

Extremvärdesanalys av havsvattennivån i Skanör

Viktor Karlberg

VT 21

Magisteruppsats i statistik, STAN40 (15 hp)

Statistiska institutionen

Lunds universitet

Handledare: Peter Gustafsson

Examinator: Björn Holmquist

Abstract

In May 2020 the Land and Environment Court concluded that Vellinge municipality are granted the right to build a total of 21 kilometers of walls and embankments for protection against a rising sea water level. To analyse the problems that rising sea levels cause, an Extreme Value approach is applied to sea level data from the area. Both Block Maxima and Peaks over Threshold models are fitted and return levels are calculated. A return period estimate of the 1872 storm, Backafloeden, is also determined and analysed for both models. Both models seems fairly reasonable with Peaks over Threshold being the superior one.

Innehållsförteckning

1	Introduktion	4
2	Teori	5
2.1	Extremvärdesteori	5
2.2	Block Maxima	6
2.2.1	Återkomstnivå för Block Maxima	6
2.3	Toppar över Tröskelvärde	7
2.3.1	Val av tröskelvärde	8
2.3.2	Avklustring	8
2.3.3	Återkomstnivå för Toppar över Tröskelvärde	9
3	Data	10
4	Resultat	11
4.1	Block Maxima	11
4.2	Toppar över Tröskelvärde	13
5	Slutsats	16
6	Referenser	17

1 Introduktion

I maj 2020 meddelade mark- och miljödomstolen att Vellinge kommun får rätt att bygga totalt 21 kilometer murar och vallar som skydd mot en stigande havsnivå kring Falsterbohalvön. Beslutet har mötts av kritik, av dem som menar att en vall kan förstöra ekologin i området, men också av dem som är oroliga för golfbanor i området. Just Falsterbohalvön är extra känsligt för de stigande vattennivåerna på grund av en kombination av låg landhöjning och snabbt stigande havsnivå.

Sveriges meteorologiska och hydrologiska institut (SMHI) har försökt prognosticera medelhavsnivåer i hela Sverige, och de visar att havsnivån vid Skanör kan komma att höjas med 30-40 centimeter till 2050, och en meter till år 2100. Detta blir problematiskt då 95% av bostadsbyggelsen i Skanör ligger på mark som befinner sig mindre än 3 meter över medelhavsnivån. Vellinge kommun menar också att det närmast liggande alternativet till vallarna är en så kallad reträtt, det vill säga att flytta all bebyggelse till mark som ligger högre upp (Kihlberg, 2020).

I den här uppsatsen appliceras en extremvärdesanalys av havsvattenståndet vid Skanör. Både Block Maxima och Toppar över Tröskelvärde anpassas och den 100-åriga återkomstnivån kalkyleras. En beräkning av återkomstperioden för den största uppmätta stormfloden görs också.

2 Teori

2.1 Extremvärdesteori

Extremvärdesteori är en statistisk metod som under de senaste åren har blivit allt vanligare, särskilt inom klimatrelaterade områden där långa tidsserier är vanliga. Låt X_1, \dots, X_n vara en sekvens av oberoende och likafördelade slumpvariabler med fördelningsfunktion $F_{X_i}(x)$.¹ Inom extremvärdesmodellering är X_i vanligtvis värden av en process som har mätts regelbundet under en viss tid. I mitt fall är X_i timvisa havsnivåer. Maximum av dessa observationer, hädanefter benämnt M_n , kommer då att vara maximum av n tidsperioder. Om n är antalet observationer under ett år kommer M_n att vara det årliga maximum. M_n 's teoretiska fördelning är enkel att hitta.

$$F_{M_n}(z) = P(M_n \leq z) = P(X_1 \leq z, \dots, X_n \leq z) = P(X_1 \leq z) \cdot \dots \cdot P(X_n \leq z) = F_{X_i}(z)^n.$$

I praktiken kan dock detta samband inte användas, då fördelningsfunktionen $F_{X_i}(x)$ är okänd. För att lösa detta problem tillämpas en normalisering av M_n : $M_n^* = \frac{M_n - b_n}{a_n}$ där ($a_n > 0$) och b_n är sekvenser av konstanter. Detta innebär att det i praktiken är gränsvärdesfördelningen för M_n^* som är intressant, och inte den svårhanterade M_n . Nedanstående del är från Coles (2001), för en mer ingående version av det som presenteras hänvisas läsaren dit.

