Cynrychioliad amharamedrig ar gyfer cyd-newidynnau amlddimensiynol mewn model gwerthoedd eithaf, â chymhwysiad eigionegol

Philip Jonathan[a]

[a] *Adran Fathemateg ac Ystadegaeth, Prifysgol Caerhirfryn, Caerhirfryn LA1 4YF*

p.jonathan@lancaster.ac.uk

Crynodeb: Cyflwynir methodoleg ystadegol er mwyn modelu gwerthoedd eithaf amgylcheddol prosesau anunfan. Seilir y fethodoleg ar fodel Pareto cyffredinoledig ar gyfer brigau dros drothwy o'r broses amgylcheddol, â chynrychioliad Voronoi ar gyfer amrywiad paramedrau'r model gwerthoedd eithaf gyda chyd-newidynnau amlddimensiynol. Defnyddir rhesymu Bayesaidd MCMC naid wrthdroadwy, yn ymgorffori samplu Metropolis-Hastings mewn Gibbs, i amcangyfrif cyd-ôl-ddosraniad holl baramedrau'r cynrychioliad Voronoi. Cymhwysir y fethodoleg i ganfod nodweddion gerwinder stormydd morol eithafol gyda chyfeiriad a thymor. Dilysir bod efelychiadau yn ôl y model a amcangyfrifwyd yn cyfateb yn dda i'r data gwreiddiol. Ymhellach, defnyddir y model i amcangyfrif uchafwerthoedd brigau dros drothwy sy'n cyfateb i gyfnodau dychwelyd llawer hwy na chyfnod y data gwreiddiol.

Allweddeiriau: gwerthoedd eithaf; cyd-newidyn amlddimensiynol; ymraniad Voronoi; rhesymu Bayesaidd; samplu Gibbs a Metropolis-Hastings; samplu MCMC naid wrthdroadwy; tonnau morol; gwerth dychwelyd;

Cyflwynwyd i Gwerddon, Haf 2021

A non-parametric representation for multi-dimensional covariates in an extreme value model, with oceanographic application

Abstract: *A statistical methodology is presented to model extreme values from non-stationary environmental processes. The methodology is based on a generalized Pareto model for peaks over threshold of the environmental process combined with a Voronoi representation for the variation of extreme value model parameters with multi-dimensional covariates. Bayesian inference using reversible-jump MCMC, incorporating Metropolis-Hastings within Gibbs sampling, is used to estimate the joint posterior distribution of all parameters of the Voronoi representation. The methodology is applied to characterise extreme ocean storm severity with direction and season. The fitted model is validated by comparing the characteristics of data simulated under the model with those of the original sample data. Further, the model is used to estimate the distribution of maxima of peaks over threshold corresponding to return periods much longer than the period of the original data.*

Keywords: extreme value; multi-dimensional covariate; Voronoi partition; Bayesian inference; Gibbs sampling and Metropolis-Hastings; reversible-jump MCMC; ocean wave; return value.

1 Rhagymadrodd

Mae canfod nodweddion digwyddiadau anaml neu eithafol yn bwysig mewn meysydd gwyddonol, diwydiannol, ariannol ac amgylcheddol. Er enghraifft, mae diogelwch rhai sy'n byw mewn ardaloedd arfordirol neu'n gweithio ar y môr yn dibynnu ar fodolaeth adeileddau megis morgloddiau, llongau a llwyfannau olew a all wrthsefyll stormydd morol garw.

Yn fras, mae canllawiau dylunio (megis NORSOK N-006, 2015; ISO19901-1, 2015) yn awgrymu y dylai adeiladwaith o'r fath wrthsefyll digwyddiad eithafol â **chyfnod dychwelyd** (*return period*) o ryw gannoedd o flynyddoedd. Er mwyn bodloni'r gofynion cynllunio felly, gwaith yr eigionegydd ffisegol a'r peiriannydd morol yw amcangyfrif priodweddau'r digwyddiadau eithafol hyn yn seiliedig ar arsylwadau a damcaniaeth ystadegol gwerthoedd eithaf.

Gyda'r hapnewidyn $M \in \mathbb{R}$ yn cynrychioli uchafwerth blynyddol rhyw allbwn $X \in \mathbb{R}$ o'r broses dan sylw, diffinnir y **gwerth dychwelyd** (*return value*) $x_P \in \mathbb{R}$ â chyfnod dychwelyd $P \in \mathbb{R}_{>0}$ ar gyfer X gan y mynegiad

$$\mathbb{P}(M \leq x_P) = 1 - \frac{1}{P} \tag{1}$$

Fel arfer, ystyrir $P \in [10, 10^6]$ mlynedd; felly, **cwantil** (*quantile*) uchel o ddosraniad M yw'r gwerth dychwelyd. Er enghraifft, gyda $P = 100$, siaradir am y gwerth dychwelyd can mlynedd, neu'r gwerth can mlynedd.

Gan nad oes gennym fesuriadau o'r amgylchedd morol am gyfnodau o gannoedd o flynyddoedd, mae amcangyfrif gwerthoedd dychwelyd (yn cyfateb i gyfnod dychwelyd o filoedd o flynyddoedd) yn anodd i'w gyflawni'n ddibynadwy. Fel arfer felly, cyfunir mesuriadau morol uniongyrchol mewn model ffisegol a elwir yn **hanesolwg** (*hindcast*) i gynhyrchu cyfres amser o led-arsylwadau H_S ynghyd â newidynnau a chyd-newidynnau perthnasol, am gyfnod o ryw ugain i gan mlynedd (e.e. Reistad et al., 2011; Bloemendaal et al., 2020). Yn ddiweddar, mae **modelau cypledig byd-eang** (*global coupled models*) yn cynnig hanesolygon am gyfnodau llawer hwy, dros 500 mlynedd (e.e. Bao et al., 2020). Defnyddir y data hyn i greu modelau ystadegol o gynffonnau dosraniadau ymylol, a chyd-gynffonnau dosraniadau newidynnau amgylcheddol o ddiddordeb, fel sail i amcangyfrif cwantilau eithafol megis gwerthoedd dychwelyd.

