengparametere, akkumulering og adressepunktene. Størrelsen på rutene kan spesifiseres etter eget ønske, avhengig av hva som er skalaen på analysene. Det ble valgt en rutenettstørrelse på 250x250meter (samme koordinatsystem som andre data), da dette sørget for inkludering av flere adressepunkter innenfor de fleste rutene, samt gjør adressepunktene lokalisering mer anonyme. Rutenettet ble klippet i flere omganger etter behov, første gang etter omrisset til kommunen. Medfølgende rutenettet produserte verktøyet også et punktlag, som gjør det mulig å ekstrahere verdien fra rasteren direkte til dette punktet. Men, ettersom disse punktene alltid faller midt i rutene, var ikke dette optimalt for mine analyser. I stedet fungerte selve adressepunktene som inngangsdataene når det skulle hentes ut verdiene fra terrengparameterne og akkumuleringsrasteren. Hver gang prosessen kjøres, skapes et nytt egenskapsfelt kalt RASTERVALU som holder på verdien fra rasteren og tilegner den til adressepunktet. Ettersom dette variabelnavnet er lite intuitivt, ble et mer passende variabelnavn laget i egenskapsfeltet. Videre ble verdiene fra RASTERVALU hentet over til det nye feltet med det passende navnet ved feltkalkulatoren. Denne fremgangsmåten ble gjort for aspekt, skråning, kurvatur og akkumulering i begge modellene.



Figur 3.12: Aspektraster til venstre og skråningsraster til høyre (i grader). Beregnet fra høydemodell med 10x10 meter oppløsning (Statens Kartverk, 2014)



Figur 3.13: Kurvaturraster beregnet fra høydemodell med 10x10 meter oppløsning (Statens Kartverk, 2014)



Figur 3.14: Rutenett plassert over skråningsraster og et utvalg adressepunkter (Statens Kartverk 2014)

Adressepunktene i denne figuren vil deretter kunne hente ut skråningsverdiene direkte på punktene ved ekstraheringsverktøyet. Etter adressepunktene har fått verdier for terrengparameterne, ble det brukt en romlig kobling (engelsk: spatial join) for å knytte adressepunktene, og avløpsnettets sammen med rutenettet vist i figur 3.14. Ettersom alle datasettene finnes i samme koordinatsystem, kan koblingen finne ut hvor objektene faller innenfor sine respektive ruter. I dialogvinduet til det romlige koblingsverktøyet finnes det flere matematiske funksjoner som kan brukes til å aggregere verdiene tilknyttet adressepunktene og avløpsnettets og dermed representere hele ruten. For eksempel: gitt at det finnes 10 adressepunkter innenfor en rute som hver har en erstatningsverdi tilknyttet seg, kan man summere disse verdiene slik at de totale summen blir gjeldende for ruten. Aggregeringsmetodene som ble benyttet var gjennomsnittet for skråning, kurvatur og aspekt. Sum ble brukt for erstatning, mens maksimum ble benyttet for akkumulering. Sistnevnte betyr at adressepunktet som har høyest verdi for akkumulering blir gitt til hele ruten, uavhengig om det finnes flere andre adressepunkter i samme rute.

Ledningene i datasettet er delt opp i linjesegmenter, det ble tatt et gjennomsnitt av alder for ledningssegmentene innenfor hver rute. Hvorvidt ledningene var felles eller separat ble opprinnelig kodet som en strengvariabel. OLS-regresjonen kan ikke ta tekst som

inngangsdata, derfor måtte denne bli kodet om til en dummy-variabel som kun kan ta verdien 0 eller 1. Ledninger som hadde felles system ble kodet 0, mens separate system ble kodet 1. Det fantes som regel begge typer innenfor hver rute. Typetall (engelsk: mode) ble brukt som aggregeringsfunksjon, dette er et sentralitetsmål som angir hvilke av observasjonene som forekommer hyppigst. Det vil si at hvis det finnes flere separate ledninger enn fellesledninger innenfor en gitt rute, vil hele ruten få verdien 1.

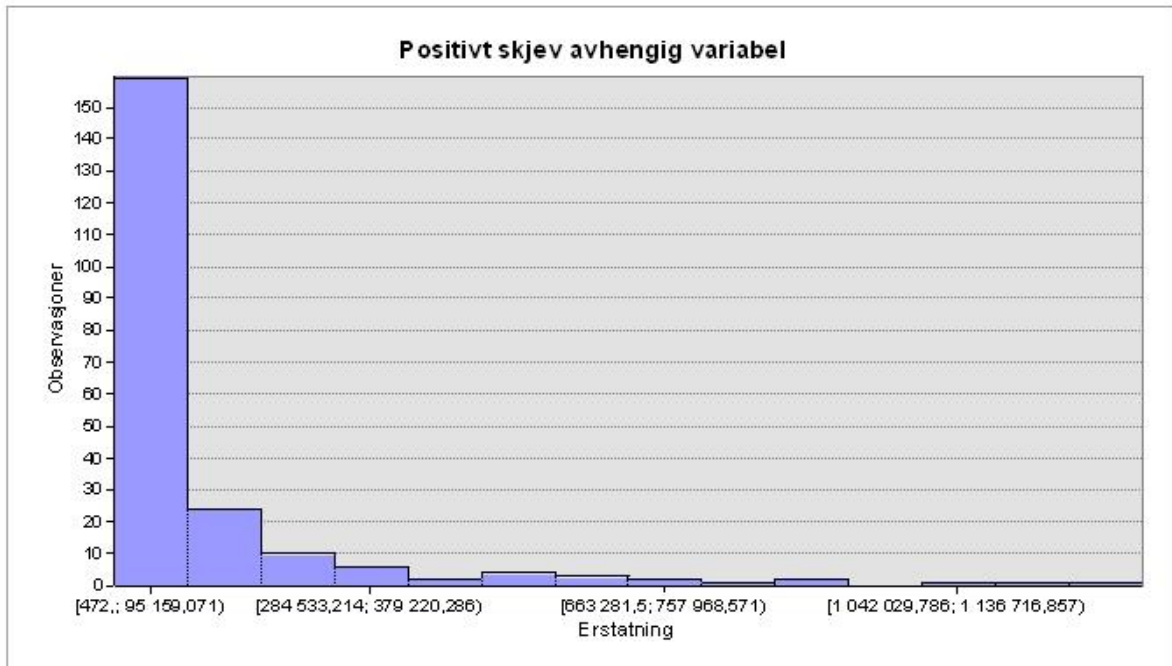
3.4.2 Transformering av variabler

I utgangspunktet var min avhengige variabel positivt skjev (se figur 3.15). Når det viser seg at den avhengige variabelen har et ikke-lineært forhold til de forklarende variablene kan det lønne seg å bruke transformasjoner. OLS er en lineær regresjonsmetode som legger til grunn at forholdene man modellerer er lineære. Når dette ikke er tilfelle, kan man transformere variablene for å se om dette skaper mer lineære forhold (ESRI, 2013g).

Log-transformasjon virker imidlertid ikke når man har nullverdier i egenskapstabellen (noen erstatningssummer viser 0). At det finnes erstatningssummer med verdien 0 kan skyldes tre ting etter mitt skjønn:

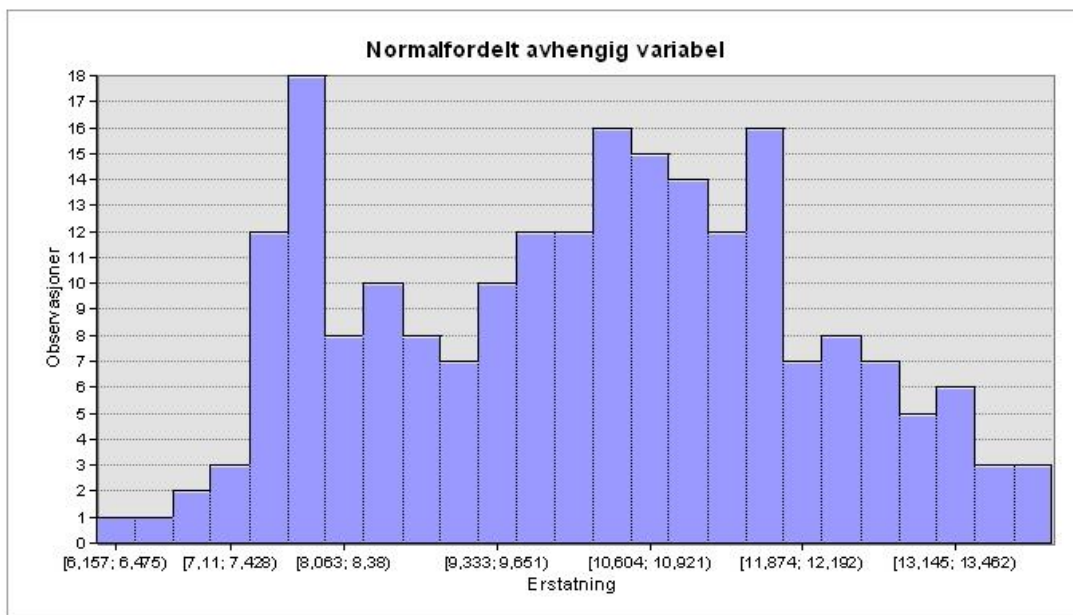
- Feilaktige data
- En sak er opprettet på skade men erstatningen er ikke registrert
- Det kan skyldes at saken er gått til regress, og opprinnelig erstatningsbeløp nullstilles (se 5.1.1) (Brevik m.fl, 2015).

Feilaktige data eller hvor erstatningsbeløpet er blitt nullstilt etter at saken er gått til regress synes mest sannsynlig. Det er lagt som forutsetning at adressepunkt med erstatningsbeløp på 0 ikke medfører riktighet, og er derfor blitt fjernet fra datasettet før OLS-regresjonen satte i gang. Det fantes også noen veldig få observasjoner med negative beløp, disse ble også fjernet. De resterende observasjonene er deretter blitt logaritmisk transformert for å få en normalfordelt kurve, og å skape mer lineære forhold til de forklarende variablene.



Figur 3.15: Avhengig variabel før transformasjon. En trend med positivt skjev fordeling

Etter log-transformasjonen kan man se en langt mer normalfordelt avhengig variabel. Dette ble oppnådd ved å legge til et nytt egenskapsfelt i datasettet. Deretter benyttes feltkalkulatoren på det nye feltet med uttrykket: `Log ([ERSTATNING])`



Figur 3.16: Logaritmisk transformert avhengig variabel som er nærmere normalfordeling

Den samme fremgangsmåten ble også fulgt for akkumuleringsvariabelen. I likhet med erstatning, hadde også denne positivt skjev fordeling. Akkumuleringsvariabelen hadde noen få 0-verdier. Disse ble imidlertid ikke slettet, isteden ble verdiene erstattet med den laveste mulige verdien, altså 1. Etter den logaritmiske transformasjonen ble forholdet manuelt inpsisert i et spredningsplott, hvor forholdet ble mer lineært med den avhengige variabelen.

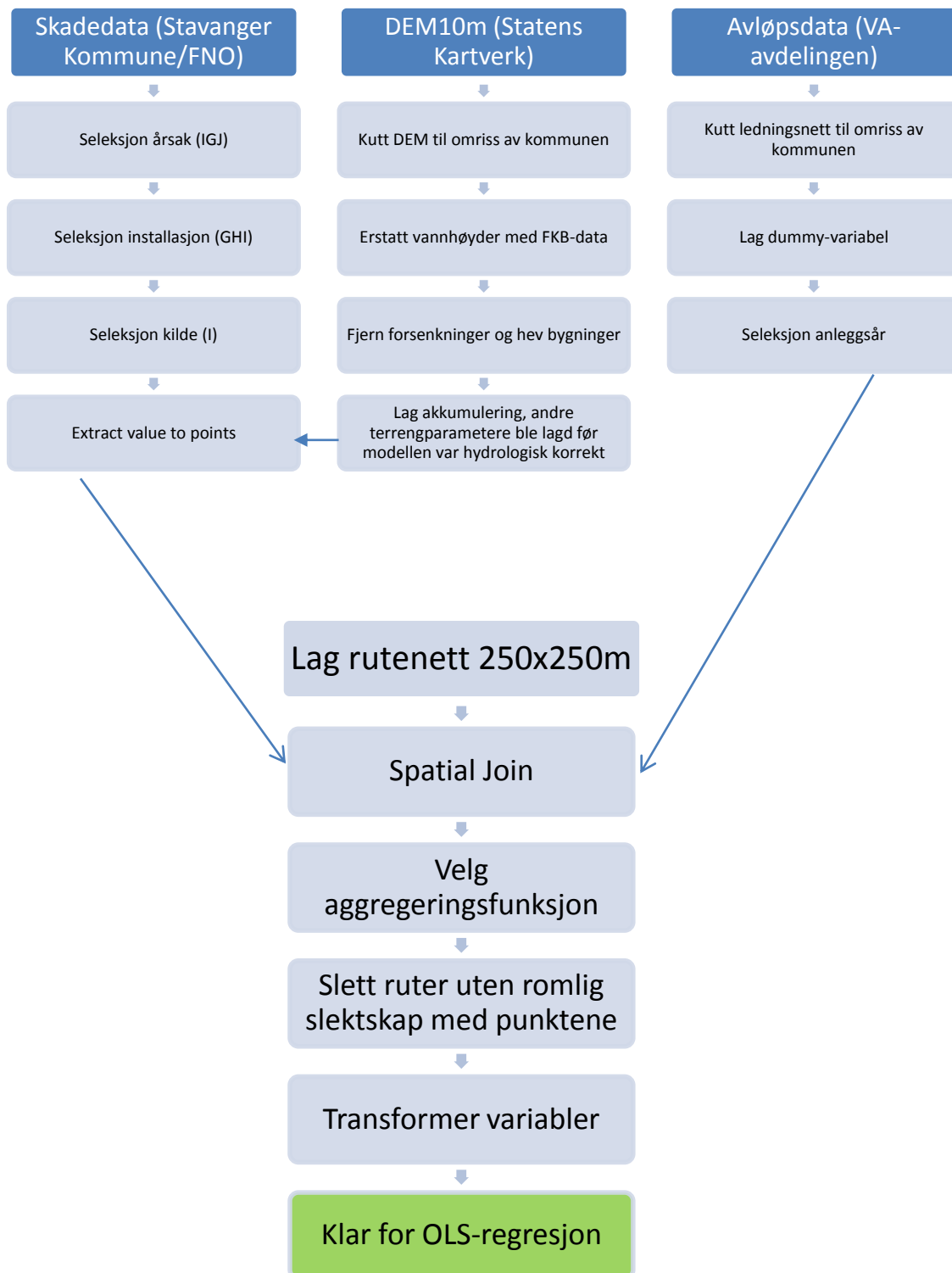
3.4.3 Oppsummering av variabelbearbeiding

Aggregeringen og transformering av variablene er omtalt i kapittel 3.4.1 og 3.4.2 men for en enklere oversikt presenteres det her i tabellform.

Tabell 1: Viser aggregeringsfunksjon og transformering brukt på variablene.

Variabel	Aggregeringsfunksjon	Tranformering
<i>Erstatning</i>	Sum	Log
<i>Skråning</i>	Gjennomsnitt	Ikke transformert
<i>Aspekt</i>	Gjennomsnitt	Ikke transformert
<i>Kurvatur</i>	Gjennomsnitt	Ikke transformert
<i>Akkumulering</i>	Maksimum	Log
<i>Anleggsår</i>	Gjennomsnitt	Ikke transformert
<i>Separat/Felles Dummy</i>	Typetall	Ikke transformert

I tillegg vises på neste side et flytskjema (figur 3.17) over bearbeidingsprosessen for å gjøre variablene klare til OLS regresjon. Skjemaet viser hvordan prosessen er blitt gjort ved høydemodellen for 10x10 meter oppløsning, mens for utsnittet ble det utført tilnærmet de samme stegene bortsett fra omrisskutt til kommuneutstrekningen.



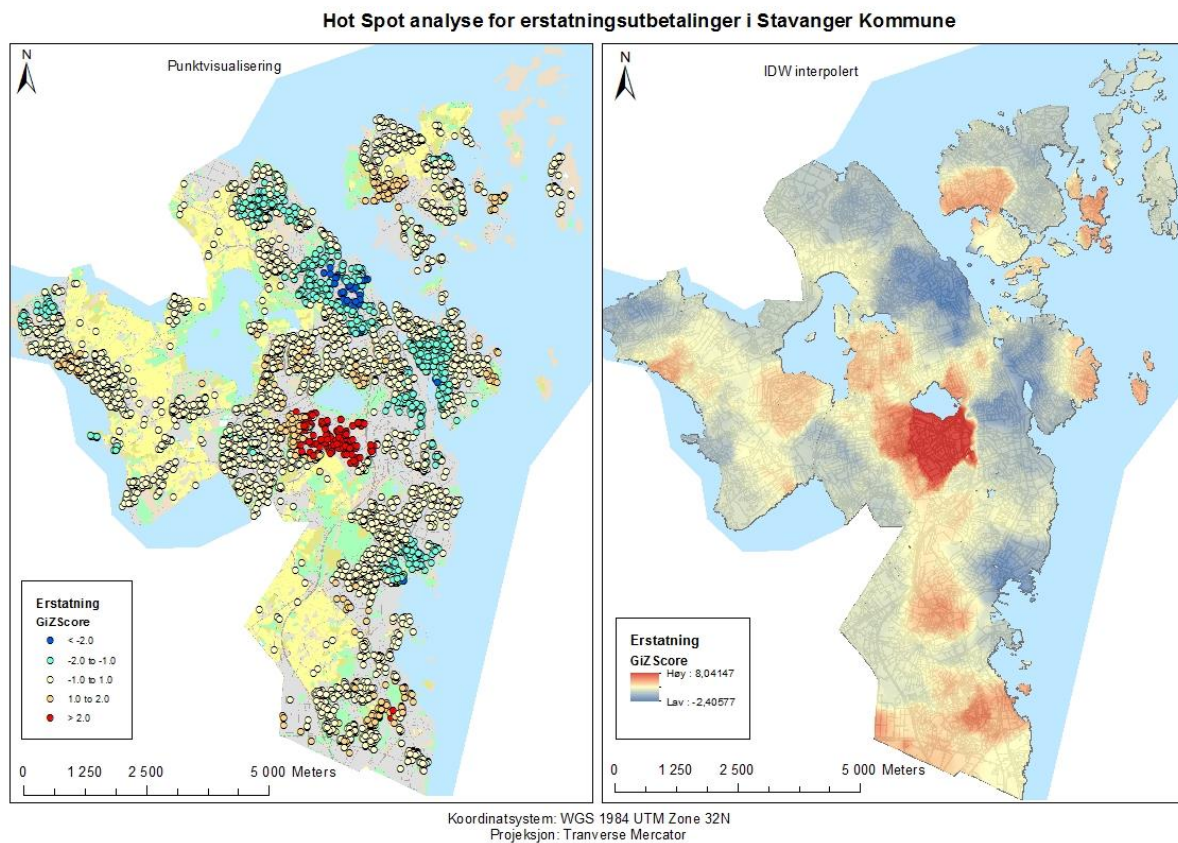
Figur 3.17: Arbeidsflyten fulgt for å gjøre datasettene klare for regresjon

4. Analyse og resultater

4.1 Hot Spot analyse for alle adressepunktene

Hot Spot analysene viser hvor det gitte analysefeltet er konsentrert geografisk. Kanskje vel så interessant som erstatningsutbetalingene, er også intensiteten av antall hendelser. Resultatene blir først presentert hvor alle adressepunktene er tatt med i figur 4.1 og 4.2, mens resultater der kun selekterte adressepunkter presenteres i figur 4.3 og 4.4.