Teorem 1 (Fisher-Tippet-Gnedenko)

Om det existerar en sekvens av konstanter ($a_n > 0$) och (b_n) så att

$$P\left(\frac{M_n - b_n}{a_n} \leq z\right) \rightarrow G(z) \text{ när } n \rightarrow \infty$$

där $G(z)$ är en icke-degenererad fördelningsfunktion, kommer G att tillhöra en av följande familjer av fördelningar:

$$I(\text{Gumbel}) : G(z) = \exp[-\exp(-\frac{z-b}{a})], -\infty < z < \infty$$

$$II(\text{Frechét}) : G(z) = \begin{cases} 0, & z \leq b \\ \exp[-(\frac{z-b}{a})^{-\gamma}], & z > b \end{cases}$$

$$III(\text{Omvänd Weibull}) : G(z) = \begin{cases} \exp[-(\frac{z-b}{a})^\gamma], & z < b \\ 1, & z \geq b \end{cases}$$

för parametrar $a > 0$, b och i II och III, $\gamma > 0$.

¹För de svenska termerna har Enger et al. (2008) samt Statistiska Föreningen (1954) använts. Där inga översättningar har hittats, har författarens egna använts.

De tre ovan definierade fördelningarna kallas för extremvärdesfördelningar. Gumbel, Frechét och Omvänd Weibull kan alla parametreras till en familj med en fördelningsfunktion som nedan:

$$G(z) = \exp\left\{-\left[1 + \gamma\left(\frac{z - \mu}{\sigma}\right)\right]^{-1/\gamma}\right\} \quad (1)$$

Denna fördelning kallas för en generaliserad extremvärdesfördelning (GEV).

GEV-modellen har tre parametrar, μ ; lägesparametern ($-\infty < \mu < \infty$), σ ; skalparametern ($\sigma > 0$) och γ ; formparametern ($-\infty < \gamma < \infty$). Typ II och III av extremvärdesfördelningarna motsvarar fallen när $\gamma > 0$, samt $\gamma < 0$. När $\gamma = 0$ ges gränsvärdesfördelningen av formeln ovan när $\gamma \rightarrow 0$ vilket kommer att leda till en extremvärdesfördelning av typ I, Gumbelfördelningen.

Vilken av modellerna som appliceras är beroende av vilken data som ska undersökas.

De två vanligaste sätten att modellera extremvärden tas upp i nästa avsnitt.

2.2 Block Maxima

Block Maxima utnyttjar Teorem 1 och använder fördelningen för maxima. En blockstorlek väljs, vanligtvis till ett år, och den största observationen i varje block anpassas sedan till GEV. De årliga maxima analyseras sedan och de olika parametrarna maximum-likelihood skattas. Block Maxima metoden anses vara ett slöseri av data, detta då endast den största observationen i varje block är med i analysen. Även om den näst största observationen i ett block kan vara större än den största observationen i nästa block kommer i Block Maxima analysen endast den senare vara inkluderad (Beirlant, 2004).

2.2.1 Återkomstnivå för Block Maxima

Efter att en extremvärdesmodell har anpassats till data är det vanligt att analysera återkomstnivån för olika extrema händelser. För årliga maxima kan fördelningens extrema kvantiler hittas genom inversfunktionen av uttrycket (1)

$$z_p = \begin{cases} \mu - \frac{\sigma}{\gamma}[1 - (-\ln(1-p))^{-\gamma}], & \gamma \neq 0, \\ \mu - \sigma \ln[-\ln(1-p)], & \gamma = 0, \end{cases} \quad (2)$$

där $G(z_p) = 1 - p$.

Nivån z_p kallas för återkomstnivån associerad med återkomstperioden $1/p$. Nivån z_p förväntas att överskridas i genomsnitt en gång varje $1/p$ år. Nivån z_p överskrids av det årliga maximat i ett godtyckligt år med sannolikhet p . Vanligtvis är det långa återkomstperioder, vilket innebär att det är små värden på p som är av intresse (Coles, 2001).

2.3 Toppar över Tröskelvärde

Toppar över Tröskelvärde (Töt) är ett annat vanligt sätt att modellera extrema händelser och anses mer effektiv än Block Maxima då den tar hänsyn till fler (stora) värden. Enkelt förklarar inkluderar Töt alla värden som är större än ett fast tröskelvärde, så kallade överskridningar. Just valet av tröskelvärdet är den största utmaningen vid Töt-modellering. En ställs inför ett bias/varians-dilemma där ett för högt tröskelvärde kommer att ge få observationer i modellen och således en hög varians, medan ett för lågt tröskelvärde kommer att leda till att icke-extrema värden inkluderas i analysen vilket medför en bias. Hädanefter kommer tröskelvärdet att betecknas u .