Fel arfer defnyddir **uchder ton arwyddocaol** (*significant wave height*, e.e. Kinsman, 2012) H_S fel mesur o'r egni yn yr arwyneb morol mewn safle penodol dros gyfnod penodol τ o amser (e.e. hanner awr, neu dair awr yn hanesyddol); un diffiniad o H_S yw $4\sigma_S$, gyda σ_S yn cynrychioli gwyriad safonol yr arwyneb morol. Yn y gwaith hwn, mesurir H_S mewn metrau (m). Mae'n hysbys bod nodweddion H_S ei hun yn amrywio'n systematig â chyd-newidynnau megis cyfeiriad storm (Ewans a Jonathan, 2008), yr amser o'r flwyddyn (Jonathan ac Ewans, 2008), dyfnder y môr (Battjes a Groenendijk, 2000), a **chyrch** (*fetch*, e.e. Kinsman, 2012). Er mwyn amcangyfrif nodweddiad digonol felly, rhaid ystyried effeithiau cyd-newidynnau yn benodol mewn modelau **anunfan** (*non-stationary*). Nid yw'n bosib rhagweld sut yn union y bydd priodweddau H_S yn amrywio â chyd-newidynnau. O'r herwydd, mae cynrychioliadau hyblyg, amharametrig

ar gyfer y ddibyniaeth gyd-newidynnol yn angenrheidiol, gan ddefnyddio arsylwadau a modelu ystadegol i amcangyfrif priodweddau optimaidd y cynrychioliad mewn unrhyw gymhwysiad.

Gan mai prosesau atmosfferig graddfa eang sy'n gyfrifol am stormydd morol, mae'n amlwg e.e. bod gwerthoedd H_S mewn mannau cyfagos yn debygol o fod yn debyg i'w gilydd. Yn wir, mae gwerthoedd eithafol H_S o safleoedd cyfagos yn arddangos **dibyniaeth eithafol ofodol**. Os mai'r bwriad yw modelu H_S yn ofodol felly, rhaid hefyd ystyried y ddibyniaeth hon (e.e. Davison et al., 2012). Yn y gwaith hwn, dim ond arsylwadau o un lleoliad sydd o dan sylw, felly gellir anwybyddu effeithiau dibyniaeth ofodol.

Mewn modd tebyg, mae gwerthoedd dilynol H_S mewn amser (yn cyfateb i gyfres o **gyflyrau môr** (*sea states*) storm, pob un o hyd τ) yn debygol o fod yn debyg, oherwydd y ddibyniaeth amserol sydd hefyd yn bresennol. Er mwyn disgrifio'r ddibyniaeth amserol mewn model gwerthoedd eithaf, gellir defnyddio model hierarchaidd. Y cam cyntaf yw creu model gwerthoedd eithaf ar gyfer H_S sy'n cyfateb i gyflwr garwaf pob storm unigol, sef H_S **brig**; cymhelliad sylfaenol dros y model hierarchaidd yw bod arsylwadau H_S brig yn debygol iawn o fod yn annibynnol (yn amodol ar gyd-newidynnau). Wedyn, yn amodol ar H_S brig a chyd-newidynnau'r cyflwr brig, gellir amcangyfrif model am gynnydd a gostegiad H_S mewn storm dros gyfres o gyflyrau dilynol (e.e. Randell et al., 2015) sy'n disgrifio'r ddibyniaeth amserol. Yn y gwaith hwn, dim ond y model gwerthoedd eithaf ar gyfer H_S brig a ystyrir.

Mae'n hysbys bod nodweddion amgylcheddol morol eithafol yn gyd-ddibynnol; h.y. rhaid nodweddu priodweddau tonnau (fel H_S brig), gwyntoedd a cheryntau eithafol ar y cyd, yn hytrach na'u hystyried yn annibynnol. Yn ystadegol, golyga hyn fod yn rhaid amcangyfrif **cynffonnau cyd-ddosraniadau** tebygolrwydd hapnewidynnau amgylcheddol, nid dosraniadau ymylol y newidynnau unigol (e.e. Heffernan a Tawn, 2004). Yn y gwaith hwn, dim ond cynffon ymylol dosraniad H_S brig a fodelir.

Mae damcaniaeth gwerthoedd eithaf yn briodol fel sail i fodelu eithafon yn unig. Yn ymarferol mewn cymhwysiad, nid yw'n glir yn union sut i asesu bod arsylwad yn "eithafol". Er enghraifft, wrth fodelu **brigau dros drothwy** (*peaks over threshold*), cymerir yn ganiataol fod gwerthoedd y brigau dros drothwy yn dilyn dosraniad arbennig, sef y **Pareto cyffredinoledig** (*generalized Pareto*, GP). Mae'r rhagdybiaeth hon yn addas yn unig pan fo'r trothwy a ddefnyddir yn "uchel". Ond os gosodir trothwy yn rhy uchel, nid oes digon o arsylwadau brigau dros drothwy yn weddill i amcangyfrif paramedrau'r dosraniad GP mewn modd derbyniol; yn gyffredinol, mae ansicrwydd yr amcangyfrifon yn chwyddo wrth i'r trothwy godi. Yn amlwg, mae dewis trothwy priodol yn bwysig; mae profion diagnostig yn bodoli i hwyluso'r gwaith (e.e. Coles, 2001). Gwell fyth hwyrach yw amcangyfrif ansicrwydd casgliadau o ganlyniad i ansicrwydd trothwy, yn enwedig yng nghyswllt modelu anunfan (e.e. Northrop et al., 2017). Yn y gwaith hwn, dewisir trothwy anunfan pendant ar gyfer casgliadau, ac mae'r casgliadau yn amodol ar y dewis hwn. Nid ystyrir ansicrwydd trothwy.

Yn gyffredinol, gwaith dyrys yw nodweddu cynffon dosraniad sampl o ddata â dulliau gwerthoedd eithaf, yn enwedig pan fydd effeithiau cyd-newidynnau yn gorfodi ystyried modelau anunfan. Felly, mae meintioli ansicrwydd casgliadau yn sylfaenol bwysig er

mwyn gwneud penderfyniadau priodol yn y pen draw. Yn y cyswllt hwn, ceir manteision amlwg o wneud casgliadau ystadegol o fewn fframwaith rhesymu Bayesaidd (e.e. Coles a Powell, 1996; Randell et al., 2016).