4.1.1 Antall hendelser / erstatning

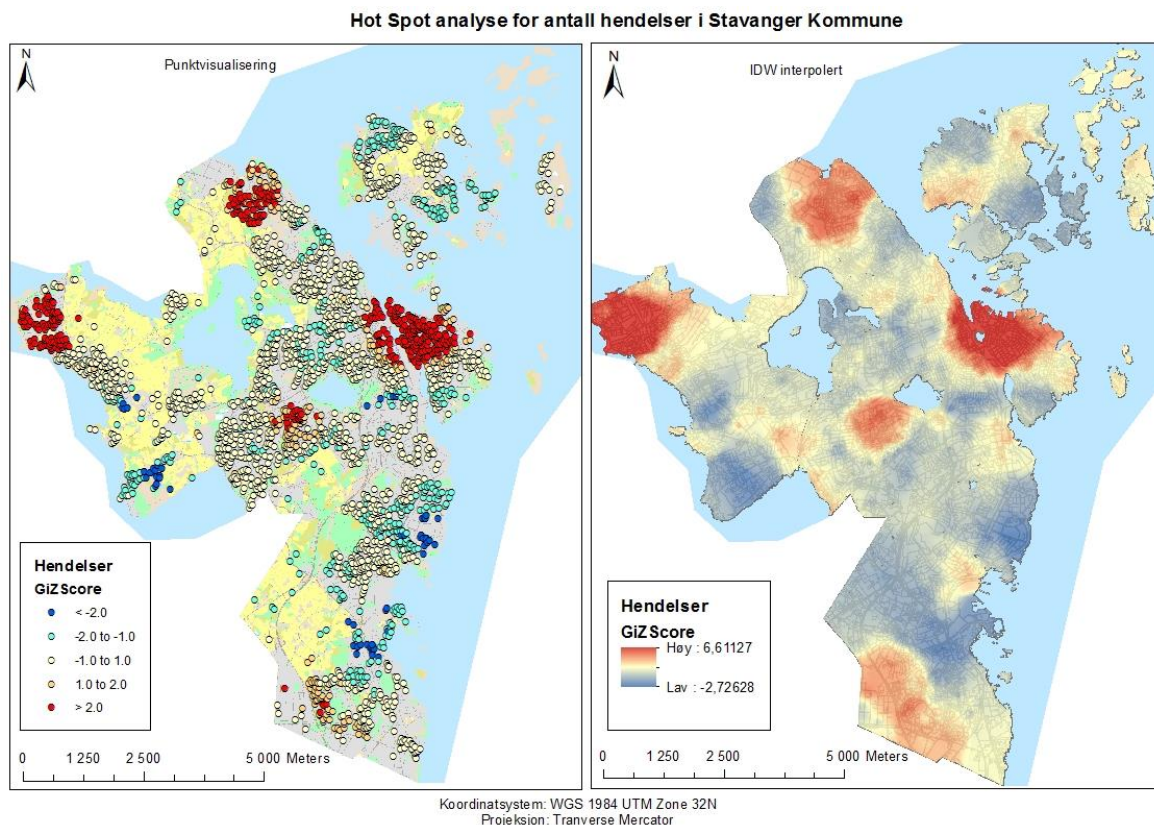


Figur 4.1: Hot Spot analyse av erstatningsutbetalinger i Stavanger Kommune hvor alle adressepunkter er tatt med. Visualisert til venstre som punkter og interpolert til høyre. (Statens Kartverk, 2015)

I figur 4.1 er det to visualiseringsmåter. Merk at det kun er punktvisualiseringen som kan tolkes som de 'ekte' resultatene, da interpoleringen også anslår verdier hvor det ikke finnes

adressepunkter. Dette er ikke en nøyaktig fremstilling og er kun ment som et alternativ for å være mer visualiseringsvennlig for kartleser. Som figur 4.1 viser, er det en ganske klar konsentrasjon av høye erstatningsutbetalinger sør for Mosvannet (Tjensvoll/Ullandhaug), symbolisert med røde punkter. Alle adressepunktene er tatt med i denne analysen, uavhengig av årsak, kilde og installasjon som er det egentlige analyseformålet. Grunnen til dette er at det vil være interessant å se om det finnes et mønster, også for det totale datasettet. I tillegg vil det være interessant å se hvorvidt det er noen likheter eller klare forskjeller med det 'egentlige' analyse materialet.

I tillegg til erstatningsutbetalingene er det også interessant å se hvordan hendelsene er konsentrert geografisk.

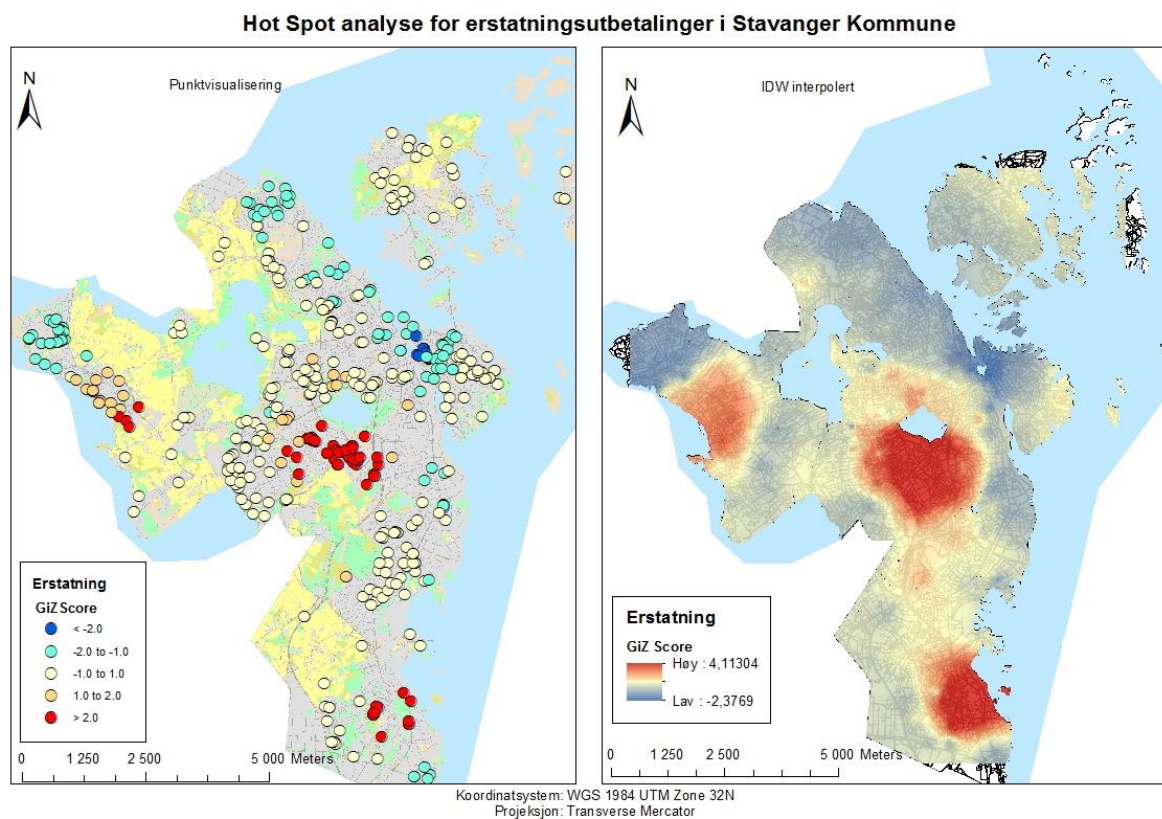


Figur 4.2: Hot Spot analyse over antall hendelser i Stavanger Kommune hvor alle adressepunkter er tatt med. Visualisert som punkter til venstre og interpolert til høyre. (Statens Kartverk, 2015)

Som det vises i figuren, er det flere statistisk signifikante hotspots. Deriblant Stavanger Sentrum i øvre kant på høyresiden, Tasta i øvre kant på venstre side og Kvernevik ytterst til venstre. Dette skyldes den tette bebyggelsen som blir innebefattet i den lave XY toleransen valgt, samtidig som det er mange skadehendelser på lite areal her. Et av de potensielle

problemområdene som det ble snakket om innledningsvis med kommunen, Jåttå (sørøst for Mosvannet), viser seg faktisk å være et statistisk signifikant cold spot i kommunesammenheng. Det samme gjelder både for intensitet av hendelser og erstatningsutbetalingene.

For de to neste kartene, er antall observasjoner redusert betraktelig. Fra 6744 observasjoner, finnes det nå 497. Som omtalt i kapittel 3.1, er det sortert ut på variablene som var fokus for oppgaven.

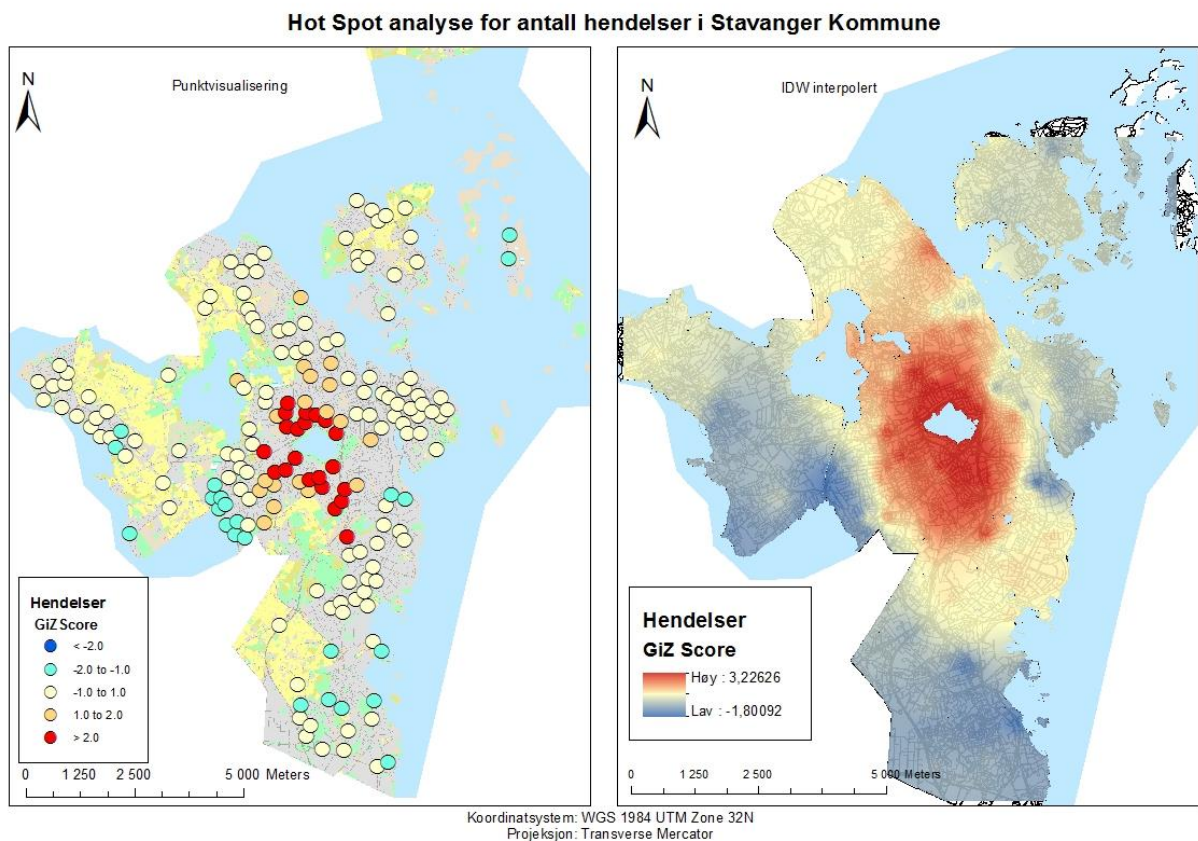


Figur 4.3: Hot Spot analyse over erstatningsutbetalinger i Stavanger Kommune hvor kun selekterte adressepunkter er tatt med. Visualisert som punkter til venstre og interpolert til høyre. (Statens Kartverk, 2015)

I motsetning til forrige hot spot analyse av erstatning, finnes det nå hot spots i området nær Gausel, nede til høyre i bilde samt mellom Sunde og Revheim til venstre i bilde. De samme adressepunktene var med i den generelle analysen også, men var ikke signifikant når den lokale summen ble målt proporsjonalt med summen av alle adressepunkter. Det største blikkfanget i kartet, er at det er det samme området sør for Mosvannet (midten av kartet) som utpeker seg hva angår høye erstatningsutbetalinger. Dette viser at også for det selekterte utvalget av variabler, er det samme området som utpeker seg som et problemområde. Dette

kan tyde på at det er skjedd store vannskader i dette området som har ført til høye erstatningsutbetalinger på mange nærliggende adresser. Samtidig kan det hende at disse adressepunktene også var utslagsgivende i analysen med alle adressepunktene.

Neste kart, viser at det identifisere problemområdet i tillegg til høye erstatningsutbetalinger, også har en høy intensitet av hendelser.



Figur 4.4: Hot Spot analyse over antall hendelser i Stavanger Kommune hvor kun selekterte adressepunkter er tatt med. Visualisert som punkter til venstre og interpolert til høyre. (Statens Kartverk, 2015)

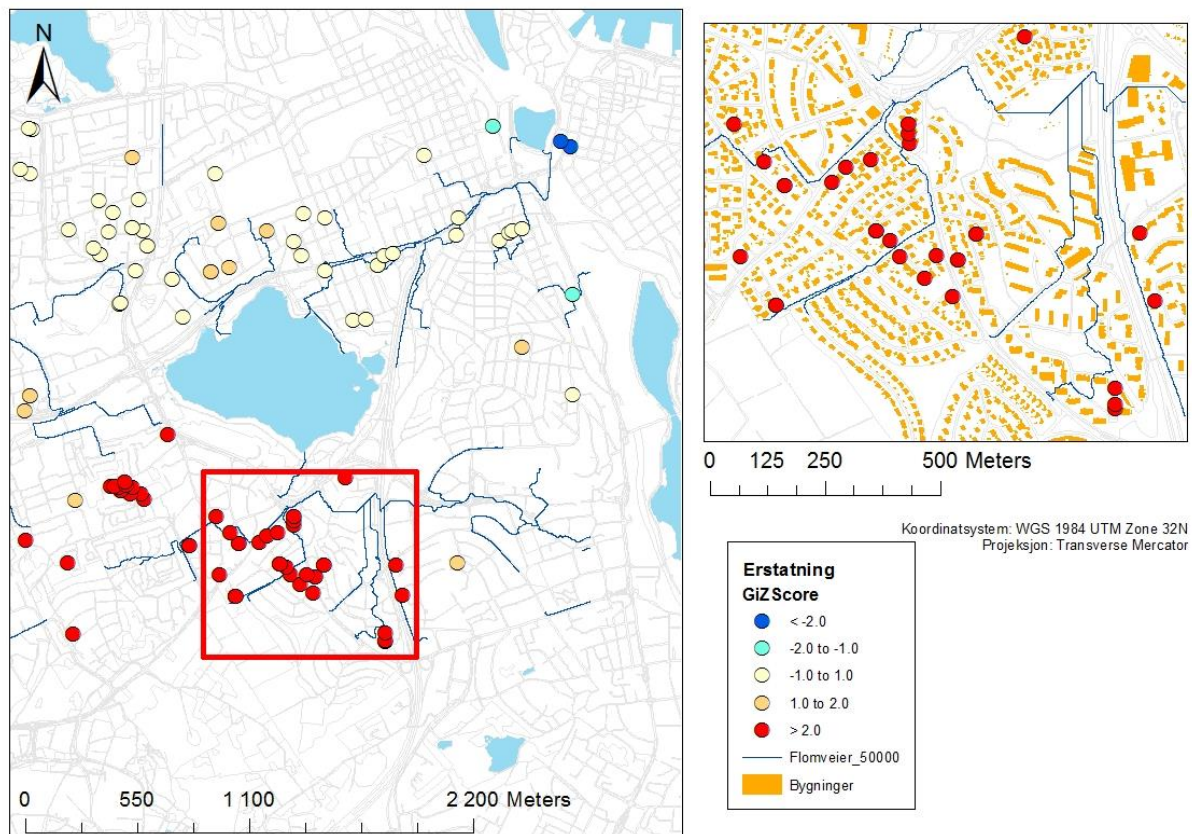
Igjen blir øyet umiddelbart dratt mot område rundt Mosvannet. Kartet er relativt ulikt analysen som tok for seg alle adressepunkter, hvor for eksempel Stavanger Sentrum, Tasta og Kvernevik ikke lenger er signifikante. Dette skyldes blant annet reduksjonen av observasjoner, samt at den økte XY toleransen til 75 meter ble utvidet.

I neste delkapittel vil det bli vist hvordan flomveiene er beregnet i det identifiserte problemområde som utmerket seg mest, området rundt Mosvannet. Hot Spot analysen for erstatninger med kun relevante årsaker blir inkludert i figurene for å vise hvordan disse

samsvarer. Det blir presentert beregnede flomveier med utgangspunkt i begge høydemodellene.