Låt X_1, X_2, \dots vara en sekvens av oberoende och identiskt fördelade slumpvariabler, med marginalfördelning $F_{X_i}(x)$. Definiera sedan en extrem händelse som en händelse där X är större än ett högt tröskelvärde, u . Fördelningen för ett godtyckligt X ur sekvensen X_i ges då av formeln:

$$P(X > u + y | X > u) = \frac{1 - F_X(u + y)}{1 - F_X(u)}, y > 0. \quad (3)$$

Återigen är den teoretiska fördelningen inte särskilt användbar då $F(x) = F_{X_i}(x)$ är okänd. Precis som i Block Maxima appliceras istället en approximation som är generell för höga tröskelvärderna. Nedanstående härledning är från Coles (2001).

Teorem 2 *Låt X_1, X_2, \dots vara en sekvens av oberoende och identiskt fördelade slumpvariabler, med marginalfördelning F och låt*

$$M_n = \max(X_1, \dots, X_n)$$

Låt X vara en godtycklig term i sekvensen X_i , och anta att F uppfyller Teorem 1, så att för stora n ,

$$P\left(\frac{M_n - b_n}{a_n} \leq z\right) \approx G(z)$$

där

$$G(z) = \exp\left\{-\left[1 + \gamma\left(\frac{z - \mu}{\sigma}\right)\right]^{-1/\gamma}\right\}$$

för $\mu, \sigma > 0$ och γ .

För ett u som är stort nog, kommer $(X - u)$ givet att $X > u$ att följa fördelningen:

$$H(y) = 1 - \left(1 + \frac{\gamma y}{\tilde{\sigma}}\right)^{-1/\gamma}, \{y > 0 \text{ och } (1 + \gamma y/\tilde{\sigma}) > 0\} \quad (4)$$

där

$$\tilde{\sigma} = \sigma + \gamma(u - \mu) \quad (5)$$

Familjen av fördelningar som definieras av uttrycket i (4) kallas för generaliserad paretofördelning (GPD).

2.3.1 Val av tröskelvärde

Som tidigare nämnts är det största problemet vid Töt-modellering valet av tröskelvärde. En av de vanligaste metoderna för val av tröskelvärde, och något som har visat sig framgångsrikt, är en grafisk metod som diskuteras av Scarrot och MacDonald (2012) vid namn *Parameter Stability Plot*². Enligt Teorem 2: om en generaliserad Paretofördelning är en lämplig modell för överskridningar av ett tröskelvärde u_0 , kommer överskridningar av ett högre tröskelvärde u också att följa en generaliserad Paretofördelning. Formparametern i de båda fördelningarna är identisk. Låt σ_u vara skalparametern för ett tröskelvärde $u > u_0$. Enligt ekvation (5) kommer

$$\sigma_u = \sigma_{u_0} + \gamma(u - u_0) \quad (6)$$

Således kommer σ_u att variera med u , om inte $\gamma = 0$. Detta problem kan lösas med hjälp av en omparametrisering av skalparametern

$$\sigma^* = \sigma_u - \gamma u,$$

vilken nu är konstant med avseende på u enligt ekvation (6). Följdaktligen bör skattningar av γ och σ^* vara konstanta för tröskelvärden större än u_0 , om u_0 är ett lämpligt tröskelvärde för att överskridningarna ska följa en generaliserad Paretofördelning. På grund av stickprovsvarians och olika stickprovsstorlekar kommer skattningarna att skiljas lite åt, dock bör de ligga inom medelfelet. Den här metoden innebär att γ och σ^* plottas mot u och u_0 väljs till det lägsta värdet för u där skattningarna är så nära konstanta.

2.3.2 Avklustring

Som tidigare nämnts utgår Töt-metoden ifrån att observationerna är oberoende. Detta stämmer dock inte i mitt fall då havsnivån en viss timme är starkt beroende av havsnivån de föregående timmarna. Detta innebär att höga havsnivåer förväntas komma i kluster och därmed är inte antagandet om oberoende korrekt. Det vanligaste tillvägagångssättet för att ändå kunna applicera Töt är en så kallad avklustring (Beirlant, 2004).