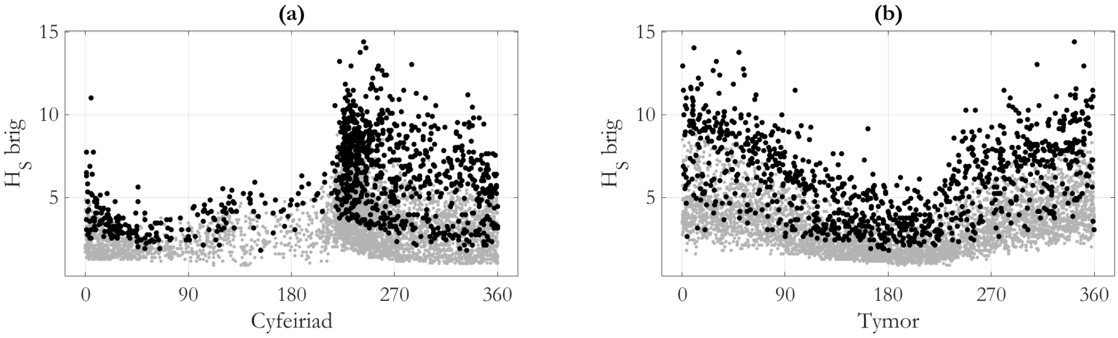
Mae cynnwys yr erthygl fel a ganlyn. Yn Adran 2, cyflwynir cymhwysiad ysgogol i'r gwaith presennol, sef canfod nodweddion eithafon H_S brig ym Môr y Gogledd gan ddefnyddio model brigau dros drothwy anunfan. Rhoddir braslun o'r cefndir mathemategol yn Adran 3, gan gynnwys disgrifiad o'r dosraniad Pareto cyffredinoledig, ymraniad Voronoi a chynrychioliadau cyd-newidynnol. Yn Adran 4 ceir trosolwg o'r amcangyfrif ystadegol Bayesaidd, yn benodol **tebygoliaeth** (*likelihood*) y sampl, y **rhagfanyleb** (*prior specification*), ac amcangyfrif MCMC **naid wrthdroadwy** (*reversible jump Markov chain Monte Carlo, RJ-MCMC*). Yn Adran 5, ceir eglurhad o'r model gwerthoedd eithaf a chanlyniadau'r cymhwysiad ym Môr y Gogledd. Cynigir trafodaeth a chasgliadau yn Adran 6.

2 Cymhwysiad cymhellol

Defnyddir data H_S, cyfeiriad storm a thymor o hanesolwg Reistad et al. (2011) ar gyfer lleoliad yng ngogledd Môr y Gogledd. Mae'r data gwreiddiol o'r hanesolwg yn nhermau H_S, cyfeiriad a thymor am ddilyniant o gyflyrau'r môr dros gyfnod di-dor o $P_0 = 60$ mlynedd. Gan fod cyflyrau môr dilynol yn ddibynnol ar ei gilydd, defnyddir y dull a esbonnir yn Ewans a Jonathan (2008) i ynysu a chanfod priodweddau brigau'r stormydd cyfatebol, am y cyfnod cyfan. Dangosir y set gyfan o werthoedd H_S brig mewn llwyd yn Ffigur 1; y rhain yw ffocws y dadansoddiad presennol.

Tuedda stormydd geirwon ym Môr y Gogledd i ddod o'r gorllewin a Môr Iwerydd. Gall cyfeiriadau stormydd amrywio'n fawr, beth bynnag. Rhydd cyrchoedd hirion i gyfeiriad yr Iwerydd, y Môr Norwyaidd a rhan ddeheuol Môr y Gogledd ei hun gyfle i donnau mawrion ddatblygu. Gall cysgodion tiroedd mawr Prydain a Norwy hefyd ffocysu tonnau i gyfeiriadau penodol. Gwelir yr amrywiad systematig yng ngwerthoedd H_S brig yn ôl cyfeiriad a thymor yn Ffigur 1. Gan mai model brigau dros drothwy a ddefnyddir gennym yn y gwaith hwn, dengys y ffigur hefyd y gwerthoedd hynny sydd dros y trothwy (a drafodir ymhellach isod) mewn du.

Yn y gwaith hwn, diffinnir y cyd-newidyn 'cyfeiriad' fel y cyfeiriad y daw **brig storm** (*storm peak*) ohono, mewn graddau yn glocwedd o'r gogledd, e.e. gyda 0° yn golygu storm o'r gogledd, a 270° yn golygu storm o'r gorllewin. Defnyddir tafluniad o 'diwrnod y flwyddyn' ar i $[0, 360)$ i ddiffinio 'tymor'. Yn y ffigur, gwelir effaith ffocysu stormydd o'r Iwerydd yn glir ym mhanel (a) tua 200°. Dengys panel (b) effaith tymhorol amlwg. Rhaid modelu'r amrywiadau hyn yn ddilys mewn model gwerthoedd eithaf, er mwyn rhagfynegi priodweddau cynffon dosraniad H_S brig yn arbennig.



Ffigur 1: H_S brig yn erbyn cyfeiriad (a) a thymor (b). Dangosir gwerthoedd brigau dros y trothwy gwerth eithaf anunfan mewn du. Mesurir H_S brig mewn metrau (m), a chyfeiriad a thymor mewn graddau, fel yr esbonnir yn y testun.

3 Cefndir mathemategol

3.1 Dosraniad brigau dros drothwy

Ystyrir hapnewidyn $Y \in \mathbb{R}$ (e.e. H_S brig y storm) a fector o gyd-newidynnau $\boldsymbol{\theta} = (\theta_1, \theta_2, ...) \in \mathcal{D}$. Yn y gwaith hwn, er mwyn pendantrwydd ystyrir cyfeiriad (θ_1) a thymor (θ_2) storm fel cyd-newidynnau, ar y parth $\mathcal{D} = [0, 360) \times [0, 360)$. Nesaf, diffinnir trothwy uchel anunfan $\psi = \psi(\boldsymbol{\theta}) \in \mathbb{R}$ yn amrywio gyda $\boldsymbol{\theta}$ mewn modd priodol (e.e. yn nhermau cwantil uchel o ddosraniad y newidyn amodol $Y|\boldsymbol{\theta}$; gweler e.e. Hafaliad (11) yn Adran 5 a Northrop et al., 2017). Wedyn, cymerir bod Y yn newidyn **uchaf sefydlog** (*max-stable*, e.e. Beirlant et al., 2004). Fel canlyniad, yn ôl Pickands (1975), mae brigau dros drothwy Y, h.y. y newidyn amodol $Y|Y > \psi(\boldsymbol{\theta})$ yn dilyn dosraniad Pareto cyffredinoledig (GP), os yw'r trothwy yn ddigon uchel. Dwysedd y dosraniad GP ar gyfer **siâp** $\xi = \xi(\boldsymbol{\theta}) \in \mathbb{R}\backslash\{0\}$ a **graddfa** $\sigma = \sigma(\boldsymbol{\theta}) \in \mathbb{R}_{>0}$ yw

$$f_{\mathrm{GP}}(y|\xi, \sigma, \psi) = \frac{1}{\sigma}\left[\left(1 + \frac{\xi}{\sigma}(y - \psi)\right)^{-1/\xi - 1}\right]_+ . \tag{2}$$

Yn yr achos $\xi = 0$, ffurf y dwysedd yw $(1/\sigma)\exp(-(y-\psi)/\sigma)$ gyda'r **cynhaliad** (*support*) $\mathbb{R}_{>0}$ ar gyfer σ. Nodir bod y paramedrau ξ a σ, fel ψ, yn anunfan, h.y. yn amrywio gyda $\boldsymbol{\theta}$.