4.2 Flomveikart

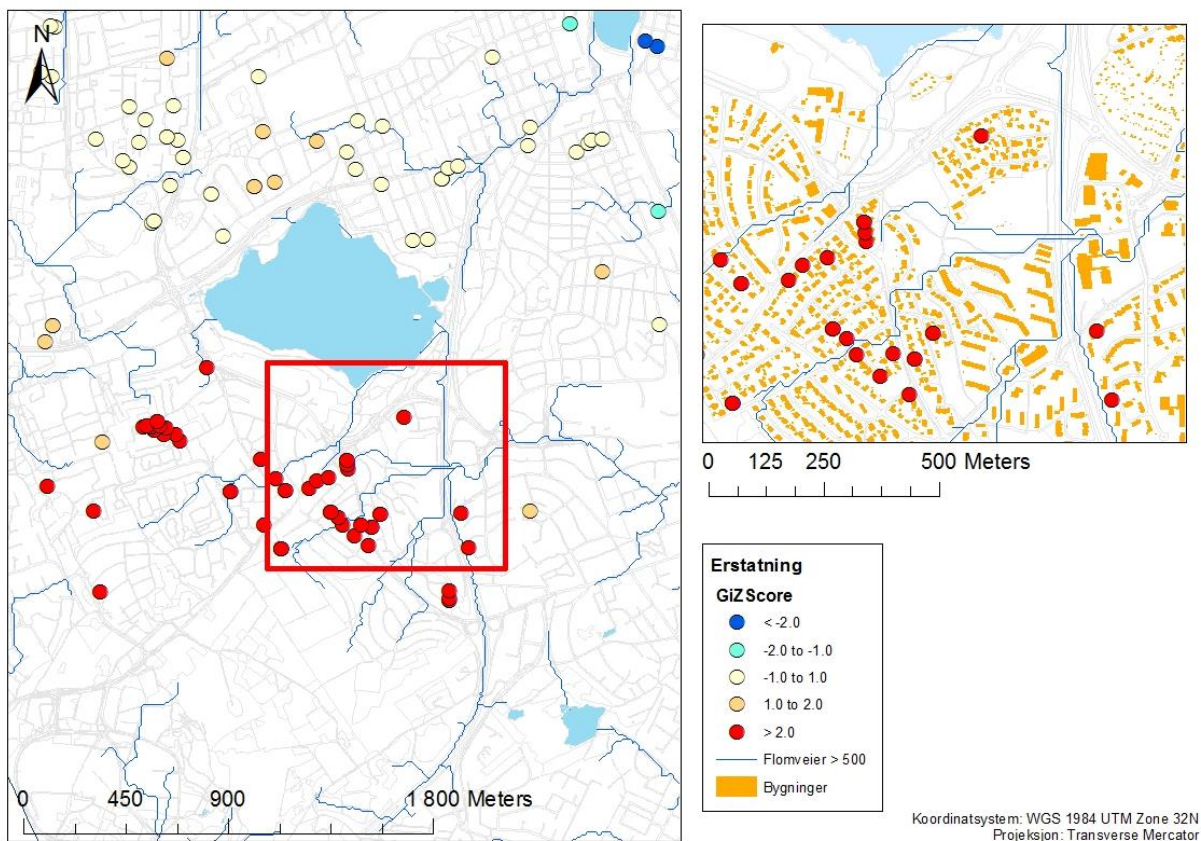
Nedenfor vises de beregnede flomveiene for høydemodellen med 1x1 meter oppløsning. Hot Spot analysen for erstatningsutbetalinger med kun relevante årsaker er inkludert.



Figur 4.5: Kart over beregnede flomveier for utsnittet av kommunen som tar utgangspunkt i høydemodellen med 1x1 meter oppløsning. Til høyre er det identifiserte problemområde fra Hot Spot analysen forstørret.

Terskelverdien er satt til 50 000, noe som indikerer antall celler som skal til for å bli klassifisert som en flomvei. Terskelen er brukerdefinert og er tatt med utgangspunkt i flomrapporten fra Meiforth (2013) som hadde lignende datamateriale til rådighet. Flomveiene viser et relativt godt samsvar med den signifikante hot spoten for erstatningsutbetalinger. Samtidig er det tydelig at det også går flomveier i områder hvor det ikke er utvist noen signifikans.

Samme hot spot analyse ble videre brukt i samsvar med flomveiene beregnet fra høydemodellen med 10m oppløsning.



Figur 4.6: Kart over beregnede flomveier for hele kommunen som tar utgangspunkt i høydemodellen med 10x10 meter oppløsning. Til høyre er deler av det identifiserte problemområde fra Hot Spot analysen forstørret (Statens Kartverk, 2015)

Terskelverdien er her satt til 500, altså 100 ganger lavere enn forrige. Dette med bakgrunn i at cellene er 10 ganger større i XY retning og antall kvadratmeter blir dermed likt. Flomveier som er under 500 blir utelatt fra dette kartet, først og fremst for at det ikke kan sannsynliggjøres at verdier under dette kan skape økonomiske skader på bygninger (Meiforth, 2013). Flomveikartet er relativt likt det forrige, men det ser ikke ut til å drenere direkte ned i Mosvannet. Sammen med akkumuleringen vist i kartene, er også terrengparametere og avløpsspesifikke variabler inkludert i regresjonene som i to påfølgende delkapitlene skal forklare hvorfor erstatningsutbetalingene varierer som de gjør blant de selekterte årsakene jf. kapittel 3.1

4.3 OLS resultater (hele kommunen)

Det ble kjørt to OLS-regresjoner i denne oppgaven. Først en som brukte verdiene ekstrahert fra høydemodellen med 10 x 10 meter oppløsning, deretter en annen som brukte verdiene ekstrahert fra høydemodellen med 1x1meter oppløsning. Rutenettstørrelsen var lik for begge regresjonene, 250x250 meter ble brukt som enhet for aggregering. Resultatene som

presenteres er hentet fra pdf-filen som blir skrevet ut i forbindelse med kjøring av OLS-verktøyet samt autokorrelasjonsverktøy.

Tabell 2: Resultat for hver enkelt av variablene i OLS-regresjon for hele kommunen

Variabel	Coefficient	StdError	t-Statistic	Probability[b]	Robust_SE	Robust_t	Robust_pr[b]	VIF[c]
Intercept	31,084697	20,526985	1,514333	0,131486	22,690371	1,369951	0,172202	-----
Aspekt	-0,001615	0,001154	-1,399966	0,163038	0,001121	-1,440746	0,151186	1,079526
Kurvatur	0,183049	0,220010	0,832004	0,406359	0,191148	0,957629	0,339362	1,065472
Skråning	-0,022447	0,037196	-0,603465	0,546865	0,034148	-0,657331	0,511700	1,039023
LogAcc	0,335274	0,066665	5,029245	0,000002*	0,074617	4,493236	0,000014*	1,066924
Anleggsår	-0,010728	0,010434	-1,028222	0,305043	0,011534	-0,930120	0,353386	1,698256
Dummy	0,214138	0,316073	0,677494	0,498850	0,325277	0,658324	0,511064	1,720334

Koeffisienten forteller både om styrken og forholdet mellom de ulike forklarende variablene og den avhengige variabelen (ESRI, 2013e) Når fortegnet er minus (-), er forholdet negativt. Variablene med et negativt forhold i denne modellen er skråning, anleggsår og aspekt. I teorien betyr dette at når erstatningsutbetalingene går ned, så vil skråningsverdiene gå opp.

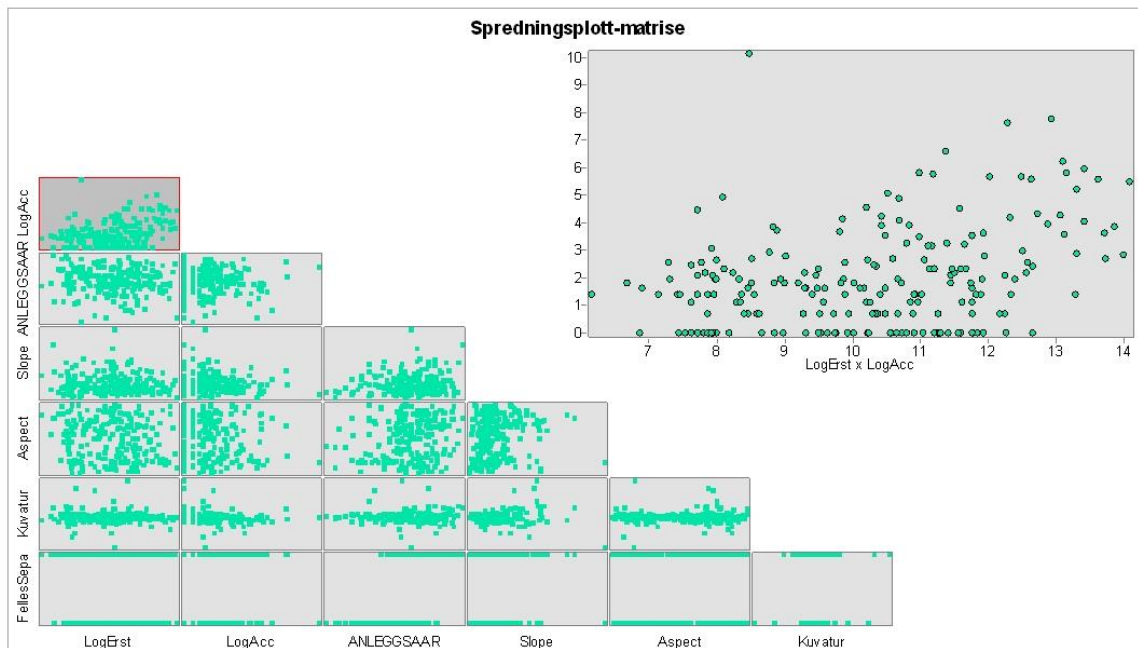
Aspektverdien har en koeffisient på -0.001. Det ble forsøkt flere transformasjoner for å skape en aspektvariabel som passet bedre til avhengig variabel, uten forbedring (se punkt to i kapittel 5.1.6)

Kurvaturen har en et positivt forhold og en koeffisient på 0.18. Dette indikerer at høyere erstatningsutbetalinger er forbundet med konvekse landformer som er de høyeste verdiene for kurvatur eller at lavere erstatningsutbetalinger er knyttet til lave kurvaturverdier som indikerer konkave landformer.

Skråningsvariabelen har en koeffisient på -0,02. Også her er koeffisienten så nær 0 at dette kan også tolkes som ingen forhold. Denne forklarende variabelen hjelper ikke modellen i og forklare variasjonen til den avhengige variabelen.

Den første forklarende variabelen med et tydelig positivt forhold er akkumuleringsvariabelen. Den logaritmisk transformerte variabelen har en koeffisient på 0,33 og et spredningsplott av

forholdet mellom variablene viser en trend som er lineær.



Figur 4.7: Spredningsplott-matrise over variablene som viser forholdet mellom akkumulerings- og erstatningsvariablene forstørret til høyre i figuren.

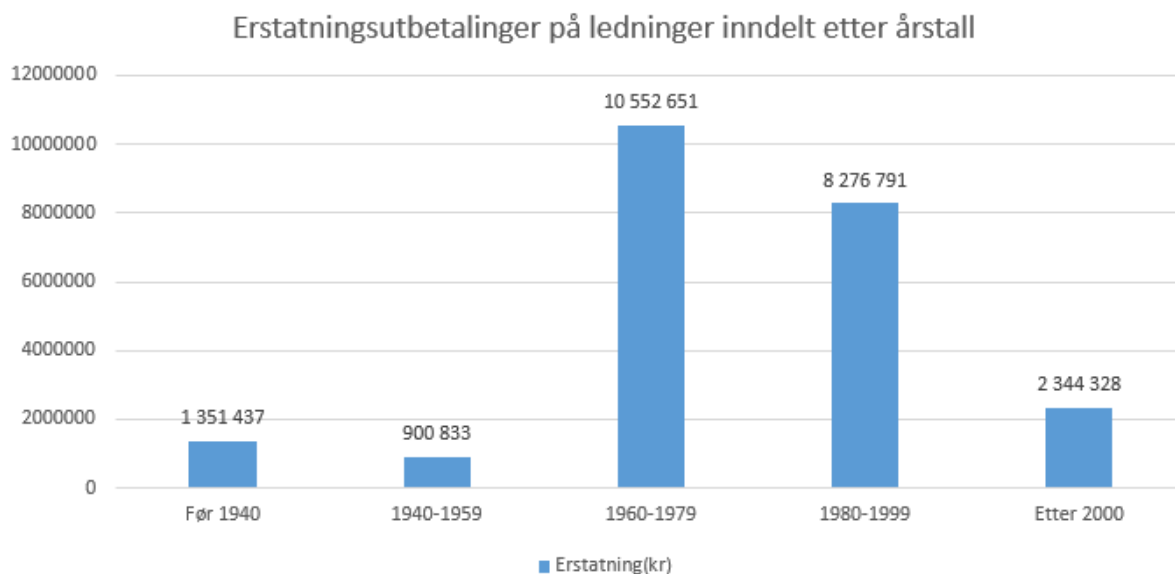
I tabell 2 har også akkumuleringsvariablen fått en asterisk(*) ved siden av henholdsvis probability[b] og robust_pr[b]. Når disse verdiene er veldig lave (mindre enn 0.05), er sjansen for at en koeffisient nær null, veldig lav. Det er derfor sannsynlig på et 95 prosent konfidensintervall at dette ikke er tilfeldig. Denne forklarende variabelen er altså assosiert med en statistisk signifikant koeffisient, og er viktig for regresjonsmodellen (ESRI, 2013e).

Variablene for anleggsår, er den gjennomsnittlige alderen på ledningene innenfor hver rute. Koeffisienten er -0.01, og det finnes ingen forhold til den avhengige variabel. Ettersom eldre ledninger var tiltenkt å ha en tilknytning til høye erstatningsverdier, var forventningen at dette forholdet skulle bli negativt. At det ikke finnes noe forhold mellom variablene betyr at den ikke hjelper modellen, men resultatet er allikevel interessant. I ettertid av regresjonen ble det laget en tabell og diagram for å manuelt inspisere variabelen isolert sett.

Tabell 3: Gjennomsnittlig alder på ledningene inndelt i fem kategorier. Antall ruter som representerer den respektive kategorien av gjennomsnittlig anleggsår.

Gjennomsnittlig anleggsår	Antall 250x250 ruter
Før 1940	3
1940-1959	16
1960-1979	77
1980-1999	98
Etter 2000	19

Den gjennomsnittlige alderen innenfor hver rute ble konvertert til en strengvariabel hvor det ble delt inn i fem kategorier. Av det totale rutenettet på 213 observasjoner, er disse også kategorisert etter hvilke år som representerte sine respektive ruter. Av denne tabellen ble det videre beregnet erstatningsbeløp innenfor hver kategori, noe som ga dette diagrammet.



Figur 4.8: Erstatningsutbetalinger på ledninger av gitte årskategorier.

Det er interessant å merke seg at etterkrigstidens ledninger markerer seg med høye erstatningsbeløp. Det skal riktignok registreres at kategorien 1960-1979 representerer flere ruter enn ledninger frem til 1959 og etter 2000. Men kategorien 1980-1999 har flest ruter og dessuten betydelig lavere erstatningsbeløp. Dette funnet forsterker hva Eigersund Kommune

(2014) påpekte om dårlig kvalitet blant etterkrigstidens ledninger, mens ledningenes kvalitet markant bedret seg ved slutten av 1970-tallet (se kapittel 2.6).

Den siste og avsluttende variabelen som ble tatt med i denne første OLS-regresjonen var dummy-variabelen som tar for seg felles kontra separate avløpssystem. Koeffisienten var på 0,21 og er et positivt forhold til den avhengige variabelen. Det vil si at når typetallet i ruten er separate system (1) er det forbundet med økte erstatningsutbetalinger.

Alle variablene har en VIF-verdi rundt 1, noe som forteller at det ikke finnes variabler med redundans som forteller den samme «historien» om den avhengige variabelen (ESRI, 2013e). Neste steg er å vurdere modellen. Både Multiple R-squared og R^2 er mål på hvor god modellen er, med et mulig verdiintervall mellom 0.0 og 1.0.

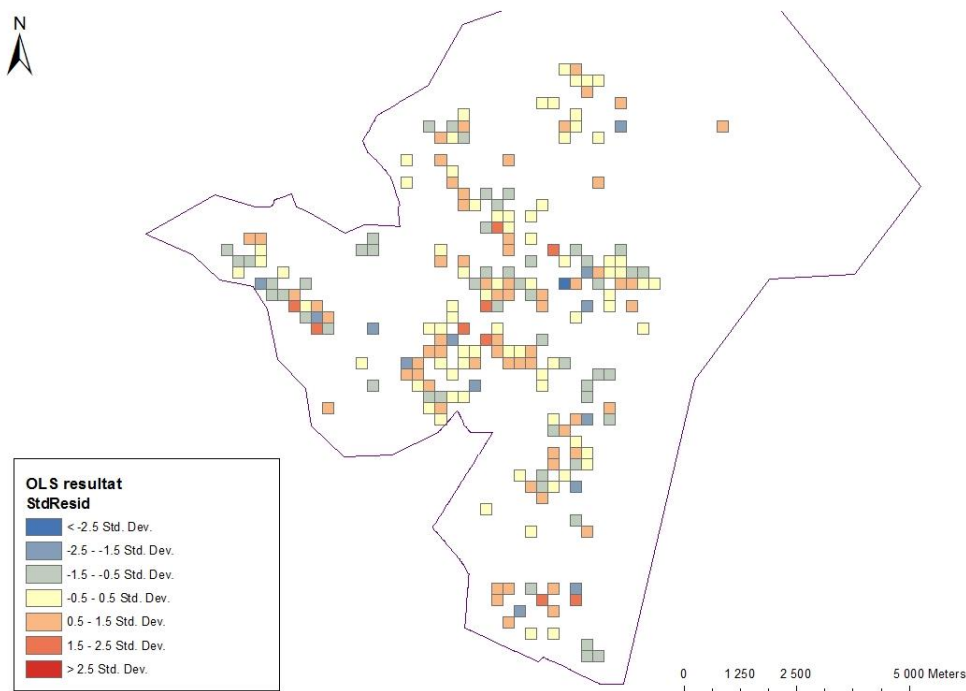
Tabell 4: OLS resultater for hele kommunen.

Observasjoner	213	Akaike's Information Criterion (AICc) [d]:	838,688204
Multiple R-Squared [d]:	0,129996	Adjusted R-Squared [d]:	0,104656
Joint F-Statistic [e]:	5,130092	Prob(>F), (6,206) degrees of freedom:	0,000062*
Joint Wald Statistic [e]:	26,798147	Prob(>chi-squared), (6) degrees of freedom:	0,000158*
Koenker (BP) Statistic [f]:	8,567802	Prob(>chi-squared), (6) degrees of freedom:	0,199383
Jarque-Bera Statistic [g]:	5,614204	Prob(>chi-squared), (2) degrees of freedom:	0,060380

Multiple R-squared viser altså en forklaringsevne på 0,12 mens R^2 viser en nedjustert forklaringsevne på 0,10. Det er sistnevnte som er mest interessant, og det betyr at de forklarende variablene forklarer ti prosent av variasjonen i den avhengige variabelen.