Idén grundar sig i att filtrera de beroende observationerna för att få en uppsättning av tröskelöverskridningar som är approximativt oberoende. Det här görs genom att först välja ett tröskelvärde u , och sedan låta överskridningar på varandra följande tillhöra samma kluster. Klustret är aktivt till r på varandra följande observationer är under u . Sedan upprepas processen tills all data har blivit analyserad. Den största observationen i varje kluster identifieras sedan, och GPD anpassas på dessa kluster-maxima som antas vara oberoende (Coles, 2001; Beirlant, 2004).

²Scarrot och MacDonald benämner metoden som *Threshold Stability Plot*.

Valet av r är även det ett varians-bias dilemma där ett för litet r kommer leda till att antagandet om oberoende inte uppfylls, medan ett för stort r kommer att innebära att kluster som kan antas vara oberoende slås ihop, och därmed förloras värdefulla data. För havsvattennivå har tidigare studier visat att stormfloder som uppstår med 48 timmars mellanrum kan antas oberoende, och detta är något som kommer att tas i beaktning vid modellering (Fredriksson et al, 2016).

2.3.3 Återkomstnivå för Toppar över Tröskelvärde

Det är också möjligt att beräkna återkomstnivå för tröskelmodeller, men metoden är lite anorlunda (Coles, 2001; Gilleland och Katz, 2016). Antag att en generaliserad paretofördelning med parametrar σ och γ är en lämplig modell för överskridningar över ett tröskelvärde. Det vill säga, för $x > u$,

$$P(X > x | X > u) = [1 + \gamma(\frac{x - u}{\sigma})]^{-1/\gamma}$$

där $\gamma \neq 0$. Låt sedan $\zeta_u = P(X > u)$. Ovanstående ekvation kan då förenklas till

$$P(X > x) = \zeta_u [1 + \gamma(\frac{x - u}{\sigma})]^{-1/\gamma} \text{ för } x > u.$$

Således är återkomstnivån x_m , som överskrids i genomsnitt en gång var m :te observation lösningen av

$$\zeta_u [1 + \gamma(\frac{x_m - u}{\sigma})]^{-1/\gamma} = \frac{1}{m}$$

vilket efter omskrivning ger

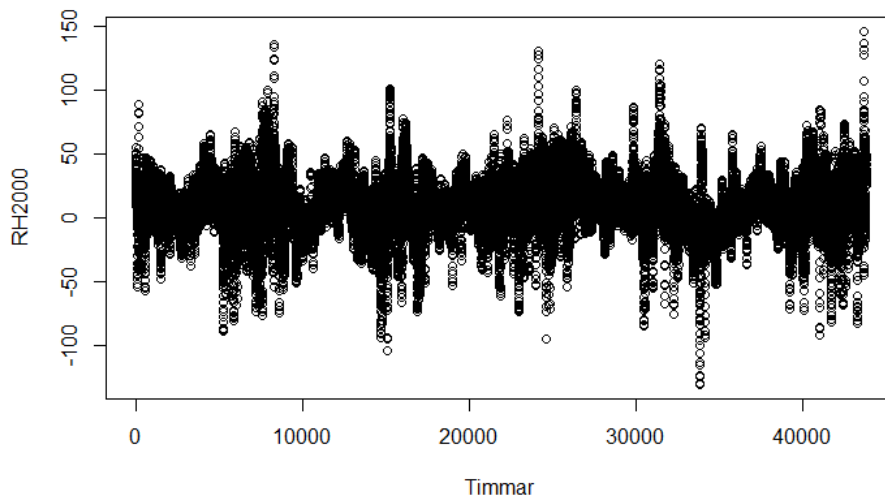
$$x_m = \begin{cases} u + \frac{\sigma}{\gamma} [(m\zeta_u)^\gamma - 1] & \gamma \neq 0 \\ u + \sigma \ln(m\zeta_u) & \gamma = 0 \end{cases}$$

givet att m är tillräckligt stort, så $x_m > u$. Sannolikheten ζ_u skattas med hjälp av data som andelen observationer som överskrider tröskelvärdet u . Om k är antalet observationer som är större än tröskelvärdet och n är totala antalet observationer så kommer $\zeta_u = \frac{k}{n}$. För att estimeras N-åriga återkomstnivån väljs m till antal observationer under N år. För timvisa observationer kommer m vara produkten av återkomstperioden, antal dagar i ett år och antal timmar på ett dygn. Till exempel, för den 100-åriga återkomstnivån blir $m = 100 \cdot 365.25 \cdot 24$.