Yn fras iawn, bwriad y gwaith presennol yw defnyddio sampl o ddata ar gyfer Y a $\boldsymbol{\theta}$ ynghyd â'r ffwythiant ψ i ddarganfod amcangyfrifion anunfan ar gyfer y paramedrau ξ a σ. Y cymhlethdod yw amcangyfrif yn union sut y mae ξ a σ yn amrywio gyda $\boldsymbol{\theta}$.

3.2 Ymraniad Voronoi

Yn gyffredinol, nid oes gennym wybodaeth bendant i ragdybio ffurf y ffwythiannau $\xi(\boldsymbol{\theta})$ a $\sigma(\boldsymbol{\theta})$. Felly rhaid eu hamcangyfrif o gasgliad ystadegol (a esbonnir ymhellach yn Adran 4). Mae'n fanteisiol felly i ddiffinio'r ffwythiannau yn nhermau cynrychioliad cyffredinol, gan ddefnyddio set o **ffwythiannau sail** (*basis functions*) ar gyfer parth

y cyd-newidynnau, $\mathcal{D}$. Mae dewis eang o ffwythiannau sail priodol. Eisioes ym maes eithafon amgylcheddol, cyhoeddwyd ar seiliau Fourier (e.e. Jonathan ac Ewans, 2011, mewn un dimensiwn, 1-D) a seiliau sblein (e.e. Randell et al., 2015, mewn 1-D a 2-D). Gweler Zanini et al. (2020) am ragor o drafodaeth. Yn y gwaith hwn, ystyrir seiliau Voronoi 2-D.

Mantais y cynrychioliad Voronoi (e.e. Bodin et al., 2009) yw ei symlrwydd. Y cwbl sydd ei angen i ddiffinio cynrychioliad Voronoi C ar $\mathbb{R}^2$ yw nifer m o gnewyll (neu ganolfannau) C_k, $k = 1, 2, ..., m$ gyda chyfesurynnau $\boldsymbol{r}_k \in \mathcal{D}$ a gwerthoedd $g_k \in \mathbb{R}$ er mwyn **ymrannu** (*partition*) y parth $\mathcal{D}$. Wedyn diffinnir ffwythiant $g_V(\boldsymbol{\theta}) \in \mathbb{R}$ ar gyfer $\boldsymbol{\theta} \in \mathcal{D}$ gan y mynegiad

$$g_V(\boldsymbol{\theta}) = g_{k^*} \qquad \text{gyda} \qquad k^* = \underset{k \in \mathbb{N}, k \leq m}{\operatorname{argmin}} \left(\texttt{pell}(\boldsymbol{\theta}, \boldsymbol{r}_k)\right), \tag{3}$$

lle y mae'r ffwythiant $\texttt{pell}(a, b) \in \mathbb{R}_{\geq 0}$ yn rhoi'r pellter rhwng a a b ar $\mathbb{R}^2$; h.y. mae $g_V(\boldsymbol{\theta})$ yn cymryd gwerth y cnewyllyn C_{k^*} â chyfesurynnau $\boldsymbol{r}_{k^*}$ sydd agosaf (yn nhermau $\texttt{pell}$) at $\boldsymbol{\theta}$. Felly mae $g_V(\boldsymbol{\theta})$ yn ffwythiant **cyson fesul rhan** (*piecewise constant*) ar $\mathbb{R}^2$.

Wrth gwrs, mae gwerthoedd y ffwythiant yn dibynnu ar y dewis o $\texttt{pell}$. Nodwedd arbennig o'r parth cyfeiriadol-dymhorol $\mathcal{D}$ yw ei fod yn gyfnodol mewn cyfeiriad a thymor; h.y. gwyddom fod pellter y pwyntiau a'r cyfesurynnau (10,10) a (350,350) o'r tarddbwynt (0,0) yn hafal. Yn ogystal, nid oes gennym reswm sylfaenol dros ystyried gwahaniaeth mewn cyfeiriad yn wahanol i wahaniaeth tymhorol yn y gwaith hwn. Awgryma hyn fod y dewis $\texttt{pell}(\boldsymbol{\theta}, \boldsymbol{r}) = \left(\sum_j \min_{k \in \mathbb{Z}} (\theta_j - r_j + 360k)^2\right)^{1/2}$ yn rhesymol ar gyfer $\texttt{pell}$. Beth bynnag, mae'n bwysig nodi bod dewisiadau amgen yn bosib, gyda goblygiadau pellgyrhaeddol o ran y casgliad ystadegol.

Er mwyn amcangyfrif cynrychioliad Voronoi addas g_V ar gyfer rhyw ffwythiant g ar $\mathcal{D}$ felly, y cwbl sydd ei angen yw dewis nifer priodol m o gnewyll C_k ynghyd â'u cyfesurynnau a'u gwerthoedd i gael y gyfatebiaeth orau rhwng g a g_V ar $\mathcal{D}$. Felly, er mwyn amcangyfrif model gwerthoedd eithaf anunfan ar gyfer H_S gyda chyfeiriad a thymor, rhaid amcangyfrif ymraniadau Voronoi gwahanol ar gyfer amrywiad ξ a σ yn Hafaliad (2) gyda chyfeiriad a thymor. Bwriad yr amcangyfrif ystadegol a drafodir yn yr adran nesaf yw disgrifio sut i ddarganfod y ddau ymraniad Voronoi hyn drwy resymu Bayesaidd.

4 Amcangyfrif ystadegol

4.1 Y sampl a'i debygoliaeth

Ystyrir sampl o arsylwadau $S = \{(y_i, \boldsymbol{\theta}_i)\}_{i=1}^n$ o werthoedd o frigau Y (sef H_S brig) dros drothwy anunfan ψ, a chyd-newidynnau $\boldsymbol{\theta}$ (sef cyfeiriad a thymor), fel y dangosir mewn du yn Ffigur 1. Tybiwn fod yr arsylwad y_i wedi ei dynnu'n annibynnol o ddosraniad GP â pharamedrau $\xi_V(\boldsymbol{\theta}_i)$ a $\sigma_V(\boldsymbol{\theta}_i)$, sef ffwythiannau Voronoi wedi eu gwerthuso ar $\boldsymbol{\theta}_i$.