Som man ser av tabellen ovenfor, er det en asterisk ved både F-statistikk og Wald-statistikk. Begge disse er generelle mål på modellens statistiske signifikans. F-statistikken er kun til å stole på dersom Koenker (BP) statistikken *ikke* er statistisk signifikant. Som det går frem av tabellen er den ikke det i denne modellen. Nullhypotesen for F- og Wald-statistikken er at de forklarende variablene *ikke* er effektive. For et 95 prosent konfidensintervall, tyder en p-verdi på mindre enn 0.05 at modellen er statistisk signifikant og at nullhypotesen kan forkastes (ESRI 2013e; ESRI 2013g). Som tabellen viser har begge målene et lavere tall enn dette, og dette tyder på en statistisk signifikant modell. Dersom det er slik at forholdene mellom avhengig og forklarende variabler endrer seg i ulike deler av kartet, vil det være en asterisk ved siden av Koenker-statistikken. Dette er ikke tilfelle i resultatene fra modellen. Videre er det heller ingen asterisk ved siden av Jarque-Bera, noe som vitner om at residualene er tilnærmet normalfordelte og dermed oppfyller kravet om pålitelighet i modellen (ESRI, 2013e).

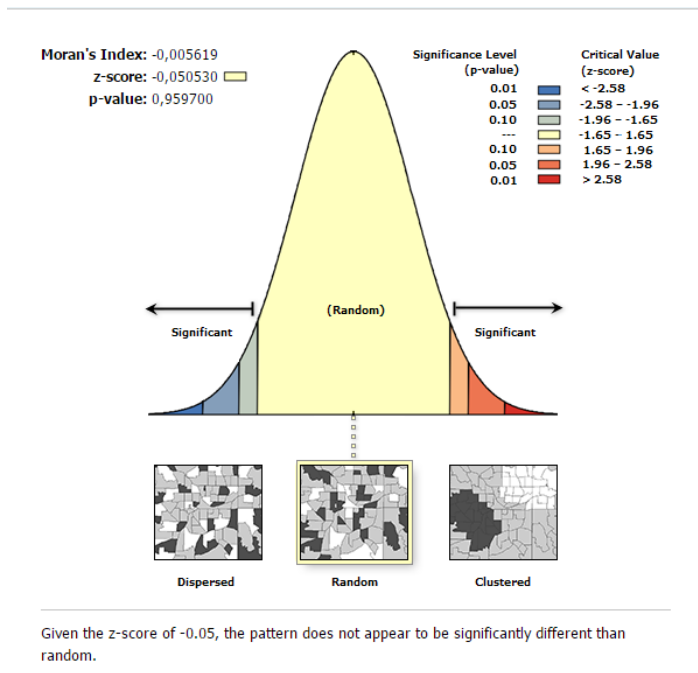
Når man kjører OLS-verktøyet, får man en diagnostikk fil som kan hjelpe i å kalibrere modellen. Det som i utgangspunktet er det interessante i OLS resultatene er R^2 verdien. Men før en R^2 -verdi kan godkjennes, må blant annet residualene undersøkes nærmere. De skal ikke utvise statistisk signifikant romlig autokorrelasjon, residualplottet skal ikke utvise noe mønster og residualene skal ideelt sett være normalfordelte (ESRI, 2013g). Hvis fordelingen avviker veldig fra normalfordelingen kan man ha en liten pålitelig modell. Et romlig autokorrelasjonsverktøy kan sjekke om residualene utviser signifikante positiv autokorrelasjon, før dette verktøyet ble kjørt ble residualene manuelt inspisert i et kart.



Figur 4.9: Residualkart som viser over- og underprediksjoner, røde ruter er overprediksjoner og blå ruter er underprediksjoner.

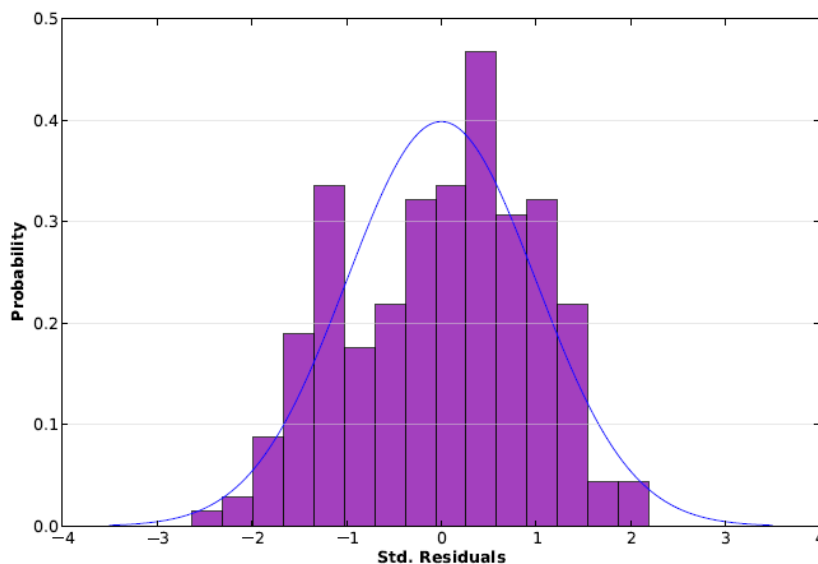
I utgangspunktet henger et rutenett sammen, men som nevnt er alle ruter uten adresser fjernet. Hvis man ser på Hot Spot kartet i figur 4.1, ser man at dette rutenettet sammenfaller med hvor disse punktene ligger. De røde rutene viser at de reelle verdiene for erstatningsutbetalinger er større enn modellen har predikert at det skulle vært. Mens de blå områdene viser hvor de reelle verdiene er lavere enn predikert av modellen.

I resultatvinduet av kjøringen av OLS-regresjonen får man en påminnelse og advarsel om at det bør kjøres et romlig autokorrelasjonsverktøy på residualene, for å sjekke at disse oppfyller kravet om ikke statistisk signifikante over- og underprediksjoner. Dersom residualene er statistisk signifikante er det tegn på at modellen mangler nøkkelvariabler, og er lite pålitelig. Nedenfor vises resultatet av det romlige autokorrelasjonsverktøyet på residualene:



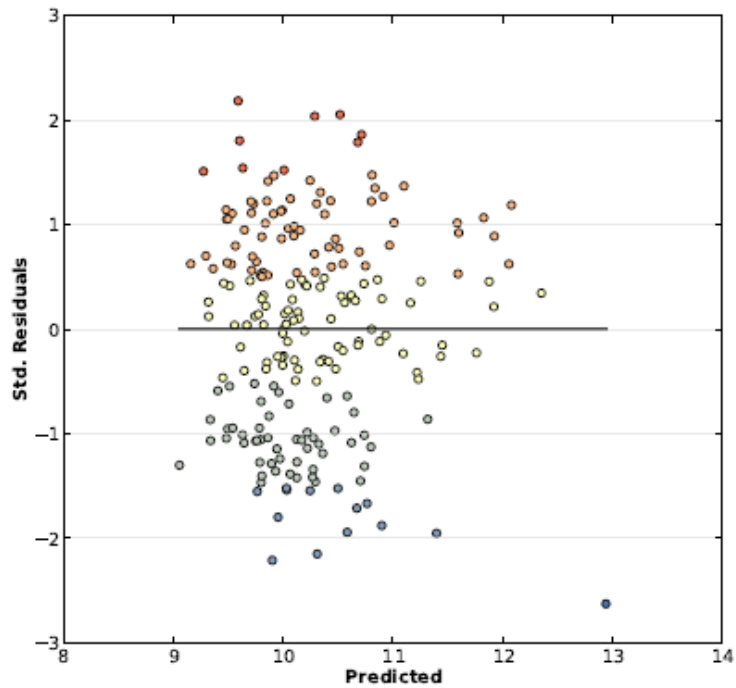
Figur 4.10: Resultat etter romlig autokorrelasjonsverktøy. Indikerer at det ikke finnes statistisk signifikante mønster av over- eller underprediksjoner

Figur 4.10 indikerer at det ikke finnes signifikante mønster av residualene, og at fordelingen synes å være tilfeldig. Videre skal residualene ideelt sett følge den blå kurven:



Figur 4.11: Histogram som viser fordelingen av residualer

Som det går frem av figuren følger fordelingen den blå normalkurven relativt bra. Den siste indikatoren på skjevhet i modellen som skrives ut i pdf filen fra OLS-verktøyet er spredningsplottet over residualene som ikke bør utvise noen form for mønster.



Figur 4.12: Spredningsplott av residualer. Ingen indikasjon på mønster.

Residualplottet viser over- og underprediksjonene i relasjon til den predikerte avhengige variabelens verdier. For en pålitelig modell skal residualplottet ikke inneha noe mønster eller struktur. Som det går frem av figuren finnes det ingen indikasjoner på mønster.

4.4 OLS resultater (utsnitt)

I dette delkapittelet presenteres resultatene fra OLS-regresjonen foretatt på det identifiserte problemområde.

Tabell 5: Resultat for hver enkelt av variablene i OLS-regresjon for utsnittet av kommunen.

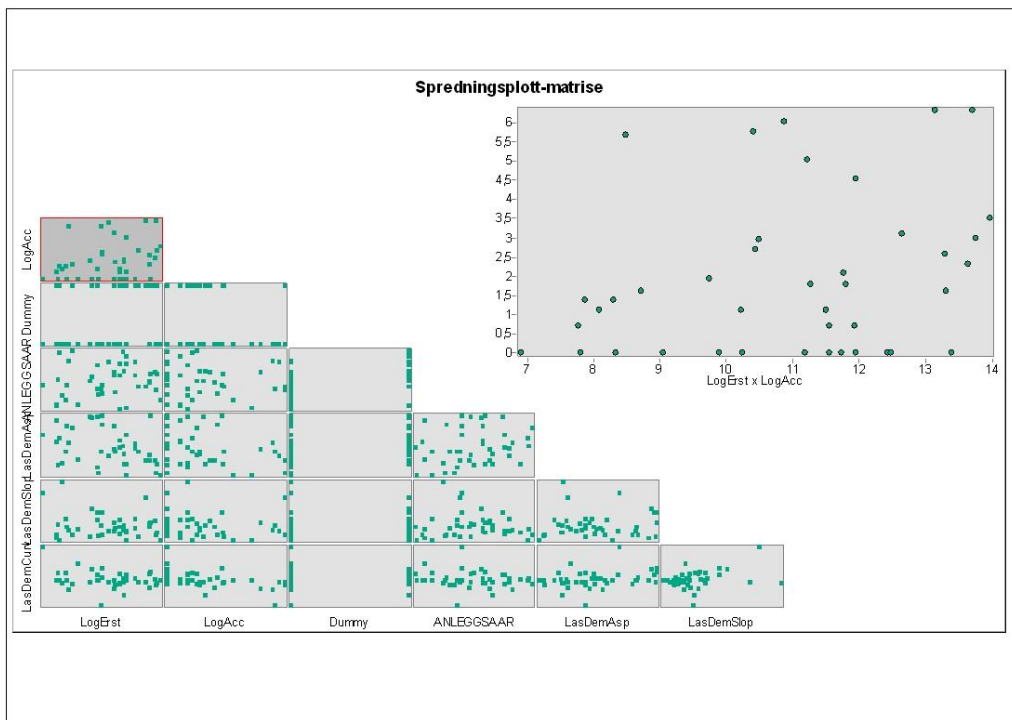
Variabel	Coefficient	StdError	t-Statistic	Probability[b]	Robust_SE	Robust_t	Robust_pr[b]	VIF[c]
Intercept	25,407037	30,786399	0,825268	0,414655	41,196571	0,616727	0,541295	-----
Kurvatur	0,013119	0,071657	0,183079	0,855765	0,057841	0,226809	0,821857	1,649156
Aspect	-0,005678	0,002656	-2,137935	0,039384*	0,002181	-2,603063	0,013337*	1,217877
Anleggsår	-0,007557	0,015789	-0,478642	0,635087	0,021093	-0,358282	0,722225	1,586599
LogAcc	0,326860	0,162703	2,008942	0,052089	0,112398	2,908062	0,006195*	1,553275
Skråning	-0,081399	0,093822	-0,867595	0,391363	0,077407	-1,051578	0,300001	1,165290
Dummy	2,027240	0,652621	3,106305	0,003685*	0,729185	2,780144	0,008588*	1,611042

Kurvaturen hadde i forrige regresjonsanalyse et positivt forhold til den avhengige variabelen. Det har den også her. Koeffisienten er positiv og har en verdi på 0,01. Men denne verdien er veldig nær null, så det vil nok være mer nærliggende og anse dette som ingen forhold.

Aspektvariabelen har samme forhold som før. Den har fortsatt et negativt fortegn, og en koeffisient nær null, men er i denne regresjonen statistisk signifikant uttrykt ved lav p-verdi.

Anleggsår har i denne regresjonen et negativt forhold. Også her er det imidlertid svært nær null og derfor ingen forhold.

Akkumuleringsvariabelen, logacc, har også i denne regresjonen et positivt forhold til den avhengige variabel. Den er nesten like sterk som i forrige regresjon, og fortsatt statistisk signifikant markert med asterisk i robust sannsynlighet. Et lineært forhold kan sees forstørret til høyre i figuren.



Figur 4.13: Spredningsplott-matrise over variablene som viser forholdet mellom akkumulerings- og erstatningsvariabelen i forstørret form til høyre.

Skråningsvariabelen har et negativt forhold til den avhengige variabelen, og en koeffisient på -0,08.

Dummy-variabelen har gått opp til en koeffisient på 2. En markant endring fra forrige regresjon, hvor den også hadde positivt forhold men ikke like sterkt. Her er også dummy-variabelen markert som statistisk signifikant ved 95 prosent konfidensintervall.

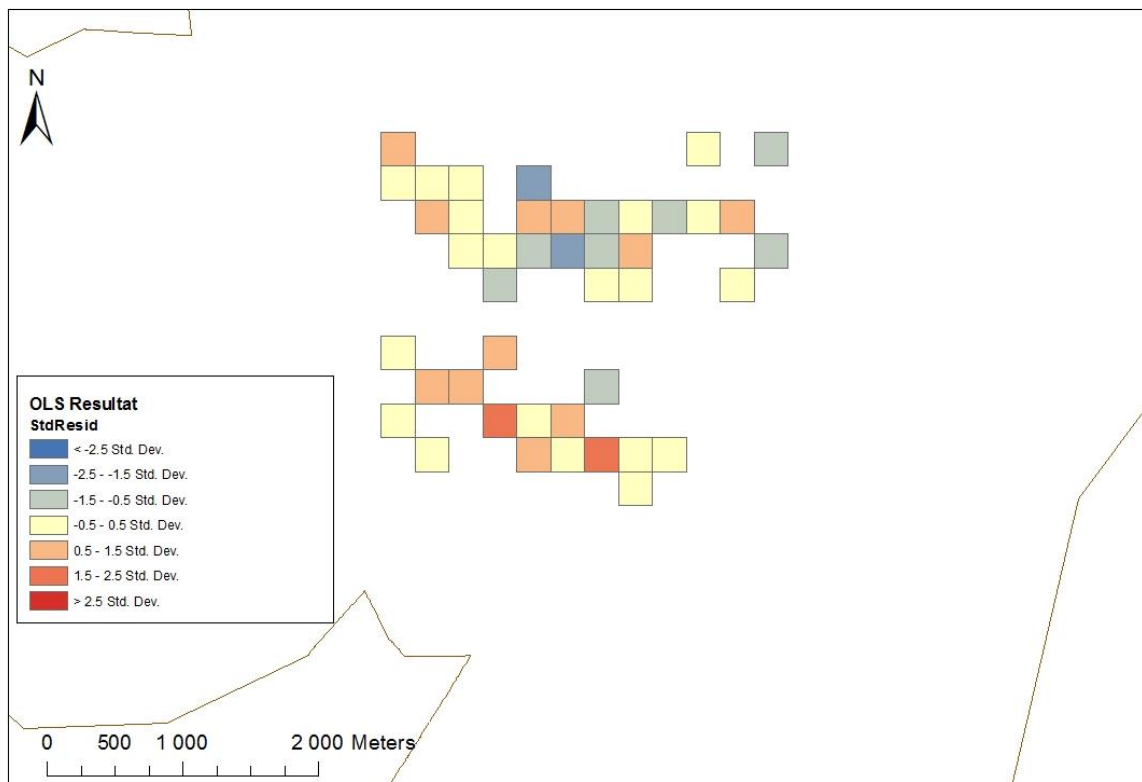
VIF-verdiene er også innenfor denne regresjonen godt innenfor den anbefalte grensen på 7.5, noe som betyr at ingen variabler utviser redundans. Resultatene totalt for modellen følger nedenfor:

Tabell 6: OLS resultater for utsnittet av kommunen

Observasjoner	43	Akaike's Information Criterion (AICc) [d]:	179,327728
Multiple R-Squared [d]:	0,380985	Adjusted R-Squared [d]:	0,277816
Joint F-Statistic [e]:	3,692826	Prob(>F), (6,36) degrees of freedom:	0,005819*
Joint Wald Statistic [e]:	51,052835	Prob(>chi-squared), (6) degrees of freedom:	0,000000*
Koenker (BP) Statistic [f]:	2,452477	Prob(>chi-squared), (6) degrees of freedom:	0,873748
Jarque-Bera Statistic [g]:	1,413250	Prob(>chi-squared), (2) degrees of freedom:	0,493306

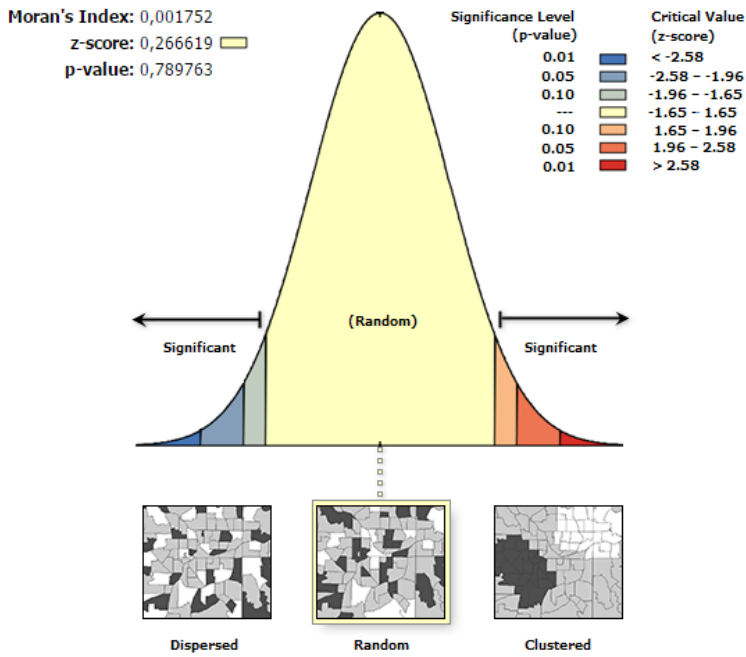
Multiple R-squared viser en forklaringssevne på 0,38, mens R^2 ligger på 0,27. Det betyr at de forklarende variablene forklarer ca. 27 prosent av variasjonen i den avhengige variabelen. En asterisk ved siden av F- og Wald-statistikken forteller at modellen som helhet ikke har koeffisienter som kombinert er 0, noe som indikerer at nullhypotesen kan forkastes. Videre er Koenker ikke statistisk signifikant, og dermed kan F- og Wald stoles på.