3 Data

Datamaterialet består av havsvattenstånd vid mätstationen i Skanör hämtat från SMHIs (n.d) öppna data. Stationen började mäta 1992 och data hämtades i maj 2020. Mätningarna består av timvisa observationer av havsvattenståndet. Det finns saknade värden, ofta i följd. Det här är troligtvis för att mätinstrumentet har gått sönder och det har tagit tid för reparatören att åtgärda problemet. Antal saknade värden är dock få, så de anses inte vara ett stort problem vid analysen. I figur 1 visas data för de första fem åren. RH2000 är Rikets Höjdsystem 2000, och är det enhetliga referenssystemet för svenska havsvattenstånd. Kort förklarar utgörs det av en fast punkt på land, som är nollpunkten i systemet. Den här nollpunkten är densamma över hela landet, och havsvattenståndet mäts utifrån den här nollpunkten (SMHI, 2019).

Som tidigare nämnts är det ofta av intresse att analysera en historisk extremhändelse, och försöka kalkylera återkomstperioder för dessa händelser. 1872 skedde den, till dags dato, största dokumenterade stormfloden i Östersjön. Stormen, som fick namnet Backafloden, krävde 271 människoliv, gjorde 15000 personer hemlösa samt förstörde 2800 byggnader i Sverige, Danmark och Tyskland (Feuchter et al, 2013). Då mätinstrumenten var begränsade råder det skilda meningar om hur högt vattnet stod, men tidigare studier menar på att i Skanör var havsvattenståndet cirka 260 cm i RH2000 (Fredriksson et al, 2016).



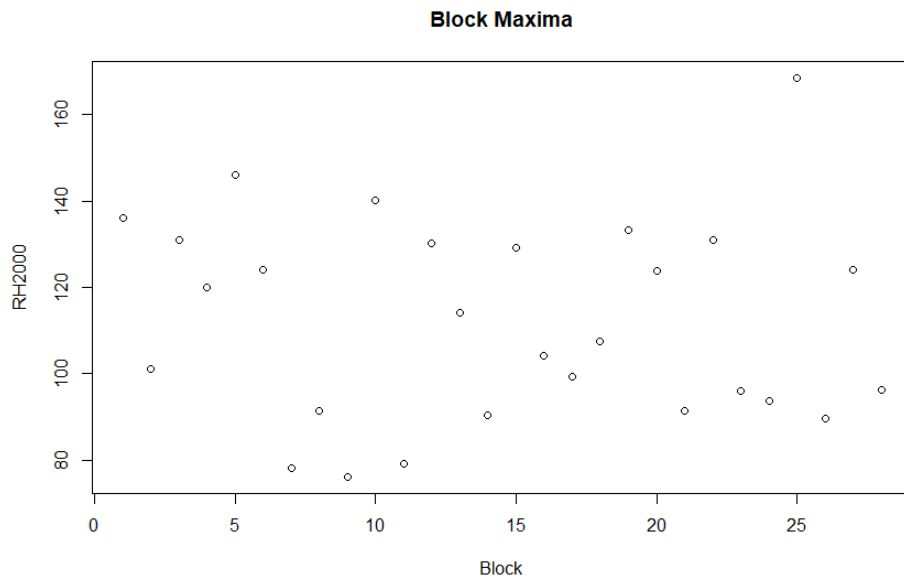
Figur 1: *Fem första åren av observationer*

4 Resultat

Analysen har gjorts i R med paketet `extRemes` (Gilleland och Katz, 2016).

4.1 Block Maxima

Block Maxima-analysen har genomförts på de årliga högsta havsvattenstånden mellan 1992-2020. Således bygger analysen på 28 observationer. Eftersom det finns en säsongseffekt måste blockstorleken väljas så att effekten inte bryter mot antaganden som modellen har. Årliga block anses bäst lämpade för att Teorem 1 ska uppfyllas, och har blivit praxis i området. Eftersom havsvattenståndet är som högst under vinterhalvåret finns det ett problem med att dela upp blocken efter kalenderår. Detta då det finns en risk att den största observationen ett år inträffar i december, och den största observationen året efter sker i januari. Det kan då bara vara ett par dagar som skiljer dessa observationer åt, och observationerna kommer inte att vara oberoende. För att minimera den här risken har jag delat upp mina block så att den första dagen varje år är 1 juli och sista dagen varje år är 31 juni. Detta är något som förespråkas av Fredriksson et al (2016). En plot av de årliga maxima som analyseras visas i figur 2.