Tebygoliaeth (*likelihood*) y sampl felly yw

$$L(\xi_V, \sigma_V | S, \psi) = \prod_{i=1}^{n} f_{\mathrm{GP}}(y_i | \xi_V(\boldsymbol{\theta}_i), \sigma_V(\boldsymbol{\theta}_i), \psi(\boldsymbol{\theta}_i)). \tag{4}$$

Fel yr amlinellwyd yn Adran 3.2, disgrifir y cynrychioliad Voronoi ar gyfer $\eta \in \{\xi, \sigma\}$ gan y set o $3m_\xi + 3m_\sigma + 2$ paramedr $\Omega = (\mu_\xi, \{\boldsymbol{\rho}_{k,\xi}\}_{k=1}^{\mu_\xi}, \{\gamma_{k,\xi}\}_{k=1}^{\mu_\xi}, \mu_\sigma, \{\boldsymbol{\rho}_{k,\sigma}\}_{k=1}^{\mu_\sigma}, \{\gamma_{k,\sigma}\}_{k=1}^{\mu_\sigma})$. Defnyddir llythrennau Groegaidd μ, ρ a γ (yn lle m, r a g) i ddynodi mai hapnewidynnau i'w hamcangyfrif bellach yw priodweddau'r ymraniad Voronoi. Serch nodiant symlach, ysgrifennwn $\Omega = (\mu_\xi, \{\boldsymbol{\rho}_{k,\xi}\}, \{\gamma_{k,\xi}\}, \mu_\sigma, \{\boldsymbol{\rho}_{k,\sigma}\}, \{\gamma_{k,\sigma}\})$. Pan fo angen, cyfeirir at y set o werthoedd $\{y_i\}_{i=1}^{n}$ fel $\{y_i\}$ er hwylustod hefyd. Amcan y dadansoddiad ystadegol presennol yw amcangyfrif y set Ω. Fel y soniwyd eisioes, ar gyfer eglurdeb yn y gwaith hwn, ystyrir bod y trothwy ψ wedi ei bennu ymlaen llaw.

4.2 Rhagfanyleb

Mewn rhesymu Bayesaidd, cyfunir y debygoliaeth L a **rhagfanyleb** (*prior specification*) o gyd-ddosraniad holl baramedrau'r model i gyfrifo cyd-ôl-ddosraniad y paramedrau. Mae'r rhagfanyleb a ddefnyddir yn weddol gyffredinol, fel a ganlyn. Tybir bod y rhifau μ_ξ, μ_σ (yn annibynnol) yn dilyn dosraniadau Poisson fel y gwelir mewn gwyrdd yn Ffigurau 3(c) a 4(c). Cymerir bod $\gamma_{k,\xi} \in \mathcal{R}_\xi$, $k = 1, 2, ..., \mu_\xi$ a $\gamma_{k,\sigma} \in \mathcal{R}_\xi$, $k = 1, 2, ..., \mu_\sigma$ (yn annibynnol) yn dilyn dosraniadau unffurf fel y gwelir mewn gwyrdd yn Ffigurau 3(d) a 4(d). Noder na all $\gamma_{k,\sigma}$ fod yn negyddol. Dangosodd blynyddoedd o ymchwil i H_S brig (e.e. Ewans a Jonathan, 2008; Jonathan ac Ewans, 2013) fod gwerthoedd ξ yn tueddu i fod yn y cyfwng $(-0.4, 0.1)$, ac felly defnyddir y cyfwng hwn ar gyfer $\mathcal{R}_\xi$; yn ogystal, gwyddys fod y cyfwng $\mathcal{R}_\sigma = (0.1, 3)$ yn addas ar gyfer disgrifio brigau dros drothwy o H_S brig, ar gyfer dewisiadau nodweddiadol o drothwy. Yn olaf, tybir bod lleoliad $\boldsymbol{r}_{k,\eta}$ pob un o'r cnewyll $C_{k,\eta}$ yn annibynnol yn dilyn dosraniad unffurf ar y parth $\mathcal{D}$, am $k = 1, 2, ..., \mu_\eta$ ac $\eta \in \{\xi, \sigma\}$.

4.3 Samplu Gibbs a Metropolis-Hastings

Amcangyfrifir cyd-ôl-ddosraniad y set Ω o baramedrau drwy ddefnyddio rhesymu Bayesaidd, gan ddefnyddio samplu MCMC **naid wrthdroadwy** (*reversible jump*, e.e. Green, 1995; Zanini et al., 2020). Bwriad y samplu yw creu dilyniant hir o amcangyfrifon o'r paramedrau yn Ω sydd yn y pen draw yn cydgyfeirio at hapsampl dibynnol o gyd-ôl-ddosraniad Ω yn ôl y model.

Gan gymryd i ddechrau fod gwerthoedd μ_ξ a μ_σ wedi eu pennu yn m_ξ a m_σ yn ôl eu trefn, defnyddiwn samplu Gibbs (e.e. Gamerman a Lopes, 2006) i amcangyfrif cyd-ddosraniad y paramedrau $\{\{\boldsymbol{\rho}_{k,\eta}\}, \{\gamma_{k,\eta}\}\}_{\eta \in \{\xi,\sigma\}}$ sy'n weddill. Syniad y broses yw

samplu yn eu tro ddwyseddau **amodol llawn** (*full conditional*)

$$\begin{array}{rcl} f(\boldsymbol{\rho}_{k,\xi}|S,\Omega\backslash\boldsymbol{\rho}_{k,\xi}) & \propto & f(S|\boldsymbol{\rho}_{k,\xi}\Omega\backslash\boldsymbol{\rho}_{k,\xi}) \times f(\boldsymbol{\rho}_{k,\xi}) \\ f(\gamma_{k,\xi}|S,\Omega\backslash\gamma_{k,\xi}) & \propto & f(S|\gamma_{k,\xi},\Omega\backslash\gamma_{k,\xi}) \times f(\gamma_{k,\xi}) \\ f(\boldsymbol{\rho}_{k,\sigma}|S,\Omega\backslash\boldsymbol{\rho}_{k,\sigma}) & \propto & f(S|\boldsymbol{\rho}_{k,\sigma},\Omega\backslash\boldsymbol{\rho}_{k,\sigma}) \times f(\boldsymbol{\rho}_{k,\sigma}) \\ f(\gamma_{k,\sigma}|S,\Omega\backslash\gamma_{k,\sigma}) & \propto & f(S|\gamma_{k,\sigma},\Omega\backslash\gamma_{k,\sigma}) \times f(\gamma_{k,\sigma}) \end{array} \quad (5)$$

ar gyfer $k = 1, 2, ...$ Gellir profi, o dan amgylchiadau gweddol gyffredinol, bod y gadwyn o werthoedd a grëir yn cydgyfeirio â hapsampl o'r ôl-ddosraniad $\Omega|S$. Gan nad yw'n bosib ysgrifennu unrhyw un o'r dosraniadau uchod ar **ffurf gaeedig** (*closed form*), rhaid defnyddio samplu Metropolis-Hastings (e.e. Gamerman a Lopes, 2006) i gyflawni'r samplu uchod. Yr fyr, mae'r broses fel a ganlyn. Cymerwn ein bod eisioes wedi cyrraedd cyflwr W ar gyfer Ω ar ôl i iteriad o samplu Gibbs. Nawr, dewiswn un newidyn ω o Ω (sydd â gwerth cyfredol w) a chynigiwn werth newydd w^* amdano, gan hefyd greu gwerth W^* o W fel canlyniad. Derbyniwn y gwerth newydd (ar gyfer gwerth ω ar iteriad $i+1$) â thebygolrwydd

$$\min\left(1, \frac{f(\{y_i\}|W^*)}{f(\{y_i\}|W)} \frac{f(w^*)}{f(w)} \frac{q(w|w^*)}{q(w^*|w)}\right). \quad (6)$$