En manuell sjekk av residualene ble også foretatt her, nedenfor vises kart med over- og underprediksjoner.



Figur 4.14: Residualkart over over- og underprediksjonene i det identifiserte problemområde.

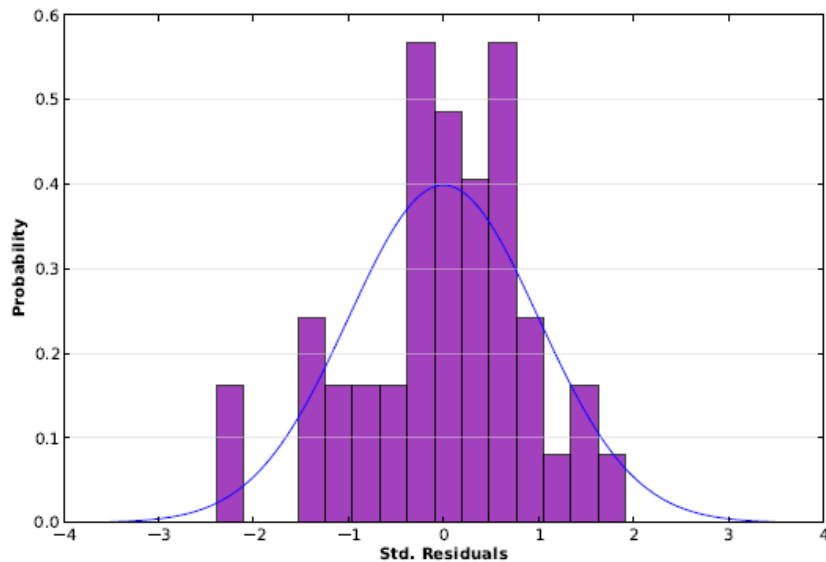
Den samme fremgangsmåten med å kjøre et romlig autokorrelasjonsverktøy på residualene ble fulgt for å godkjenne at residualene ikke er konsentrert geografisk. OLS-modellen for utsnittet har også de ønskede resultatene på romlig autokorrelasjon. Ingen statistisk signifikante konsentrasjoner av over- eller underprediksjoner.



Given the z-score of 0.27, the pattern does not appear to be significantly different than random.

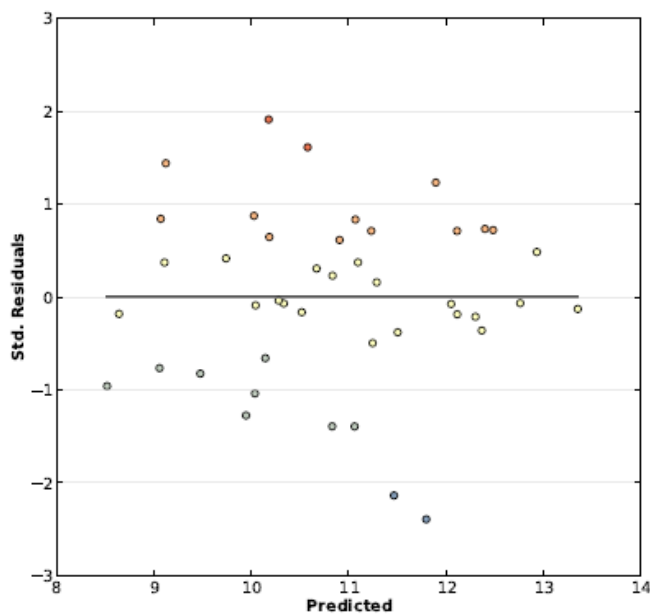
Figur 4.15: Resultat av romlig autokorrelasjonsverktøy. Mønsteret indikerer ikke statistisk signifikante over- eller underprediksjoner.

Figuren indikerer at mønsteret ikke synes å være signifikant, og blir klassifisert som tilfeldig.



Figur 4.16: Histogram som viser fordelingen over residualer.

Som nevnt skal fordelingen av residualene ideelt være normalfordelt og følge den blå linjen. Det er mer avvik fra normalfordelingen for denne regresjonen enn den forrige. Diagnostikkrapporten til verktøyet påpeker at såfremt fordelingen ikke er *veldig* forskjellig fra normalfordelingen, kan det indikere en skjev modell. Hvis denne skjevheten er statistisk signifikant, vil det vises ved en asterisk(*) på Jarque-Bera verdien. Dette finnes det imidlertid ikke i tabell 6 og etter mitt skjønn er ikke fordelingen veldig forskjellig fra normalfordeling heller. En ikke signifikant Jarque-Bera, tilnærmet normalfordeling, samt residualplott og autokorrelasjon som ikke utviser noen mønster gjorde at modellen ble beholdt til tross for noe avvik fra den blå kurven.



Figur 4.17: Spredningsplott av residualer. Ingen indikasjon på mønster.

Spredningsplottet har ingen indikasjon av mønster, og understøtter vurderingen av at regresjonen er pålitelig.

5. Diskusjon

5.1 Evaluering av datagrunnlag og analyser

Kvaliteten på resultatene av analysene er selvsagt betinget kvaliteten på inngangsdataene og vurderinger gjort i metodikken. I dette kapittelet blir datagrunnlaget, metodiske valg og analysene diskutert.

5.1.1 Evaluering av skadedata

Årsakskoding

Få av forsikringsselskapene har brukt denne kolonnen noe særlig. Nesten halvparten av observasjonene jeg hadde til rådighet har koden 9, som betyr ukjent årsak (se kodebok i vedlegg). Årsakskodingen ble ikke tatt i bruk før 2007/2008, det vil si at alle vannskader før denne tid ikke har noen svar på hvordan eller hvorfor skaden skjedde. Det er også et problem at noen av forsikringsselskapene ikke bruker underkategorier, men kun bruker denne 'ukjent-kodingen' når en vannskade oppstår. Det vil si at årsaker relatert til overflateavrenning blir likestilt i kodingen med en vaskemaskin som har hatt lekkasje. De siste årene har kodingen blitt langt mer sofistikerte, men det finnes fortsatt rom for forbedringer. Blant annet har pilotkommunene ønsket seg ytterligere underkategorisering hvor det går frem av kodingen hvorvidt overflateavrenning er skjedd fra offentlig eller privat areal. Trondheim kommune har påpekt at årsakskodingen brukt i denne oppgaven (I, G, J) ikke er godt nok relatert til nedbørsflom, men kan også være relatert til andre årsaker (Brevik m.fl, 2015). Den inkonsekvente kodingen av årsaksdataene har ført til at potensielt relevante saker, ikke er blitt tatt med i mine analyser. Samtidig skal det noteres at forsikringsselskapenes koding ikke hadde pilotprosjektets anliggende i tankene når registrering av skadehendelser startet.

Erstatningssum

Flere av skadehendelsene har null i erstatningssum. Årsaken til dette er uklart, men det kan skyldes at saken er gått til regress. Når det eventuelle regresskravet blir godkjent, nullstiller forsikringsselskapene beløpet. Noen veldig få av skadehendelsene var også registrert med

negative beløp. Det var usikkert hvorvidt det var registrert feil fortegn i datainnføringen eller om også dette skyldtes regressaker. Samtidig har forsikringsselskapene levert forskjellig når det gjelder kroner og millioner kroner (Brevik m.fl, 2015). Etter manuell inspeksjon av dataene kunne det ikke identifiseres noen tilfeller av dette i mine data.

Det er kun de økonomiske tapene som er dekket av forsikringsselskapene som er med i datasettet. Det vil si at avviste krav, egenandeler og indirekte kostnader ikke er innhentet. En studie av Nilsdal (1985) antyder at disse ikke-registrerte kostnadene kan være meget betydelige (NVE 1997:10; Nilsdal 1985). Samtidig er det ingen indikasjon på om kjellerne som er oversvømt er innredet eller ikke, og dette er sannsynligvis en betydningsfull variabel for erstatningssummen.

Generelle utfordringer:

Oppsettet for føring og koding av dataene er på ingen måte optimalt. Det eksisterer ingen pålagt form for hvordan dette skal gjøres. Samtidig er dataene i utgangspunktet ikke laget for pilotprosjektets formål da registrering av hendelsene startet lenge før prosjektets oppstart. Men med tettere samarbeid mellom kommune og forsikringsselskap virker det lovende for fremtidig bruk av skadedataene, såfremt samarbeidet fortsetter.

5.1.2 Evaluering av Hot Spot analysene

Det som potensielt er 'farlig' med visualisering av Hot Spot kartene er at de bærer med seg et inntrykk av autoritet og objektivitet. Men det er viktig å presisere at mange av valgene gjort i denne analysen er brukerbestemte, blant annet det viktige avstandsvalget for nabolagene. Dersom man endrer på dette tallet, kan resultatene drastisk forandre seg og gi et helt forskjellig bilde av 'virkeligheten' enn det som er presentert i denne oppgaven. Som kartograf er det derfor nødvendig å påpeke min egen rolle i produksjonen av disse kartene, for å understreke at dette er resultat av mine faglige vurderinger og bør ikke oppfattes som fullkommen objektivitet. Symbologien i kartet er en automatisert prosess, og det anbefales forsiktighet i å endre på dette, da både mistolking og uriktig fremstilling kan forekomme dersom man definerer symbologien selv (ESRI, 2007). Dette er naturlig nok tatt til følge, og symbologien har stått uforandret etter resultatet av den automatiserte symbologien.

Videre kan det oppstå brukerpåførte feilkilder i tilknytning til den valgte XY toleransen for hot spot analysen hvor antall hendelser vises. I bruk av aggregeringsverktøyet bruker man en

selvvalgt avstand som skal definere hvilke adressepunkter som skal slås sammen, for deretter å bli brukt som analysefelt. Det anbefales å bruke en lav XY toleranse ifølge ESRI (2013d), men dette oppleves som et generelt og vagt råd. Et verktøy som beregner gjennomsnittlig avstand mellom nærmeste nabo ble kjørt på adressepunktene for å gi et estimat av hvilken toleranse som var fornuftig, og 15 meter ble derfor valgt. For hot spot analysen av hendelser hvor kun de relevante årsakene var selektert ut, ble XY toleransen som sagt hevet til 75 meter for å sørge for at punktene skulle ha flere naboer. I utgangspunktet skal alle punkter ha minst én nabo (ESRI, 2013b), men dette var etter mitt skjønn, umulig å overholde, da noen få av adressepunktene var langt fra naboer. Hvis toleransen skulle blitt ytterligere utvidet til også å inkludere disse (opptil 900 meter) ville dette overtrådt retningslinjene om å velge en lav XY toleranse. Valget falt derfor på en slags kompromissløsning mellom de to. Hensikten med analysene var å identifisere problemområdene, men det registreres at ikke alle adressepunkt hadde en nabo og det må noteres som en mulig feilkilde.

For hot spot analysene med utgangspunkt i erstatningsutbetalinger var avstandene relativt enkle å finne ut av, da autokorrelasjonsverktøyet identifiserte det optimale nabolaget for visualiseringen. I motsetning til OLS-regresjonene, ble ikke erstatningsbeløp på 0 fjernet fra disse analysene. Først og fremst fordi antall observasjoner var langt større, og en rettferdiggjøring av at alle 0-verdier var feildata eller nullstilt regress ikke lå tilstede etter mitt skjønn. Ettersom de fortsatt representerer registrerte skadehendelser ble de inkludert i analysene for begge hot spot analysene som brukte erstatningsutbetalinger. Videre var hovedformålet med analysen å identifisere problemområdene, som fortsatt var mulig til tross for dette.

Resultatene fant at det var relativt like hot spots, uavhengig av hvilke seleksjoner som ble utført. Spesielt gjaldt dette for erstatningsutbetalingene, hvor området sør for Mosvannet ble et kraftig hot spot i begge tilfellene. Dette er noe overraskende, da potensielle problemområder ble diskutert med Stavanger Kommune i startfasen av oppgaven og dette område ikke ble nevnt. Det er allikevel nevnt i media at dette området har vært utsatt for skadehendelser, for eksempel en artikkel fra Stavanger Aftenblad skriver at 35 kjellere måtte pumpes i dette området som følge av flom og mangel på kapasitet i avløpssystemet (Stavanger Aftenblad, 2012). Videre går det frem av artikkelen at innredede kjellere med 10-15 cm vann, vil kreve mye jobb å rette i. Dessuten mange adressepunkter i det identifiserte problemområde skadedatoer med samme årstall som artikkelens publiseringsdato, noe som kan tyde på at hendelsene det refereres til i Aftenbladet har vært kostbare.

5.1.3 Evaluering av høydemodeller

I utgangspunktet er det ikke noen umiddelbare feilkilder i den LiDAR-deriverte høydemodellen. Punktettheten er høy (5,15 m²) og den gjennomsnittlige punktavstanden er lav (0.44m). Ifølge flere kilder (Su og Bork 2006; Liu 2008, Lloyd og Atkinson 2002) er IDW en god interpoleringsmetode å benytte når punktavstanden er såpass lav. Valg av oppløsning på 1meter er etter mitt skjønn rettfærdiggjort av den høye punktettheten, og kunne muligens vært redusert helt ned til 0.5m uten å redusere datakvaliteten. Dette ville imidlertid ha økt prosesseringstiden, noe som med tilgjengelig maskinvare ville vært lite hensiktsmessig. Videre er vannhøydene fra LiDAR-målingene erstattet med en jevn og konstant høyde fra FKB, som Romano (2004) påpekte var nødvendig på grunn av bølgelengden og dårlige målinger i vann med laser.

Høydemodellen med 10m oppløsning fra Statens Kartverk er best egnet for generelle analyser. Den klarer ikke fange opp detaljer i topografien på samme måte som den andre, og vil med det utelate potensielt viktig informasjon fra terrengparametrene og akkumuleringsanalysen. I utgangspunktet ville det vært mest hensiktsmessig å bruke LiDAR-data for hele kommunen, men ettersom maskinvaren ikke håndterte dette måtte en alternativ løsning benyttes. Generelle analyser laget fra Statens Kartverk (2014) sin høydemodell og en detaljert analyse med høy oppløsning ble den resulterende løsningen. Mye bearbeiding har blitt gjort for å gjøre modellen hydrologisk korrekt, men oppløsningen fra Statens Kartverk (2014) er etter min mening ikke høy nok til å beregne nøyaktige flomveier i urbane strøk. Modellen var imidlertid viktig for å få kjørt en OLS-regresjon på hele kommunen, slik at alle adressepunktene skulle få en verdi fra terrengparametre og akkumulering.

5.1.4 Evaluering av terrengparametre

Ettersom oppløsningen til høydemodellen er ulik vil også de ekstraherte terrengparametrene være ulike. En for grov oppløsning vil kunne føre til en betydelig forskjell mellom den egentlige høyden i terrenget og høydemodellen, hvor viktig informasjon kan gå tapt. Selv høyoppløste modeller kan være problematiske ved interpolering, da falske geografiske objekter som ikke hører hjemme i modellen blir beregnet og tatt med (Hengl, 2006). For eksempel er kurvaturvariabelen ekstremt sensitiv for feilkilder som oppstår etter interpolering (Olaya 2009; Hengl m.fl 2004).

Resultatene av terrengparametere vil altså variere med oppløsningen til høydemodellen (Hengl og Evans, 2009). Ifølge Florinsky og Kuryakova (2000) kan det være hensiktsmessig å teste ut terrengparametere på forskjellige oppløsninger for å finne den optimale oppløsningen som gir høyest korrelasjonskoeffisient til den avhengige variabel i regresjonsmodellen. Hengl (2006) påpeker allikevel at en slik optimal oppløsning vanskelig lar seg finne, da terrengparameterne nødvendig ikke gir høyest korrelasjonskoeffisient på samme oppløsning. For eksempel vil kurvaturvariabelen kunne fungere godt på 10x10m oppløsning, mens skråningsvariabelen helst burde vært hatt en høyere oppløsning for å predikere avhengig variabel. Hengl påpeker videre at testing av kombinasjoner av oppløsninger vil være tidkrevende og kun gjelde studieområdet som undersøkes (Hengl, 2006).

Med to ulike oppløsninger i høydemodellene er det testet hvilke korrelasjonskoeffisienter terrengparameterne har til avhengig variabel. Det har imidlertid ikke blitt gjort forsøk på for eksempel 10x10m kurvaturverdier sammen med 1x1 skråningsverdier i samme regresjonsmodell. Samtidig lyktes det ikke å skape et godt forhold mellom verken skråning eller aspekt til erstatningsutbetalinger. Dette kan tyde på at en ikke-lineær regresjonsmetode ville vært bedre egnet.

5.1.5 Evaluering av akkumuleringsraster

Åtte dreneringsretninger en grov forenkling av hvordan overflatestrømminger fungerer. D8-algoritmen er urealistisk når vannet i praksis kan drenerer til et punkt på uendelige forskjellige vis (Rød, 2015). Begrensningen av 8 retninger og medfølgende ulemper, gjorde at Tarboton (1997) foreslo en ny metode som bygget på D8, kalt D_{∞} (D-uendelig). D8 kan modellere konvergens (flere celler som drenerer til én celle). Etersom D8-algoritmen er en *single-neighbour flow* kan den ikke modellere divergens (en celle som drenerer til flere celler) (Gruber og Peckham 2009). Dette er den største svakheten til D8-metoden. Det finnes andre programvarer som kunne blitt benyttet for å nyttiggjøre seg av D-uendelig metoden, blant annet programvaren GRASS, men D8 metodens enkelhet og direkte integrasjon i ArcGIS ble foretrukket. Samtidig er et forbehold med D8 algoritmen at alt vannet går til avrenning. Ingenting blir kalkulert med i infiltrasjonsraten til jordsmonnet. Sånn sett er teorien til Horton (1933) om infiltrasjonsevne ikke direkte anvendbar i denne algoritmen, men ble allikevel inkludert i teorikapittelet da det etter mitt skjønn er viktig for å forstå begrensningen i analysemetoden som er brukt.