Figur 2: *Årliga maxima*

Både GEV-modeller och Gumbel-fördelningen anpassades och jämfördes med varandra. Då de olika valideringsplottarna var snarlika gjordes ett Likelihood-kvottest, något som är en vanligt förekommande statistisk metod i ämnesområdet, för att se om den ena modellen

föredrogs före den andra. Då testet var icke signifikant valdes Gumbel-modellen som den bättre modellen, på grund av att den anses enklare.

Gumbel	Estimat
Lägesparameter, μ	101.33 (3.93) ³
Skalparameter, σ	19.67 (2.90) ⁴
Formparameter, γ	0
100-årig återkomstnivå	191 (162-221) ⁵
Återkomstperiod för Backafloden	3200 (415 - 73000) år ⁶

Tabell 1: *Maximum Likelihood-estimat för Gumbelmodellen, 100-årig återkomstnivå för modellen samt återkomstperioden i antal år för Backafloden.*

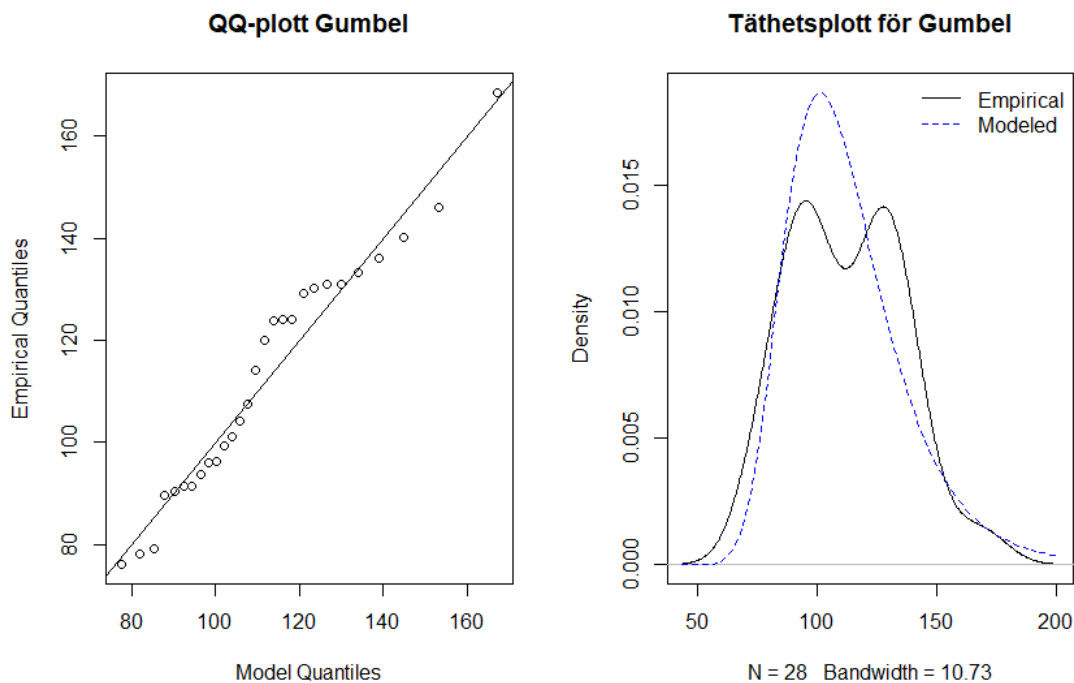
Det relativt breda konfidensintervallet på den 100-åriga återkomstnivån är inte särskilt förvånande, då modellen endast bygger på 28 observationer. Extremvärdesmodeller har visat sig framgångsrika på prognoser genom extrapolering, men för att få ett mer precist resultat för återkomstnivån behövs mer data. Att så få observationer är inkluderade i modellen syns också i de två diagnostikplottarna i figur 3.