Yn y mynegiad uchod, $q(a, b)$ yw'r **dwysedd cynnig** (*proposal density*, e.e. Gamerman a Lopes, 2006). Yn y gwaith presennol, defnyddiwn ddwyseddau cynnig cymesur **hapgerddediad** (*random walk*) gyda $q(a, b) = q(b, a)$, sy'n golygu symleiddio pellach i Hafaliad 6. Mae'r pedwar term sy'n weddill yn y mynegiad yn hysbys i ni, sef tebygoliaethau'r sampl yn ôl W ac yn ôl W^*, a gwerthoedd **rhagddwyseddau** (*prior densities*) ar gyfer w^* a w. Felly gallwn werthuso'r mynegiad yn Hafaliad 6 a phenderfynu a ddylid derbyn y gwerth newydd w^* ar gyfer ω ai peidio.

4.4 Samplu naid wrthdroadwy

Esbonnir uchod sut i gyflawni samplu Metropolis-Hastings mewn Gibbs, drwy iteru â Hafaliadau 5 a 6. Rhydd hyn fodd o amcangyfrif ôl-ddosraniad $\Omega|S$ yn ôl y model, ond bod $\mu_\xi = m_\xi$ a $\mu_\sigma = m_\sigma$ wedi eu pennu. Yn yr adran hon, ystyriwn sut i ehangu'r samplu i amcangyfrif cyd-ôl-ddosbarthiad yr holl newidynnau yn Ω, gan gynnwys μ_ξ a μ_σ, drwy ddefnyddio samplu MCMC naid wrthdroadwy.

Gellid ystyried y broses naid wrthdroadwy fel estyniad o'r samplu Metropolis-Hastings uchod. Ond bellach, yn hytrach na chynnig gwerthoedd newydd ar gyfer newidynnau sy'n bodoli'n barod yn unig, cynigiwn **drosiadau** (*transitions*) sy'n creu neu'n difa newidynnau. Mae trosiad **creu** yn creu cell newydd (ar gyfer ymraniad Voronoi ξ neu σ): felly rhaid creu newidyn newydd ar gyfer cyfesurynnau'r cnewyllyn, a newidyn newydd ar gyfer gwerth y gell Voronoi. Rhaid hefyd addasu gwerthoedd y newidynnau ar gyfer yr ymraniad Voronoi sy'n bodoli eisioes i gael yr hen gynrychioliadau a'r newydd i gyfateb. I'r gwrthwyneb, mae trosiad **difa** yn dileu un o'r celloedd Voronoi sydd eisioes yn bodoli. Rhaid cael gwared ar y newidynnau ar gyfer cyfesurynnau'r cnewyllyn

perthnasol a gwerth y gell. Hefyd rhaid addasu'r gwerthoedd sy'n weddill eto i gael yr hen gynrychioliadau a'r rhai newydd i gyfateb.

Serch symlrwydd y disgrifiad, ystyriwn y sefyllfa ar ôl i iteriad o samplu naid wrthdroadwy, wedi cyrraedd cyflwr $\Omega = W$. Gyda'r nodiant λ_η am y set $(\mu_\eta, \{\boldsymbol{\rho}_{k,\eta}\}, \{\gamma_{k,\eta}\})$, $\eta \in \{\xi, \sigma\}$, ystyriwn drosiad o $\lambda_\eta = \ell$ i ℓ^*, ac felly hefyd o $\Omega = W$ i W^*. Gwerth λ_η cyn y trosiad arfaethedig yw $\ell = (m, \{\boldsymbol{r}_k\}, \{g_k\})$, ac $\ell^* = (m^*, \{\boldsymbol{r}_k^*\}, \{g_k^*\})$ ar ôl y trosiad. Ehangir Hafaliad 6 i'r ffurf

$$\min\left(1, \frac{f(\{y_i\}|W^*)}{f(\{y_i\}|W)} \frac{f(\ell^*)}{f(\ell)} \frac{q(\ell|\ell^*)}{q(\ell^*|\ell)} \left|\frac{\partial(\ell^{e*})}{\partial(\ell^e)}\right|\right). \tag{7}$$

Mae dewis trosiadau yn ôl Hafaliad 7 yn sicrhau y bodlonir amodau **cydbwysedd manwl** (*detailed balance*) gan y gadwyn a grëir. Yn yr hafaliad, $f(\ell)$ yw rhagddwysedd λ_η cyn y trosiad arfaethedig. Gellir ffactorio hwn i gael

$$f(\ell) = f(m) f(\{\boldsymbol{r}_k\}|m) f(\{g_k\}|\{\boldsymbol{r}_k\}, m) \tag{8}$$

lle y mae ffurfiau'r dwyseddau ar y dde eisioes wedi eu pennu yn Adran 4.2. Mae mynegiad cyffelyb ar gael ar gyfer $f(\ell^*)$, rhagddwysedd λ_η ar ôl y trosiad.