Videre er det forutsatt at sluk, kummer og ledninger er fulle og ikke har kapasitet til å ta imot mer vann. Den hydrologiske modellen er i begge høydemodellene blitt bearbeidet med å fjerne forsenkninger, erstatte vannhøydene med FKB data og heve bygninger med 5 meter. Men det finnes andre viktige objekter med hydrologisk betydning i landskapet som er viktig for beregning av flomveier (Duke m.fl, 2003). Veiunderganger, bruer og kulverter er slike objekter som har betydning for dreneringen. Spesielt i høydemodellen med 10x10 meter oppløsning vil små objekter kunne forsvinne. For den LiDAR deriverte høydemodellen er det mulig å se fortauskanter og kantstein grunnet bedre oppløsning. Det kunne ha vært hensiktsmessig å ha inkludert disse objektene i høydemodellen. Dette er ikke gjort og kan være en svakhet ved modellene. En studie av Murhphy m.fl (2008) sjekket betydningen av inkludering av slike objekter i to høydemodeller. En konvensjonell DEM på 10m oppløsning og en LiDAR-derivert DEM på 1m oppløsning. Begge høydemodellene hadde en hydrologisk korrekt versjon og en som ikke var bearbeidet, hvor resultatene i etterkant ble sammenlignet med manuelle målinger i området. For den hydrologisk korrekte høydemodellen med 10m oppløsning var resultatene bedre enn den som ikke var bearbeidet. Men studien fant også, at høydemodellen med 1m oppløsning *ikke* ga bedre beregninger av flomveier ved å gjøre den hydrologisk korrekt. De beste resultatene for beregning av flomveier var den ubearbeidede versjonen. Dette forklares ved at LiDAR-dataene allerede var veldig nøyaktige, og at introduksjon av andre data i modellen svekket nøyaktigheten og skapte falske objekter i modellen. Konklusjonen i studien var videre at høyere oppløsning var av større betydning enn å gjøre modellen hydrologisk korrekt (Murphy m.fl, 2008).

Valg av terskelverdi for flomveiene var også en utfordring. Olivera m.fl (2002) påpeker at høydedatabasen i USA bruker 5000 for å klassifisere bekker. Men dette var i 2002, og oppløsningen var på 30 x 30m. Selvsagt spiller oppløsningen en rolle her. I flomveirapporten til Meiforth (2013) ble det brukt 50.000 som terskelverdi for flomveier som sannsynligvis kan gjøre skade i Trondheim, med 1m oppløsning. 50.000 ble også valgt for høydemodellen i denne oppgaven med 1m oppløsning, mens 500 ble brukt for den med 10m oppløsning da denne er 10x så stor i X og Y retning. I kartene presentert i kapittel 4.2 er det ikke inndelt flomveier i klasser over terskelverdien, hvilket betyr at alle flomveier får den samme tykkelsen og fargen. Dette gir ikke et variert bilde av de ulike flomveiene som er beregnet. Men hensikten var å visualisere alle flomveier som potensielt er skadelige, og etter min mening gir dette et ryddigere kart som er mindre forvirrende for kartleser. Flomveiene som gikk ut i Mosvannet utviste merkelige sikk-sakk eller vinkelrette dreneringsretninger og var

såpass forstyrrende i kartet, at det ble plassert et vannpolygon over for å gjøre kartet mer leservennlig etter anbefaling fra Meiforth (2013). I figur 4.5 er det enkelte steder små streker som er beregnede flomveier i ytterkantene av bildet, dette er hvor LiDAR-utsnittet slutter og det ser unaturlig ut. Videre ble det ikke benyttet noen vektraster i beregningen, som kunne vært mulig å oppdrive fra online værtjenester. Ettersom volumberegninger ikke var det primære formålet, ble standardverdien til ArcGIS valgt istedenfor. Inspeksjon av resultatene som ble generert ble kun observert i selve programvaren og i flyfoto, det ble ikke foretatt noen befaringer i området, dette kunne gitt økt innsikt i problematikken.

Figur 4.5 og 4.6 viser at det går flomveier midt i det identifiserte problemområde. Samtidig er det også flomveier hvor det ikke er avdekket noen hot spot. At det finnes et samsvar mellom problemområde og flomveiene er allikevel tydelig. Utløpet til Mosvannet ligger på ca 36 meters høyde, mens hot spot ligger på mellom 50-75 meters høyde. Den naturlige dreneringsretningen vil gå mot vannet, og går via dette boligfeltet. Rett sør for Mosvannet synes også en beregnet flomvei som ikke henger sammen med nettverket, denne er blitt kuttet av ved en rundkjøring hvor det finnes en veiundergang. Dette kan tyde på at veiundergangen burde ha blitt senket ned slik som det ble gjort i rapporten til Meiforth (2013), og det synes ikke å stemme med Murphy m.fl (2008) sine funn at de mest nøyaktige flomveiene blir beregnet på en ubearbeidet LiDAR-derivert høydemodell.

5.1.6 Evaluering av OLS

Byggingen av en OLS-modell er en iterativ prosess som innebærer å fjerne, legge til og transformere variabler til man oppnår den best kalibrerte modellen. Noen avhengige variabler er mer intuitive og rett frem enn andre. For å tolke resultatene og diskutere modellen er det lagt til grunn ESRI (2013g) og James m.fl (2013) sine sjekklister for å kvalitetssikre og godkjenne resultatene:

1. Er minst én av de forklarende variablene X_1, X_2, \dots nyttige for forklare variasjonen i y ?
2. Matcher forholdene det forventede resultatet?
3. Utviser noen variabler redundans?
4. Er modellen min *biased*²?
5. Er alle nøkkelvariabler funnet?
6. Hvor godt blir den avhengige variabelen forklart?

² Ordet *biased* er et generelt begrep hvis betydning avhenger av kontekst. I denne oppgaven er det lagt til grunn skjevfordeling eller ikke-lineæritet

1.

Utvelgelsen av de forklarende variablene bør være forankret teoretisk til avhengig variabel. Alle terrengparametere, akkumulering samt anleggsårvariabelen var det i større eller mindre grad. Men forklarende variabler kan også inkluderes dersom de bygger på sunn fornuft, eller er fokus for oppgaven. For betydningen av typen avløpssystem ble det ikke konsultert teori i nevneverdig grad, bortsett fra det generelle. Via en kontaktperson i VA-avdelingen i Stavanger Kommunen ble det fremmet et ønske om å se nærmere på separate kontra felles systemer og derfor er inkluderingen av denne variabelen etter mitt skjønn rettfærdiggjort i modellen. Som man kunne se av resultatene viste det seg imidlertid at ikke alle de forklarende variablene hjalp modellen. Det var flere koeffisienter nær null, indikatorer på ingen forhold. I utgangspunktet kan det være hensiktsmessig å fjerne disse, såfremt det ikke er en god grunn til hvorfor de burde være der (ESRI, 2013g). I mitt skjønn var disse grunnene tilstede og variablene ble beholdt til tross for noen svake/ingen forhold.

Alle variablene ble i tillegg undersøkt for ikke-stasjonæritet. Når forholdet er ikke-stasjonært, betyr det at forholdet endrer seg når lokaliseringen endrer seg. En relativt normal hendelse for romlige data. I OLS-regresjonen er det den innebygde *Koenker's studentized Breusch-Pagan* som avdekker om slike forhold eksisterer ved å markere ved en asterisk. Dersom denne er statistisk signifikant, kan man kun stole på den robuste sannsynligheten for å avgjøre om det er variabelen er statistisk signifikant. Mine variabler utviste ingen statistisk signifikant ikke-stasjonæritet. Dette ville imidlertid ikke vært direkte negativt, da variabler med signifikant Koenker-verdi ville vært interessant å bruke i en GWR-regresjon (engelsk: *geographically weighted regression*) (ESRI, 2013g). Oppsummert vil det kunne sies at i begge regresjonsmodellene fantes gode prediktorer av y, mens andre hadde positive forhold uten å være signifikante og noen var dårlige prediktorer.

2.

Før regresjonsverktøyet ble satt i gang, var det forventede forholdet notert på forhånd. Etter kjøringen ble koeffisientene visuelt inspisert for å avkrefte eller bekrefte forventningene. Det er utfordrende når regresjonsmodellen rapporterer om forhold som ikke gir de resultatene man forventer. I mine regresjonsmodeller skjedde dette ved flere anledninger. I noen tilfeller ble

problemet løst ved å transformere variablene, mens i andre var det ikke mulig å etablere det forventede forholdet.

De forventede resultatet for akkumuleringen var et sterkt positivt forhold til erstatningsutbetalinger. Dette er relativt selvforklarende da erstatningsutbetalingene er selektert ut på vannskader, med to av årsakene innenfor seleksjonen som direkte går på overvannsproblematikk og ytre påvirkning. På toppen av dette ble også maksimumsfunksjonen benyttet i aggregeringen, noe som sørget for at den høyeste akkumulasjonsverdien blant adressepunktene i ruten ble gjeldende for hele ruten. Forventningene ble innfridd av resultatet, indikert ved at variabelen var statistisk signifikant i begge regresjonene.

For aspektvariabelen derimot, var forventningen noe usikker. I følge teorien skal nordvendt aspekt ha sammenheng med mindre innstråling og er generelt våtere enn sydvendt aspekt. Variabelen er imidlertid sirkulær med verdier fra 0 til 360° hvor retningen nord både har den laveste og høyeste verdien på henholdsvis 0-22.5° og 337.5-360°. Dette er ikke-lineært. Flere transformasjoner ble benyttet for å forsøke og gjøre denne variabelen bedre egnet for modellen. Først ble det forsøkt å bryte opp variabelen i to lineære komponenter ved bruk av cosinus og sinus transformasjoner. Kort fortalt blir det da en nord-sør komponent, og en øst-vest komponent (Pereira 1989; Jenness 2007). To egenskapsfelt ble opprettet hvor retningen cos-aspekt og sin-aspekt ble beregnet. Sinus verdiene gir verdier mellom -1 og 1 som er henholdsvis direkte vest og øst, mens cosinus verdiene som har samme verdiintervall hvor -1 er direkte sør og 1 er nord. En ny OLS-regresjon ble forsøkt med disse variablene, uten å lykkes i å skape et bedre forhold til erstatningsutbetalinger. Verdien for Akaike's informasjonkriterium (engelsk: Akaike's Information Criterion (AICc)) ble høyere samtidig som R²-verdien ble lavere. Generelt sett er den beste modellen den som har lavest AICc-verdi og høyest R²-verdi. En annen metode for å transformere aspekt ble laget av Robert og Cooper (1989, sitert i Barber og Vanderzanden 2009, s.57) hvor verdiene blir inndelt fra 0 til 1. Nordvendt aspekt blir gitt 0-verdien mens 1 blir gitt til sydvendt aspekt. Resultatet blir da en kontinuerlig variabel mellom dette intervallet. Et verktøy som benytter seg av denne metoden og er kompatibel med ArcGIS er kalt Slope/Aspect transformation. Verktøyet er brukerlagd, og finnes ikke direkte i selve programvaren til ESRI (Evans, 2003).

Etter bruk av verktøyet ble regresjonen kjørt enda en gang, men heller ikke denne gangen lyktes det å skape et godt forhold til erstatningsvariabelen. Den opprinnelige

aspektvariabelen, uten transformasjoner ble derfor beholdt, til tross for at den ikke hjalp modellen. OLS forutsetter at forholdene er lineære, og når dette ikke er tilfelle vil variabelen gi svake resultater. En studie av Ranhao m.fl (2008) påpekte at topografiske variabler som aspekt og skråning ikke egner seg i lineære regresjonsmetoder. Dette gjaldt riktignok for nedbør som avhengig variabel, men det er trolig at en ikke-lineær metode ville vært bedre egnet også her.

Forventningen for kurvatur var at det skulle være positivt. Variabelen er en kombinasjonen av plan- og profilkurvatur som betyr at de høyere verdiene er knyttet til konvekse former hvor vannhastigheten er høyere og strømmingene er divergente. Resultatet ble positivt i begge regresjonene men med tydeligere positivt forhold for hele kommunen. De ble allikevel ikke statistisk signifikante, og det blir vanskelig å trekke noen slutninger basert på det.

Forventningen for anleggsår var at det skulle være et negativt forhold. Ettersom eldre rør gjerne er forbundet med slitasje og lekkasjer, skulle det tilsi at erstatningsbetalingene gikk opp desto eldre rørene ble. Resultatet ble riktignok negativt, men med koeffisienter nær 0 ved begge regresjonene, indikerer dette ingen forhold. Samtidig er det interessant å notere at i rutene hvor den gjennomsnittlige alderen på rørene var fra 1960-1979 (se figur 4.8) er tilknyttet de høyeste erstatningsutbetalingene. Dette kan tyde på at det er materialvalg og anleggsmetode fremfor alderen i seg selv, som er betydningsfulle for skadeutbetalinger tilknyttet rørene jf. rapporten fra Eigersund Kommune (2014), som påpekte at etterkrigstidens rør hadde lavest kvalitet. Dette stemmer for øvrig overens med en studie fra Sverige som fant at ledninger fra 1960-1970 årene forårsaker de fleste skadene og har høyest kostnader (VVS Företagan, 2005)

Dummy-variabelen var den siste avløpsspesifikke variabelen som var med i regresjonene. Sunn fornuft tilsier at det moderne separate systemet skal være bedre enn fellessystemet, som er av eldre dato. Når rør av fellessystem skal skiftes, er det dessuten vanlig å bytte til separat system (Fredrikstad Kommune, 2007). For hele kommunen utviste variabelen ikke et statistisk signifikant forhold, men hadde en positiv koeffisient. For utsnittet av kommunen derimot, hvor det identifiserte problemområde befinner seg, ble variabelen statistisk signifikant.

Ettersom fellessystem er kodet 0 og separatsystemet 1, vil tolkningen være at i ruter som har typetallet 1, og dermed flest separate system, har høyere verdi av erstatningsutbetalinger enn ruter som har mest fellessystem. Et relativt interessant funn, da dette strider mot hva som er

normal praksis i utskiftninger av rør. Ettersom koeffisienten var langt høyere for det identifiserte problemområde, kan det indikere at betydningen av ledningssystemet er relevant for erstatningsutbetalingene i dette område. Typetall som aggregeringsmetode betyr at ledningssystemet som forekommer oftest, representerer hele ruten. For eksempel, hvis en ledning med separat system er delt opp i mange linjesegmenter, mens fellessystemet innenfor samme rute er en kontinuerlig linje uten å være oppdelt, vil det separate systemet gjelde for ruten selv om det skulle vise seg at antall meter med ledning er mest felles system. Dette kan være en mulig feilkilde til metoden benyttet for å aggregere variabelen. Oppsummert ga ikke skråning og aspekt det forventede resultatet, til tross for transformasjoner og bearbeiding, men ble allikevel beholdt på grunn av teoretiske røtter til hvordan vannet flyter.

3.

Dersom man har redundans blant forklaringsvariablene vil det si at samme porsjon av den avhengige variabelen er forklart av flere forklaringsvariabler. Dette problemet oppstår når flere av de forklarende variablene ikke er uavhengige av hverandre, noe som er en forutsetning for OLS-regresjon. Målet på dette er kalt varians inflasjonsfaktor (engelsk: *Variance Inflation Factor*) og forkortes til VIF i diagnostikkrapporten (Rød 2015). Terskelverdien for VIF er på 7.5, dersom flere av de forklarende variablene har en verdi som overstiger dette, bør de fjernes en og en, ellers vil modellen være veldig ustabil og kunne gi en kunstig høy R^2 (ESRI, 2013e). Ingen av variablene i regresjonene indikerte en VIF-verdi på over 7.5, og sånn sett ble dette potensielle problemet ikke berørt.

4.

Målene som forteller at resultatene er skjeve eller ikke-lineære (engelsk: *biased*) er indikert ved residualer som ikke er normalfordelte, og en statistisk signifikant Jarque-Bera verdi. En potensiell årsak til at modellen er skjev kan være mangel på nøkkelvariabler, eller outliers (observasjoner som er langt utenfor de andre observasjonene). I akkumuleringsvariabelen fantes det én ganske klar outlier (se figur 4.7). Denne kan i teorien fjernes for en bedre modell, men bare dersom man mener at observasjonen skyldes en feil i datasettet. Akkumuleringsverdiene er et direkte produkt av den bearbeidede høydemodellen, og hvis det finnes feil i denne vil det naturligvis påvirke resultatet. Det ble forsøkt å fjerne denne observasjonen først, og riktignok førte det til en høyere R^2 (14%). Allikevel ble denne observasjonen beholdt i regresjonen, da det ikke kunne bekreftes at denne skyldtes en feil i datasettet.

Normalfordelte residualer er som nevnt viktig for å bekrefte at modellen ikke utviser skjevhet. Ingen av OLS-regresjonene hadde residualer som var perfekt normalfordelte (se figur 4.11 og 4.16) Diagnostikkfilen i verktøyet forteller at dersom histogrammet er *veldig* forskjellig fra normalkurven, kan det tyde på skjevhet. Hvis denne er statistisk signifikant, vil det også bli representert med en (*) ved Jarque-Bera verdien. Etersom sistnevnte ikke var tilfelle for noen av modellene, og fordelingen av residualer var relativt likt normalkurven, ansees dette resultatet som innenfor grensen av hva som er akseptabelt. Videre ble romlig autokorrelasjonsverktøy brukt i tillegg for å se hvorvidt residualene utviste signifikant autokorrelasjon, men begge modellenes residualer falt innenfor kategorien tilfeldig noe som understøtter at modellen ikke utviser skjevhet. Oppsummert fantes ingen direkte tegn på skjevhet i modellen, verken på romlig autokorrelerte residualer, statistisk signifikant Jarque-Bera verdi eller stor forskjell mellom histogram av residualer og normalfordelt kurve.

5.