³Medelfel inom parentes

⁴Medelfel inom parentes

⁵Normalapproximerade konfidensintervall inom parentes

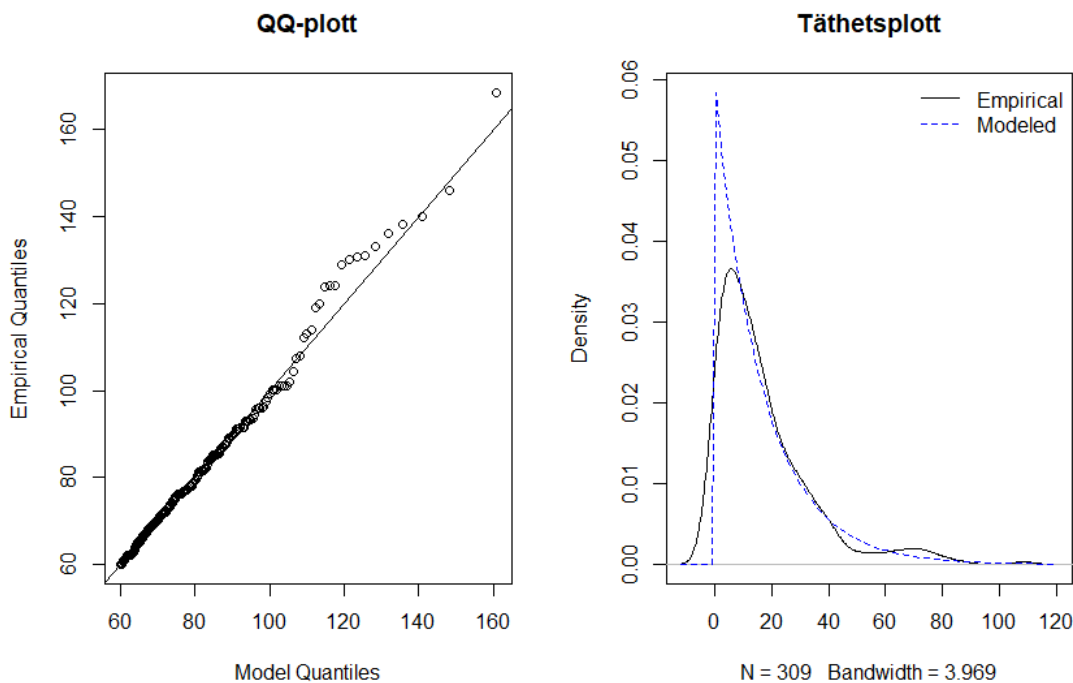
⁶Profil-likelihood konfidensintervall inom parentes



Figur 3: Diagnostikplottar för Gumbelmodellen

4.2 Toppar över Tröskelvärde

Då Toppar över Tröskelvärde använder sig av mer data i analysen än Block Maxima ser modellen bättre ut. Diagnostikplottarna som används för att validera modellen visar i figur 4 på en godtagbar anpassning till mätningarna av havsvattennivån. Tröskelvärdet valdes efter den metod som nämns i sektion 2.3.1, där ett tröskelvärde på 60 cm i RH2000 verkade mest lämpligt. Detta då 60 var det lägsta värdet där parametrarna var konstanta även om tröskelvärdet höjdes. Avklustring, vilket diskuterades i sektion 2.3.2, applicerades på dessa data varvid r valdes till 48 timmar.



Figur 4: Diagnostikplottar för Töt; $u=60$, $r=48$.

Generaliserad Pareto	Estimat
Skalparameter	16.92 (1.39) ⁷
Formparameter	0.013 (0.05) ⁸
100-årig återkomstnivå	184 (160-208) ⁹
Återkomstperiod för Backafloden	4500 (700 - 59732) år ¹⁰

Tabell 2: *Maximum Likelihood-estimat för generaliserade paretomodellen, 100-årig återkomstnivå för modellen samt återkomstperioden i antal år för Backafloden.*

Estimaten för den 100-åriga återkomstnivån för de båda modellerna är någorlunda lika, något som indikerar på stabilitet i modellerna. Tröskelmodellen har dock något smalare konfidensintervall, och är troligtvis mer tillförlitlig baserat på analysen av diagnostikplottarna.

Backafloden är av allt att döma en väldigt extrem händelse. En ska dock vara försiktig med att tolka dessa resultat allt för ingående. Resultaten från modellerna är inte konsistenta, samt

⁷Medelfel inom parentes

⁸Medelfel inom parentes

⁹Normalapproximerade konfidensintervall inom parentes

¹⁰Profil-likelihood konfidensintervall inom parentes

den långa återkomstperioden och de breda konfidensintervallen tyder på att modellerna inte lyckas att modellera denna händelse särskilt väl. Resultaten är inte särskilt förvånande då det är ett synnerligen grovt antagande att de senaste 28 årens extremvärden och ett extremt högt vattenstånd för nästan 150 år sedan skulle följa samma fördelning.

5 Slutsats

Extremvärdesanalys har visat sig vara en god modell för att analysera havsvattenståndet i Skanör. Särskilt tröskelmodellen gav bra diagnostiska resultat och verkar således vara en helt rimlig modell för att analysera havsvattenståndet i Skanör. Analysen blir dock lidande av den relativt korta tidsserien som finns tillgänglig. Det här går, tyvärr, inte att lösa på annat sätt än att vänta på att åren ska gå och således få fler observationer att bygga modeller på. Det här belyser ett av problemen vid extremvärdesmodellering. Extrema händelser är, per definition, få och det innebär att de olika modellestimaten, särskilt extrema återkomstnivåer kommer att ha en stor varians.