Mae'r dwysedd cynnig $q(\ell^*|\ell)$ yn Hafaliad 7 yn gysylltiedig â'r trosiad (neu'r naid) o gynrychioliad $\lambda_\eta = \ell$ i ℓ^*, ac yn arbennig â newid yn nifer y newidynnau yn λ_η. Felly cynrychiola $q(\ell|\ell^*)$ ddwysedd cynnig y trosiad gwrthdro. Yn ffodus, gellir ffactorio'r gymhareb $q(\ell^*|\ell)/q(\ell|\ell^*)$ fel a ganlyn

$$\frac{q(\ell|\ell^*)}{q(\ell^*|\ell)} = \frac{q(m|m^*)}{q(m^*|m)} \frac{q(\{\boldsymbol{r}_k\}|\{\boldsymbol{r}_k^*\}, m^*, m)}{q(\{\boldsymbol{r}_k^*\}|\{\boldsymbol{r}_k\}, m, m^*)} \frac{q(\{g_k\}|\{g_k^*\}, \{\boldsymbol{r}_k^*\}, \{\boldsymbol{r}_k\})}{q(\{g_k^*\}|\{g_k\}, \{\boldsymbol{r}_k\}, \{\boldsymbol{r}_k^*\})}. \tag{9}$$

Y term olaf yn Hafaliad 7 yw'r **Jacobiad** (*Jacobian*) am y trawsffurfiad newidynnau. Er mwyn gwerthuso'r Jacobiad, rhaid ymestyn y set λ_η o newidynnau i λ_η^e, fel bod yr un nifer o newidynnau yn λ_η^e (cyn y trosiad arfaethedig) ag yn λ_η^{e*} wedi'r trosiad. Yn ffodus, parthed y trawsffurfiadau newidynnau a ddefnyddir yn y gwaith presennol ar gyfer trosiadau creu a difa, mae gwerth y Jacobiad yn hafal i un.

Mae gwerthuso'r termau yn y gymhareb cynnig yn Hafaliad 9 yn weddol hawdd, ond rhaid ystyried natur y trosiad arfaethedig. Serch pendantrwydd, ystyriwn drosiad creu. Yn y cyswllt hwn, $q(m^*|m)$ yw'r tebygolrwydd $p_C \in (0,1)$ o gynnig trosiad creu; mae hwn yn rhydd i ni ei bennu. Yn debyg, $q(m|m^*)$ yw'r tebygolrwydd $p_D \in (0,1)$ o gynnig y trosiad difa gwrthdro, a'n dewis ni. $q(\{\boldsymbol{r}_k\}|\{\boldsymbol{r}_k^*\}, m^*, m)$ yw'r tebygolrwydd o ddifa un gell Voronoi benodol yn y trosiad gwrthdro; gan fod $m^* = m + 1$ yn bresennol, y tebygolrwydd o ddifa un gell yn arbennig yw $1/(m+1)$. $q(\{\boldsymbol{r}_k\}|\{\boldsymbol{r}_k^*\}, m^*, m)$ yw dwysedd creu cyfesurynnau cnewyllyn y gell newydd, sef $1/|\mathcal{D}|$ o dan amodau'r trosiad arfaethiedig, gyda $|\mathcal{D}|$ yn cynrychioli mesur parth y cyd-newidynnau. $q(\{g_k\}|\{g_k^*\}, \{\boldsymbol{r}_k^*\}, \{\boldsymbol{r}_k\})$ yw'r dwysedd cynnig ar gyfer y set $\{g_k\}$ yn y trosiad gwrthdro; gan y cyfrifir y rhain **yn benderfynedig** (*deterministically*) heb unrhyw elfen o hap o'r gwerthoedd cyfatebol yn $\{g_k^*\}$, mae gwerth y dwysedd cynnig hwn yn hafal i un. Yn olaf, $q(\{g_k^*\}|\{g_k\}, \{\boldsymbol{r}_k\}, \{\boldsymbol{r}_k^*\})$ yw'r dwysedd cynnig ar gyfer gwerth y gell Voronoi newydd a grëir. Yn y gwaith presennol, cymer hwn ffurf Gaussaidd gyda chymedr $\tilde{g}$ yn hafal i werth lleoliad y cnewyllyn

newydd yn yr ymraniad Voronoi cyn trosi, a gwyriant safonol κ yr ydym yn rhydd i'w bennu. Ffurf y gymhareb cynnig ar gyfer trosiad creu felly yw

$$\frac{q(\ell|\ell^*)}{q(\ell^*|\ell)} = \frac{p_D}{p_C}\frac{|\mathcal{D}|}{m+1}(2\pi)^{1/2}\kappa \exp\left(\frac{(g-\tilde{g})^2}{2\kappa^2}\right). \tag{10}$$

Dengys dadl debyg mai ffurf y gymhareb cynnig ar gyfer trosiad difa yw

$$\frac{q(\ell|\ell^*)}{q(\ell^*|\ell)} = \frac{p_C}{p_D}\frac{m}{|\mathcal{D}|}\frac{1}{(2\pi)^{1/2}\kappa} \exp\left(-\frac{(g-\tilde{g})^2}{2\kappa^2}\right). \tag{11}$$

Mae'r drafodaeth uchod yn dangos sut i fabwysiadu trosiadau naid wrthdroadwy mewn rhesymu Bayesaidd, a sut felly i gyflawni'r rhesymu cyfan. Rhaid bod yn ofalus mewn rhai amgylchiadau (e.e. i osgoi cynnig trosiad difa pan fo $m = 1$), a dylid gwneud dewisiadau ar gyfer gwerthoedd yr **hyperbaramedrau** (*hyper-parameters*) p_C, p_D a κ, a gwyriant safonol y cynnig hapgerddediad yn Hafaliad 6, er mwyn cael y dilyniant MCMC i **gymysgu** a chydgyfeirio â chyd-ôl-ddosraniad $\Omega|S$ mor hwylus â phosib. Gweler Zanini et al. (2020) am ragor o wybodaeth.

5 Canlyniadau'r cymhwysiad

Cymhwyswyd y fethodoleg uchod i amcangyfrif priodweddau eithafon H_S brig ym Môr y Gogledd, gan ddefnyddio'r data a gyflwynwyd yn Adran 2. Yn ogystal, gwiriwyd y fethodoleg drwy astudiaethau â data wedi eu hefelychu o brosesau yr oedd eu priodweddau yn hysbys i ni. Canfuwyd bod yr amcangyfrifon yn cyfateb yn dda i'r gwirionedd yn yr efelychiadau. Mae'r adran hon yn cynnig trosolwg o ganlyniadau'r cymhwysiad i H_S.