Indikatoren på at man mangler én eller flere nøkkelvariabler er romlig autokorrelerte residualer. Da vil både over- og underprediksjonene være geografisk konsentrert. I kjøringen av OLS analysen er dette den første advarselen som dukker opp, da dette er viktig å avkrefte før man aksepterer resultatene fra diagnostikkrapporten. I kjøringen av romlig autokorrelasjonsverktøyene ble det avdekket at ingen av regresjonene utviste tegn på romlig autokorrelasjon (se figur 4.10 og 4.15). Det fantes ingen statistisk signifikant z-score som indikerer at nøkkelvariabler manglet fra modellen. I figurene 4.9 og 4.14 viser begge kartene noen røde ruter sør for Mosvannet. Dette indikerer at de reelle observasjonene for erstatningsutbetalingene er større enn modellen har predikert at det skulle være. Samtidig finnes det blå ruter flere andre steder, som viser at de reelle observasjonene for erstatningsutbetalingene er lavere enn den predikerte verdien. Men, ingen av disse over- eller underprediksjonene synes å være signifikant annerledes enn en tilfeldig fordeling. Dette trenger allikevel ikke bety at alle relevante forklarende variabler er funnet. For eksempel ville betydningen av innredede kjellere, samt en dummy variabel med urbane/rurale områder vært potensielt viktige variabler for erstatningsutbetalingene.

6.

Vurderingen for hvor godt den avhengige variabelen er forklart, er R^2 en viktig indikator på. Denne angir hvor godt de forklarende variablene, forklarer variasjonen i den avhengige variabelen. Man kan imidlertid ikke stole på denne verdien ukritisk. Det er nødvendig at de fem omtalte sjekkpunktene nevnt ovenfor er 'bestått'. I noen tilfeller kan R^2 fungere godt innenfor noen områder, eller innenfor et gitt intervall av verdier, samtidig som den ikke fungerer i det hele tatt i andre områder. Dette kan ikke tolkes direkte fra R^2 . Koeffisientene kan heller ikke stoles fullt på dersom residualene viser tegn til romlig autokorrelasjon. Ved høy VIF-verdi kan man få ekstremt høye, men kunstige, R^2 verdier. Dette vil gjøre modellen svært ustabil (ESRI, 2013g).

Samtidig må en påpeke at tolkningen av R^2 gjøres subjektivt. Hva som er en god forklarende modell, avhenger fullstendig av problemstillingen en ønsker å besvare. Nullhypotesen er at koeffisientene *ikke* hjelper modellen. Kan denne forkastes? En samlet F-statistikk som er statistisk signifikant i begge modellene tilsier at den kan det. Modellen er effektiv i å forklare variasjonen i erstatningsutbetalinger men samtidig er forklaringssevnen på 10 prosent relativt lav. Modellen med 27 prosent forklaringssevne er på en annen side relativt god, med tanke på det komplekse fenomenet som modelleres. Videre er ikke en lav R^2 -verdi synonymt med en dårlig modell (ESRI, 2013e). I enkelte områder av vitenskapen er en R^2 verdi på omkring 20 prosent oppsiktsvekkende, mens i andre fagområder må R^2 verdien helst opp på 80-90 prosent før noen hever et øyenbryn (ESRI, 2013g). Viktigere enn en høy R^2 verdi, er etter mitt skjønn at modellen er pålitelig og ikke har nøkkeltall som indikerer at troverdigheten er svak. Med bakgrunn i de fem overnevnte punktene, anses sjekklisten som bestått og resultatene pålitelige.

6. Konklusjon

Målet med regresjonsmodellene var økt forståelse for dataene identifisert i hot spot analysene, og å bruke denne innsikten til å svare på problemstillingen. Hot spot analysene identifiserte problemområder, både kjente og nye og viste at erstatningsutbetalinger samsvarer med flomveiene beregnet fra begge høydemodellene. Regresjonsmodellen som var myntet på kun det avdekkede problemområdet ga den høyeste R^2 verdien. Mens modellen som hadde hele kommunen som utgangspunkt, hadde en lav forklaringssevne. Dette er ikke nødvendigvis ensbetydende med en dårlig modell. Modellering av erstatningsutbetalinger er et komplekst fenomen som sannsynligvis kan ha flere variabler som er effektive i å predikere variasjonen. Nyttig informasjon er også ervervet ved at skråning og aspekt tilsynelatende ikke egner seg i god grad til modellering av den avhengige variabelen, og kunne vært bedre egnet i en ikke-lineær regresjonsmetode.

I innledningskapittelet ble det nevnt at klimatilpasning handler om erkjennelsen av et klima i endring, hvor tiden for handling er nå. Pilotprosjektet mellom forsikringsselskapene og pilotkommunene er en vitnesbyrd om nettopp dette. Et samarbeid hvor begge parter har erkjent at hyppigere hendelser vil komme, og hvor nye løsninger tas i bruk for å forebygge, hindre og handle proaktivt. Skadedataene har vært et viktig skritt i denne sammenheng og i denne oppgaven ble de brukt til å besvare flere spørsmål:

Hovedspørsmål

1. Kan vi, ved hjelp av tilgang på forsikringsselskapenes naturskadestatistikk, generere kunnskap om nye problemområder?

Svar: I startfasen av oppgaven ble det diskutert med mine kontaktpersoner i Stavanger Kommune om hvilke områder som var sannsynlige å peke seg ut som problemområder. Hot spot analysene avdekket et område sør for Mosvannet som hadde geografiske konsentrasjoner av høye erstatningsutbetalinger, både hvor alle adressepunktene var tatt med og hvor selekterte årsaker var hentet ut. Videre var intensiteten av hendelser her også signifikant i analysen med selekterte årsaker. Dette problemområde ble ikke diskutert eller tatt opp i møter med kommunen. Istedenfor ble andre områder utpekt som sannsynlige problemområder som ville avdekkes med analysene. Jeg konkluderer derfor at tilgang på forsikringsselskapenes naturskadestatistikk har generert kunnskap om nye problemområder.

Delspørsmål

1. Vil kartlegging av skadedata fra forsikringsselskapene identifisere problemområder i kommunen og dermed gi kommunen et bedre kunnskapsgrunnlag for å forebygge klimarelaterte skader?

Svar: Kartleggingen av skadedataene har gitt økt innsikt i hvor problemområdene er. Både alene og i kombinasjon med støttende datamaterialet. Skadedataene er satt i kontekst og sammenhenger kan tolkes på bakgrunn av sammenkoblingen med øvrige data. Med henvisning til regresjonsmodellene vises det til at karakteristikkene til ledningsnettene er blitt undersøkt. I følge mine resultater spiller ikke alderen på avløpsnettene stor betydning for erstatningsbeløpene i en lineær regresjon. Det viser seg allikevel at alderskategorien 1960-1979 er særlig knyttet til høye erstatningsutbetalinger. Det skal ikke antas at dette er ny kunnskap, men selv en bekreftelse av kommunens antakelser vil være en styrking i kunnskapsgrunnlaget.

Samtidig ble det for det identifiserte problemområde avdekket i regresjonen at de ulike avløpssystemene har betydning for erstatningsutbetalingene. Betydningen er statistisk signifikant i det identifiserte problemområde hvor separate systemer synes å være forbundet med høye erstatningsutbetalinger. Dette er et interessant funn med tanke på den rådende utskiftningsstrategien hvor eldre felles systemer blir erstattet med separate. Det finnes allikevel feilkilder knyttet til aggregeringsmetoden av variabelen, og antall observasjoner i det identifiserte problemområde er relativt lavt. Det skal derfor utvises forsiktighet i å slå fast dette som fakta, men funnet bør etter mitt skjønn nyansere kunnskapsgrunnlaget noe. Videre undersøkelser bør utføres for å se nærmere på betydningen av avløpssystemer for erstatningsutbetalinger.

2. Vil modellering av flomveier basert på høyoppløselige terrengmodeller gi et realistisk bilde av hvor flomveiene vil gå, og stemmer dette overens med problemområder identifisert ved kartlegging av skadedataene?

Svar: Algoritmen som er benyttet for å modellere flomveier inneholder en del feilkilder. Blant annet kun åtte dreneringsretninger, og premisset om at all nedbør går til avrenning, samt at sluk og ledninger er ved full kapasitet. Videre er høydemodellen med 10x10 meter oppløsning såpass grov at små, men viktige objekter ikke inkluderes. Flomveikartene

beregnet fra 1x1 meter modellen er bedre i dette aspektet. Kartet over flomveiene synes å samsvare bra med FKB-data for Møllebekken, men ettersom det ikke er gjort noen befaringer i kommunen er det krevende å slå fast hvor nøyaktig de beregnede flomveiene er. Setter man opp flomveiene ved siden av hot spot analysen for erstatning ser man allikevel at det går flomveier tvers igjennom det identifiserte problemområde, noe som samsvarer godt med de selekterte årsakene som er hentet ut. Videre var akkumuleringsvariabelen statistisk signifikant i begge regresjonsmodellene, noe som vitner om at de beregnede flomveiene i høy grad kan knyttes til erstatningsutbetalingene.

3. Kan en kobling med forsikringsdata opp mot andre temadata gi en økt innsikt i hvorfor identifiserte problemområder er problematiske?

Svar: Regresjonene har vist at akkumuleringsverdier er av stor betydning for hvor det finnes vannskader. Dette var ikke uventet. Derimot fikk dummy-variabelen høyere koeffisient enn forutsett, spesielt for utsnittet hvor den var statistisk signifikant. Kobling av forsikringsdata mot avløpsdatasettet har gitt økt innsikt i hvorfor det identifiserte problemområdet er problematisk. Samtidig viser terrengparameterne liten kollektiv forklaringsgrad i den avhengige variabelen, noe som kan tyde på at ytterligere temadata bør inkluderes i videre arbeid, gjerne flere egenskapsdata knyttet til ledningsdatasettet som f.eks ledningsmaterialet og betydningen av rurale/urbane områder.

6.1.1 Videre arbeid

Den metodiske tilnærmingen til oppgaven med hot spot analyser og videre til statistiske beregninger via OLS-regresjon er noe undertegnede mener passer godt sammen. Såfremt kommunene har tilgang på skadedataene fra forsikringsselskapene kan denne metodikken utføres og videreutvikles. Kobling av skadedataene til LiDAR-deriverte høydemodeller med supplerende FKB-data hjelper å sette dataene i kontekst og hente ut detaljerte terrengparametere samt flomveier. Dersom flomveiberegninger er hovedformålet i andre analyser, vil det være hensiktsmessig å benytte datasett med høyest tilgjengelig oppløsning, samt senke ned veiunderganger, kulverter og bruer. Videre kan også en mer sofistikert algoritme enn D8-metoden benyttes.

7. Referanser

- Aftenposten. (2012). *Disse står ansvarlige ved flom- og rasskader*. Hentet 07.05.2015 fra: <http://www.aftenposten.no/nyheter/iriks/Disse-star-ansvarlig-ved-flom--og-rasskader-6961981.html>
- Barber, J., Vanderzanden, D. (2009). *The Region 1 Existing Vegetation Map Products (vmap) Release 9.1.1*. United States Department of Agriculture Forest Service, Missoula, Montana. Rapport 09-03. Hentet 29.04.2015 fra: http://www.fs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb5342838.pdf
- Bater, W. C., Coops, C. N. (2009). Evaluating error associated with lidar-derived DEM interpolation. *Computers & Geosciences*, 35(2), s. 289-300
- Blaga, L. (2012). Aspects regarding the significance of the curvature types and values in the studies of geomorphometry assisted by GIS. *Analele Universității din Oradea – Seria Geografie*, 22(2), s.327-337
- Brevik, R., Aall, C., Rød, K. J. (2015). *Pilotprosjekt om testing av skadedata fra forsikringsbransjen for vurdering av klimasårbarhet og forebygging av klimarelatert naturskade i utvalgte kommuner*. Sogndal: Vestlandsforskning. Hentet 12.02.2015 fra: <https://www.fno.no/Nyheter/2015/02/godt-mottatt-pilotprosjekt/>
- Caravaglia, S., Sharma, A. (2000). A smart guide to dummy variables: Four applications and macro. NESUG Conference Proceedings, Philadelphia PA. Hentet 23.03.2015 fra: <http://www.ats.ucla.edu/stat/sas/library/nesug98/p046.pdf>
- Childs, C. (2004). *Interpolating surfaces in ArcGIS spatial analyst*. ESRI Education Services. Hentet 27.04.2015 fra: <http://www.esri.com/news/arcuser/0704/files/interpolating.pdf>
- Christensen, P. (2012). 35 kjellere måtte pumpes. I: *Stavanger Aftenblad*. Hentet 10.05.2015 fra: <http://www.aftenbladet.no/nyheter/lokalt/stavanger/35-kjellere-matte-pumpes-3007226.html>
- DeMers, N.M. (2005). *Geographic Information Systems*. (3. utg.). New Mexico, USA: John Wiley & Sons, Inc

- Duke, D.G., Kienzle, W.S., Johnson, L.D., Byrne, M.J. (2003). Improving overland flow routing by incorporating ancillary road data into Digital Elevation Models. *Journal of Spatial Hydrology*, 3(2), s. 1-27
- Eck, J.E., Chainey, S., Cameron, J.G., Leitner, M., Wilson, R.E. (2005). *Mapping crime: understanding hotspots*. U.S Department of Justice, Washington, DC. Hentet 23.04.2015 fra: <http://discovery.ucl.ac.uk/11291/1/11291.pdf>
- Eigersund Kommune, 2014. *Sektorplan for vann og avløp, 2014-2025 i Eigersund kommune*. Høringsutkast. Hentet 23.04.2015 fra: <http://www.eigersund.kommune.no/getfile.php/2752588.1621.svwrufebes/Sektorplan+for++vann+og+avl%C3%B8p.pdf>
- ESRI. (2015). *How Incremental Spatial Autocorrelation works*. Hentet 24.03.2015 fra: <https://desktop.arcgis.com/en/desktop/latest/tools/spatial-statistics-toolbox/how-incremental-spatial-autocorrelation-works.htm>
- ESRI. (2014a). *Optimized Hot Spot analysis (Spatial Statistics)*. Hentet 07.05.2015 fra: <http://resources.arcgis.com/en/help/main/10.2/index.html#//005p00000058000000>
- ESRI. (2014b). *Statistical analysis for LAS datasets*. Hentet 27.01.2015 fra: <http://resources.arcgis.com/en/help/main/10.2/index.html#//015w00000049000000>
- ESRI. (2013a). *Assessing lidar coverage and sample density*. Hentet 22.02.2015 fra: <http://resources.arcgis.com/en/help/main/10.1/index.html#//015w0000003t00000000>
- ESRI. (2013b). *Hot Spot Analysis (Getis-Ord Gi*) (Spatial Statistics)*. Hentet 24.04.2015 fra: <http://resources.arcgis.com/en/help/main/10.1/index.html#//005p00000010000000>
- ESRI. (2013c). *How Incremental Spatial Autocorrelation works*. Hentet 18.01.2015 fra: <http://resources.arcgis.com/en/help/main/10.1/index.html#//005p00000055000000>
- ESRI. (2013d). *Integrate (Data Management)*. Hentet 24.04.2015 fra: <http://resources.arcgis.com/en/help/main/10.1/index.html#//00170000002s000000>
- ESRI. (2013e). *Interpreting OLS Results*. Hentet 22.02.2015 fra: <http://resources.arcgis.com/en/help/main/10.1/index.html#//005p00000030000000>

ESRI. (2013f). *Regression analysis basics*. Hentet 23.02.2015 fra:

http://resources.arcgis.com/en/help/main/10.1/index.html#/Regression_analysis_basics/005p00000023000000/

ESRI. (2013g). *What they don't tell you about regression analysis*. Hentet 23.02.2015 fra:

<http://resources.arcgis.com/en/help/main/10.1/index.html#/005p00000053000000>

ESRI. (2012a). *Determining flow direction*. Hentet 18.01.2015 fra:

http://webhelp.esri.com/arcgisDEsktop/9.3/index.cfm?TopicName=Determining_flow_direction

ESRI. (2012b). *How Hot Spot Analysis (Getis-Ord G_i^*) works*. Hentet 23.02.2015 fra:

<http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/005p00000011000000>

ESRI. (2011a). *How Aspect Works*. Hentet 28.04.2015 fra:

http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/How_Aspect_works/009z000000vp000000/

ESRI. (2011b). *How Sink works*. Hentet 18.01.2015 fra:

<http://webhelp.esri.com/arcgisdesktop/9.3/index.cfm?TopicName=How%20Sink%20works>

ESRI. (2011c). *How Slope works*. Hentet 07.05.2015 fra:

http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/How_Slope_works/009z000000vz000000/

ESRI. (2011d). *Raster Calculator (Spatial Analyst)*. Hentet 22.02.2015 fra:

<http://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/009z000000z7000000.0.htm>

ESRI. (2010). *Understanding curvature rasters*. Hentet 28.04.2015 fra:

<http://blogs.esri.com/esri/arcgis/2010/10/27/understanding-curvature-rasters/>

- ESRI. (2009). *Collect Events (Spatial Statistics)*. Hentet 23.04.2015 fra:
http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?id=1917&pid=1915&topicname=Collect_Events_%28Spatial_Statistics%29
- ESRI. (2008). *Cell size of raster data*. Hentet 02.02.2015 fra:
http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?id=2655&pid=2652&topicname=Cell_size_of_raster_data
- ESRI. (2007). *Symbolizing the results of a hot spot analysis*. Hentet 23.04.2015 fra:
<http://blogs.esri.com/esri/arcgis/2007/08/01/symbolizing-the-results-of-a-hot-spot-analysis/>
- ESRI. (udatert(a)). *GIS Dictionary – ESRI Grid*. Hentet 07.05.2015 fra:
<http://support.esri.com/en/knowledgebase/GISDictionary/term/ESRI%20Grid>
- ESRI. (udatert(b)). *GIS Dictionary – spatial autocorrelation*. Hentet 13.02.2015 fra:
<http://support.esri.com/en/knowledgebase/GISDictionary/term/spatial%20autocorrelation>
- ESRI. (udatert(c)). *Surface Creation and analysis*. Hentet 28.04.2015 fra:
http://resources.esri.com/help/9.3/arcgisengine/java/gp_toolref/geoprocessing/surface_creation_and_analysis.htm
- Evans, 2003. *ArcGIS Geomorphometry & Gradient Metrics*. Hentet 24.04.2015 fra:
[http://s3.amazonaws.com/RTools/Gomorphometry+%26+Gradient+Metrics++\(version+2.0\).zip](http://s3.amazonaws.com/RTools/Gomorphometry+%26+Gradient+Metrics++(version+2.0).zip)
- Finans Norge, 2013. *Sterk økning i vannskader*. Hentet 12.02.2015 fra:
<https://www.fno.no/Hoved/Aktuelt/Pressemeldinger/2013/sterk-okning-i-vannskader/>
- Finans Norge. (udatert). *Om Finans Norge*. Hentet 12.02.2015 fra:
<https://www.fno.no/Hoved/OmFNO/>
- Florinsky, V.I., Kurykakova. A.G. (2000). Determination of grid size for digital terrain modelling in landscape investigations – exemplified by soil moisture distribution at a microscale. *International Journal of Geographical Information Science*, 14(8), s. 815-832

- Fredrikstad Kommune. (2007). *Avløpsledninger*. Hentet 21.04.2015 fra:
<https://www.fredrikstad.kommune.no/Tjenester/Vann-og-avlop/Avlop---offentlige-avlopsanlegg/Avlopsledninger/>
- Gallant, C. J., Wilson, P.J. (2000). Digital terrain analysis. I: Gallant, C.J., Wilson, P.J. (red.) *Terrain analysis: Principles and Applications*. John Wiley & Sons Ltd, New York, s. 1-27
- Getis, A., Ord, K.J. (1992). The Analysis of Spatial Association by Use of Distance Statistics. *Geographical Analysis*, 24, s. 189-206
- Grann, L.S. (2011). *Utrygghet for flom - en betinget verdsettingsstudie*. Masteroppgave. Institutt for økonomi- og ressursforvaltning, Universitetet for miljø- og biovitenskap
- Gruber, S., Peckham, S. (2009). Land-Surface Parameters and Objects in Hydrology. I Hengl, T., Reuter, H.I. (red.) *Geomorphometry: Concepts, Software, Applications*. Developments in Soil Science, 33. Elsevier, s. 171-194.
- Hengl, T. (2006). Finding the right pixel size. *Computers & Geosciences*, 32(9), s. 1283-1298
- Hengl, T., Gruber, S., Shrestha, P.D. (2004). Reduction of errors in digital terrain parameters used in soil-landscape modelling. *International Journal of Applied Earth Observation*, 5(2), s. 97-112
- Hengl, T., Evans, I.S. (2009). Mathematical and Digital Models of the Land Surface. I: Hengl, T., Reuter, H.I. (red.) *Geomorphometry: Concepts, Software, Applications*. Developments in Soil Science, 33, Elsevier, 33, s. 31-63
- Horton, R.E. (1933). *The role of infiltration in the hydrological cycle*. *Transactions of the American Geophysics Union*, 14, s. 446-460
- IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (red.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.
- IPCC, 2007: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate

Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (red.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp.