För att förstärka insikten om modelleringen skulle man kunna studera extremvärdesfördelningarna för angränsande mätpunkter, både för min- och max-värden av havsvattenstånden. Detta skulle innebära att modellerna får en högre trovärdighet, och även mer pålitliga resultat.

En framtida analys av havsvattenståndet hade även kunnat inkludera klimatförändringar i modellen. Det här var något jag inte hade möjlighet till men när jag försökte inkludera olika former av trender i analysen var de dock alla insignifikanta. Troligtvis på grund av att tidsserien inte är tillräckligt lång. En trend av stigande extrema havsvattennivåer är inte säkert uppenbara i data om 28-års mätningar trots att vi vet att den existerar. Detta ger problem, främst vid skattning av återkomstnivån. Även om ingen trend syns på 28 år, kommer den högst troligt att synas inom 100 år eller längre. Alltså existerar det en bias för återkomstnivåer, och dessa nivåer är högst troligt underskattade. Det är alltså av intresse för framtida studier att försöka lösa det här modelleringsproblemet och på så sätt kunna analysera framtida havsvattenstånd där klimatförändringar har inkluderats.

Ett alternativ till att undersöka om de extrema händelserna har blivit mer extrema kan vara att studera om de extrema händelserna har blivit mer vanliga. Det vill säga, sker det fler extremhändelser idag än vad det gjorde tidigare?

När vallen i Skanör står klar och det är klarlagt hur högt den kommer stå, vore det såklart av intresse att analysera hur havsvattenståndet kommer att förhålla sig till den samt beräkna sannolikheter för översvämning. Detta var något jag åsyftade att göra, men olika förhandlingar och domstolsprocesser har dragit ut på tiden och för närvarande finns inget beslut som indikerar på hur hög vallen är tänkt att bli.

6 Referenser

Beirlant, J. (red.) (2004). *Statistics of Extremes: Applications and Theory*, Hoboken, N.J.: Wiley.

Coles, S. (2001). *An introduction to Statistical Modeling of Extreme Values*, London: Springer.

Enger, J., Huitfeldt, B., Jorner, U., Wretman, J. (2008). Swedish translation for the ISI Multilingual Glossary of Statistical Terms. Tillgänglig online: statistikframjandet.se/etisk-kod/ordlista/

Feuchter, D., Jörg, C., Rosenhagen, G., Auchmann, R., Martius, O., Brönnimann, S. (2013) The 1872 Baltic Sea storm surge. In: Brönnimann, S. and O. Martius (Eds.) *Weather extremes during the past 140 years. Geographica Bernensia G89*, s. 91-98

Fredriksson, C., Tajvidi, N., Hanson, H., Larson, M. (2016). Statistical Analysis of Extreme Sea Water Levels At The Falsterbo Peninsula, South Sweden. *Journal of Water Management and Research*, vol. 72, s. 129-142.

Gilleland, E., Katz, R. (2016). extRemes 2.0: An Extreme Value Analysis Package in R. *Journal of Statistical Software*, vol. 72, no. 8, s. 1-39.

Kihlberg, J. (2020). Klartecken för Vellinge kommun att bygga skyddsvallar, Dagens Nyheter, 7 maj, Tillgänglig online: <https://www.dn.se/nyheter/sverige/klartecken-for-vellinge-kommun-att-bygga-skyddsvallar> [Hämtad 22 februari 2021].

Scarrot, C., MacDonald, A. (2012). A Review of Extreme Value Threshold Estimation and Uncertainty Quantification. *REVSTAT - Statistical Journal*, vol. 10, no. 1, s. 33-60.

SMHI (2019). Nytt referenssystem för vattenstånd i havet.

Tillgänglig online: <https://www.smhi.se/nyhetsarkiv/nytt-referenssystem-for-vattenstand-i-havet-1.148346> [Hämtad 10 mars 2021].

SMHI (n.d). Ladda ner oceanografiska observationer. Tillgänglig online:

www.smhi.se/data/oceanografi/ladda-ner-oceanografiska-observationer/ [Hämtad 25 maj 2020]

Statistiska Föreningen (1954). Nordisk Statistisk Nomenklatur. Engelsk-Nordisk och Svensk-Engelsk ordlista, Köpenhamn.