James, G., Witten, D., Hastie, T., Tibshirani, R. (2013). *An Introduction to Statistical Learning – with Applications in R*. New York, Springer

Jenness, J. (2007). *Some thoughts on Analyzing Topographic Habitat Characteristics*. Jenness Enterprises. Hentet 29.04.2015 fra:

http://www.jennessent.com/downloads/topographic_analysis_online.pdf

Johnston, K., Ver Hoef, M. J., Krivoruchko, K., Lucas, N. (2001). *Using ArcGIS Geostatistical Analyst*. ESRI Press: Redlands, CA. Hentet 27.04.2015 fra:

http://dusk2.geo.orst.edu/gis/geostat_analyst.pdf

Jones, R. (2002). Algorithms for using a DEM for mapping catchment areas of stream sediment samples. *Computers & Geosciences*, 28(9), s. 1051-1060.

Kennedy, M. (2006). *Introducing Geographic Information Systems with ArcGIS*. John Wiley & Sons, Inc., Hoboken, New Jersey

Legendre, P. (1993). Spatial Autocorrelation: Trouble or New Paradigm? *Ecology*, 74(6), s.1659-1673.

Liu, X. (2008). Airborn LiDAR for DEM generation: some critical issues. *Progress in Physical Geography*, 32(1), s. 31-49

Liu, X., Zhang, Z., Peterson, J., Chandra, S. (2007). The effect of LiDAR data density on DEM accuracy. *Proceedings of International Congress on Modelling and simulation (MODSIM07)*. Christchurch, New Zealand, s. 1363-1369

Lloyd, C.D., Atkinson, P.M. (2002). Deriving DSMs from LiDAR data with kriging. *International Journal of Remote Sensing* 23(12), s. 2519-2524

- Longley, P. A., Goodchild, M. F., Maguire, D. J. og Rhind, D. W. (2011). *Geographic Information Systems & Science*. (3. utg.). New Jersey, USA: John Wiley & Sons, Inc.
- Longley, P. A., Goodchild, M. F., Maguire, D.J og Rhind, D.W. (2005). *Geographic Information Systems and Science*. (2.utg). John Wiley & Sons, Ltd.
- Lovdata. (2012a). *Force majeure*. Hentet 03.03.2015 fra:
https://lovdata.no/artikkel/force_majeure/103
- Lovdata. (2012b). *Høyesterett – Objektivt erstatningsansvar etter forurensningsloven*. Hentet 30.04.2015 fra: https://lovdata.no/artikkel/hoyesterett_-_objektivt_erstatningsansvar_etter_forurensningsloven/141
- Lovdata. (2013). *Lov om forsikringsavtaler (forsikringsavtaleloven)*. Hentet 03.03.2015 fra:
https://lovdata.no/dokument/NL/lov/1989-06-16-69#KAPITTEL_1
- Lovdata. (2015). *Lov om vern mot forurensninger og om avfall (forurensningsloven)*. Hentet 30.04.2015 fra: https://lovdata.no/dokument/NL/lov/1981-03-13-6/KAPITTEL_4#§24a
- Mark, D.M. 1988. Network Models in Geomorphology. Modelling Geomorphological Systems. Anderson, M.G. New York, John Wiley: 73–97.
- Martz, W.L., Garbrecht, J. (1998). The treatment of flat areas and depressions in automated drainage analysis of raster digital elevation models. *Hydrological Processes*, 12(6), s.843-855
- Meiforth, J. (2013). *Modelling flood streams for urban planning in Trondheim, Norway*. Project UniGIS Professional, University of Salzburg. Hentet 01.05.2015 fra:
<http://www.trondheim.kommune.no/attachment.ap?id=59632>
- Miljødirektoratet. (2012). *Forsikring, avkorting og regress*. Hentet 21.01.2015 fra:
http://www.miljodirektoratet.no/no/Klimatilpasning_Norge/Bibliotek/Generelle-tekster/Forsikring-avkorting-og-regress/
- Miljøkommune. (2015). *Litt om ansvar og erstatning*. Miljødirektoratet i samarbeid med Landbruksdirektoratet. Hentet 21.01.2015 fra
<http://www.miljokommune.no/Temaoversikt/Overvann/Om-ansvar-og-erstatning-/>

- Miljøverndepartementet. (2010). *Referat fra fagmøte om forsikring og naturskade*. Klima- og forurensningsdirektoratet 4. Februar 2010, Oslo. Hentet 23.04.2015 fra https://www.regjeringen.no/globalassets/upload/md/vedlegg/klima/klimatilpasning/underlagsrapporter/referat_fagmote_forsikring_04022010.pdf
- Murphy, C.N.P., Ogilvie, J., Meng, R.F., Arp, P. (2008). Stream network modelling using lidar and photogrammetric digital elevation models: a comparison and field verification. *Hydrological Processes*, 22, s. 1747-1754
- Myrstad, T., Fosse, H., Langnes, S.M. (2011). Ekstremvær – hvem tar regninga? *Hegnar*. Hentet 07.05.2015 fra: <http://www.hegnar.no/juss/artikkel474700.ece>
- NGI. (2010). *An overview of LiDAR: collection to applications*. Hentet 29.04.2015 fra: <http://www.ngi.no/upload/48594/1%20What%20Is%20Lidar.pdf>
- Nilsdal, J.K. (1985). Kallaroversvømming i samband med regn. *Geohydrologiska Forskningsgruppen Chalmers Tekniske Hogskola*, Meddelande nr.79. Hentet 23.04.2015 fra: http://publications.lib.chalmers.se/records/fulltext/179513/local_179513.pdf
- Norsk Vann. (2013a). *HR-2007-522-A – Rt-2007-431*. Norges Høyesterett – dom. Hentet 22.02.2015 fra: http://va-jus.no/wp-content/uploads/2011/11/Rt_2007_431_Stavanger.pdf
- Norsk Vann. (2013b: Rapport B17). *Investeringsbehov i vann- og avløpssektoren*. Hentet 23.04.2015 fra: [www.norskvann.no/kompetanse/va-bokhandelen/rapporter/product/download/file_id-365">rapport_b17_2013.pdf](http://www.norskvann.no/kompetanse/va-bokhandelen/rapporter/product/download/file_id-365)
- Norsk Vann, (2008:162). Lindholm, O., Endresen, S., Thorolfson, S., Sægrov, S., Jakobsen, G., Aaby, L. *Veiledning i klimatilpasset overvannshåndtering*. Hentet 21.01.2015 fra: http://www.norskvann.no/component/rsfiles/view?path=tosidere/r_162_overvannshandtering.pdf
- NOU. (1996:16). *Tiltak mot flom*. Hentet 21.01.2015 fra: <https://www.regjeringen.no/contentassets/21e6f5f6fe424757a8d418a33259c223/no/pd/fa/nou199619960016000dddpdfa.pdf>

- NOU. (2010:10). *Tilpassing til eit klima i endring*. Hentet 04.12.2014 fra:
<https://www.regjeringen.no/contentassets/01c4638b3f3e4573929f3b375f4731e0/nn-no/pdfs/nou201020100010000dddpdfs.pdf>
- NRK, 2015. *Erstatning for naturskader skyter i været*. Hentet 28.04.2015 fra:
<http://www.nrk.no/fordypning/erstatning-for-naturskader-skyter-i-vaeret-1.12319203>
- NVE. (1997:10). *Anvendt Urbanhydrologi*. Hentet 23.04.2015 fra:
http://webby.nve.no/publikasjoner/publikasjon/1997/publikasjon1997_10.pdf
- O'callaghan, F.J., Mark, M. D. (1984). The Extraction of Drainage Networks from Digital Elevation Models. *Computer vision, graphics, and image processing*, 28, s. 323-344
- Olaya, V. (2009). Basic land-surface parameters. I: Hengl, T., Reuter, H.I. (red.), *Geomorphometry: Concepts, Software, Applications*. Developments in Soil Science, 33. Elsevier, s.141-169.
- Olivera, F., Furnans, J., Maidment, D., Djokic, D., Ye, Zichuan. (2002). Drainage systems. I: Maidment, R. D. (red.) *Arc Hydro. GIS for Water Resources*. Redlands, California, ESRI Press.
- Pereira, M., J. (1989). *A spatial approach to statistical habitat suitability modeling: The Mt. Graham red squirrel case study*. Doktoravhandling. The University of Arizona, Tuscon. Hentet 29.04.2015 fra:
http://arizona.openrepository.com/arizona/bitstream/10150/184821/1/azu_td_9004972_sip1_m.pdf
- Pike, R.J., Evans, I.S., Hengl, T. (2008). Geomorphometry: a Brief Guide. I: Hengl, T., Reuther, H.I. (red.). *Geomorphometry: Concepts, Software, Applications*. Developments in Soil Science, 33, Elsevier, s.1-28
- Ranhao, S., Baiping, Z., Jing, T. (2008). A Multivariate Regression Model for Predicting Precipitation in the Daqing Mountains. *Mountain Research and Development*, 28(3), s. 318-325.
- Roalkvam, J. (2014). *Identifisering av skredfarlige elver og bekker i Trondheim kommune med moderne fjernanalyse*. Masteroppgave. Geografisk Institutt, NTNU, Trondheim

- Romano, M.E. 2004. Innovation in LiDAR processing technology. *Photogrammetric Engineering and Remote Sensing* 70, s. 1202–1206
- Rød, J.K. (2015). *GIS – et verktøy for å analysere verden*. Kunnskapsforlaget. Foreløpig manuskript
- Saglie, L. K., Harvold, K. (2010). Arealplanlegging. I Vogelsang, C., Tønnesen, A., Saglie, I.L., Nenseth, V., Kasa, S., Harvold, K. (red.), *Ansvar og virkemidler ved tilpasning til klimaendringer* (s. 29-45). Oslo: CIENS. Hentet 23.04.2015 fra https://www.regjeringen.no/globalassets/upload/md/vedlegg/klima/klimatilpasning/underlagsrapporter/ciens_rapport.pdf
- Statens Kartverk. (2014). *Terrengmodeller – land*. Hentet 22.02.2015 fra: <http://kartverket.no/Kart/Kartdata/Terrengmodeller/Terrengmodell-10-meters-grid/>
- Statens Kartverk. (2015). *Oversikt over nedlastbare FKB-data i Stavanger kommune*. Hentet 05.05.2015 fra: <http://159.162.103.4/metadata/fkbprod.jsp?fylke=11&komm=1103&prod=FKB&Submit1=Vis+metadata>
- Stavanger Aftenblad. (2012). *35 kjellere måtte pumpes*. Hentet 10.05.2015 fra: <http://www.aftenbladet.no/nyheter/lokalt/stavanger/35-kjellere-matte-pumpes-3007226.html>
- Stoker, J.M., Greenlee, S.K., Gesch, D.B., Menig, J.C. (2006). CLICK: the new USGS center for LiDAR information coordination and knowledge. *Photogrammetric Engineering and Remote Sensing*, 72(6), s.613-616
- Su, J., Bork, E. (2006). Influence of vegetation, slope and lidar sampling angle on DEM accuracy. *Photogrammetric Engineering and Remote Sensing*, 72(11), s. 1265-1274
- Tarboton, D.G. (1997). A new method for determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, 33(2), s. 309-319.
- Thompson, A.J., Bell, C.J., Butler, A.C. (2001). Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modelling. *Geoderma*, 100, s. 67-89.

VVS Företagan. (2005). Vattenskadeundersökningen. Hentet 10.05.2015 fra:

http://www.hwq.se/data/Vattenskadeundersokningen_2005.pdf

8. Vedlegg

Tabell 7: Tabellen er hentet fra Brevik m.fl (2015). Egne modifikasjoner er gjort i tabellen hvor Stavanger Kommune har kodet annerledes.

Feltnavn	Datatype	Forklaring	Kommentar
Skadenummer	Tekst (17)	Tellevariabel	Ikke unik ID. Kan være opp til tre like nummer. For unik ID: skill på «Skadetype» i tillegg
Kommunennummer	Tekst (4)	Skadestedets kommunenummer	
Postnr	Heltall Short (4)	Skadestedets postnummer	Lite utfyllt. De fleste er '9999' som betyr ukjent
Bnr/gnr	Heltall Short (4)	Skadestedets gårds- og bruksnummer	Lite utfyllt. De fleste er '9999'
Skadeadresse	Tekst (27)	Skadestedets adresse, vei- /gatenavn, nummer og bokstav	Vasket grunding i FME. Store bokstaver og lite feil
Skdato	Heltall Long (8)	Inntruffet dato for skaden	På formatet YYYYMMDD, gjør det enklere å sortere
Bransje	Tekst (5)	Kode for type forsikring / forsikringsprodukt	111 = Hjem/løsøre 112 = Villa (privatbolig/enebolig) 113 = Hytte/fritidsbolig 119 = Øvrig privat 221 = Huseier/boligblokk/sameie 222 = Næringsbygg / produksjonsbygg / industri (inkl. næringsløsøre) 223 = Landbruk (inkl. gartneri) 224 = Bedrift, butikk, kontor (næringsløsøre, innredning) 299 = Øvrig næring
Installasjon	Tekst (1)	Vannskader på installasjon (basert på "VASK")	A= Vannrør innvendig åpent B = Vannrør innvendig skjult C = Avløp innvendig åpent D = Avløp innvendig skjult E = Våtrom (følgeskader) F = Vaskemaskin, oppvaskmaskin og beholder G = Utvendig vann- og avløpsanlegg, H = Vanninntrengning

			utenfra gjennom grunn I = Vanninntrengning utenfra over grunn J = Varmeanlegg, gulvvarme og radiatorer K = Sprinkleranlegg 9 = Ukjent
Skadetype	Tekst (2)	Type skade (VA, NP, AB)	VA = Skade som skyldes alle type vannskader; alt fra rørbrudd innvendig til frost og vanninntrenging utenfra NP = Naturskade (elveflom, skred, storm eller stormflo) – se felt NP_type for presisering av type naturskade (se også definisjon av ”ekte naturskade” på http://www.naturskade.no/). AB = Avbrudd; utgifter som påløper hvis skaden gjør at man ikke kan bo hjemme, eller ved bedriftsavbrudd. Kommer i tillegg til vannog/eller naturskader.
Kilde	Tekst (1)	Skade pga vannlekkasje fra ulike kilder (basert på ”VASK”)	Lekkasje fra: A = Metallrør (kobber, støpejern, stål) B = Plastrør. C = Støpte rør (betong, keramikk) D = Rørdel/skjøt, kupling, sluk E = Vanntilkoplet maskin F = Varmtvannsbereder G = Anlegg for romtemperaturregulering H = Sanitærutstyr og vaskekum, armatur I = Nedbør. Smeltevann. Grunnvann J = Søl, kondens og dusjing 9 = Ukjent
Aarsak	Tekst (1)	Årsak til vannskader (basert på ”VASK”)	Årsak: A = Produktfeil B = Prosjekteringsfeil C = Håndverkerfeil D = Brukerfeil. Uhell E = Slitasje og elde (normalt > 30 år) F = Lokal korrosjon G = Stopp i avløp. Tilbakeslag H = Frost I = Ytre påvirkning J = Drenering 9 = Ukjent

NP_type	Heltall Short (1)	Presisering av naturskadetype (skade dekkes av Norsk Naturskadepool) (basert på "NASK")	1 = Storm 2 = Stormflo 3 = Flom 4 = Skred 5 = Jordskjelv 6 = Vulkanutbrudd 9 = Ukjent
Erstatning	Heltall Long (7)	Fastsatt erstatning i hele kr	(betalt + gjenstående reserve) – EX regress (innbetaling eller forventet innbetaling) -999 999 for manglende data
Fsum	Heltall Long (9)	Forsikringssum i mill kr gjeldende på skadedato.	Forsikringssummen skal være lik med brannforsikringssummen som også brukes som grunnlag for innberetningen til Norsk Naturskadepool. (Det vil si at ev. avbruddsdekning og lignende ikke skal være med). Kan ha store unøyaktigheter da enhet varierer fra kr. til mill. kr. -9999 for manglende data (men også flere tilfeller av '0')