Mae Ffigur 2(a) yn dangos y trothwy ψ a ddefnyddiwyd yn y gwaith fel ffwythiant o gyfeiriad a thymor, a phanel (b) yn dangos cyfradd brigau H_S uwch y trothwy. Amcangyfrifwyd y rhain yn empirig o'r data. Ar gyfer ψ defnyddiwn yr amcangyfrif

$$\psi(\boldsymbol{\theta}) = \underset{h}{\operatorname{argmin}} \left| \frac{\sum_{j:\boldsymbol{\theta}_j \in \mathcal{N}_d(\boldsymbol{\theta})} \mathbb{I}(y_j \le h)}{\sum_k \mathbb{I}(\boldsymbol{\theta}_k \in \mathcal{N}_d(\boldsymbol{\theta}))} - \tau \right| \tag{12}$$

am $\boldsymbol{\theta} \in \mathcal{D}$, â thebygolrwydd $\tau = 0.8$. Mae'r trothwy yn cyfateb i gwantil uchel o ddosraniad lleol H_S brig ar y **gymdogaeth** (*neighbourhood*)

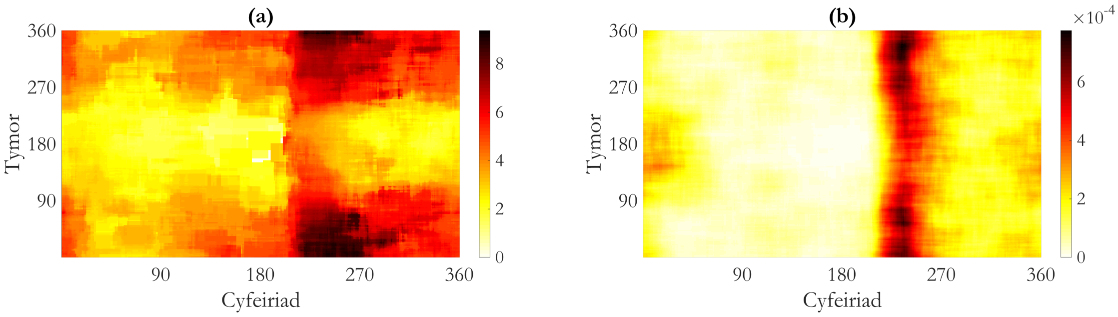
$$\mathcal{N}_d(\boldsymbol{\theta}) = \{\boldsymbol{\theta}' \in \mathcal{D} : \text{pell}(\boldsymbol{\theta}', \boldsymbol{\theta}) \le d\} \tag{13}$$

â $d = 30°$. Yn Hafaliad 12, $\mathbb{I}(x)$ yw'r dangosydd (= 1 pan fo x yn wir, ac = 0 fel arall). Gwelir bod y trothwy yn uwch yn y gaeaf fel y disgwylir, ond hefyd yn uwch i gyfeiriadau tua 200°, yn cyfateb i stormydd o gyfeiriad Môr Iwerydd sy'n cael eu ffocysu gan dir mawr Norwy.

Amcangyfrifir y gyfradd ν yn lleol mewn modd empirig cyffelyb, gan ddefnyddio

$$\nu(\boldsymbol{\theta}) = \frac{1}{|\mathcal{D}|P_0} \sum_{j:\boldsymbol{\theta}_j \in \mathcal{N}_d(\boldsymbol{\theta})} \mathbb{I}(y_j > \psi(\boldsymbol{\theta})) \tag{14}$$

â chyfaint y parth cyd-newidynnau $|\mathcal{D}|$, a chyfnod P_0 y data llawn mewn blynyddoedd. Mae priodweddau'r gyfradd yn amlygu ffocysu'r stormydd gan dir mawr Norwy ar gyfeiriadau'n agos at 200°; gwelir hefyd fod cyfradd y gaeaf yn uwch.



Ffigur 2: Amcangyfrifon empirig anunfan ar gyfer (a) trothwy ψ a (b) chyfradd ν. Cymer ψ yr un unedau â H_S.

Mae Ffigurau 3 a 4 yn crynhoi casgliadau ar gyfer ôl-ddosbarthiadiau graddfa GP σ a'r siâp cyfatebol ξ yn ôl y model yn amodol ar y sampl S. Yn y ddau ffigur, dengys panel (a) ganolrif ôl-ddosbarthiad ymylol y paramedr, a phanel (b) amrediad rhyngchwartel yr ôl-ddosbarthiad. Ym mhanel (c), ceir rhag-fàs (gwyrdd) ac ôl-fàs (du) y nifer μ o gelloedd Voronoi. Dengys panel (d) ragddwysedd (gwyrdd) ac ôl-ddwysedd (du) gwerthoedd y celloedd Voronoi. Defnyddir 10,000 iteriad ar gyfer yr amcangyfrif MCMC naid wrthdroadwy, a'r 1000 iteriad olaf ohono i amcangyfrif y mesurau a ddarlunnir.

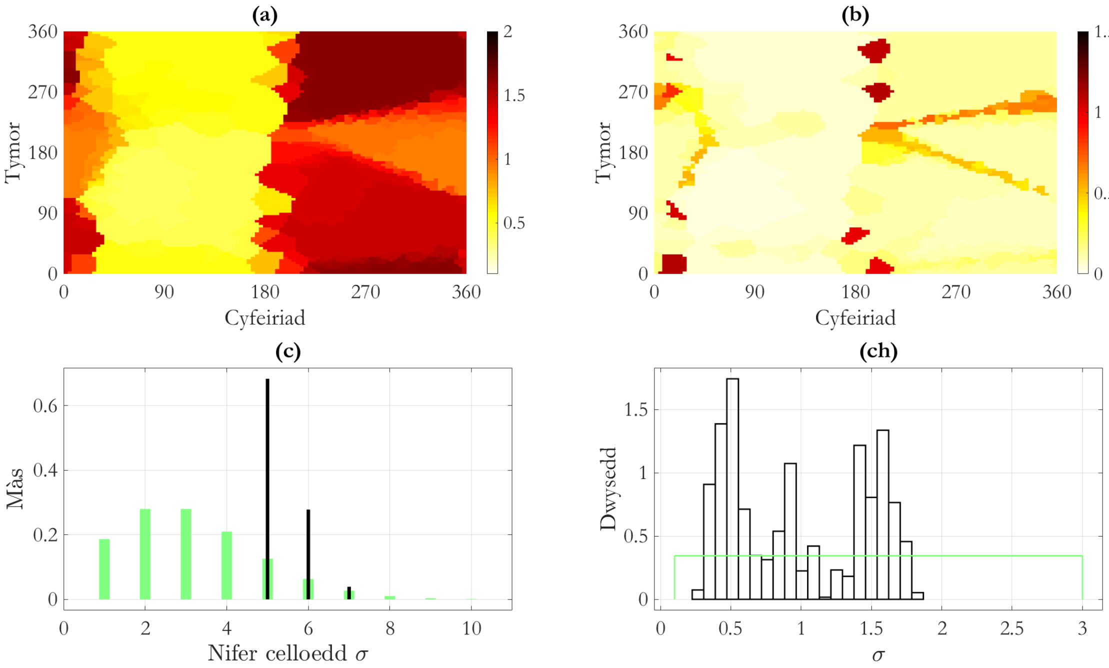
O ran $\sigma|S$ gwelir bod panel (a) yn adlewyrchu priodweddau cyfeiriadol a thymhorol H_S brig, a gwerthoedd yn fras rhwng 0.3 ac 1.9. Mae'r ansicrwydd mwyaf yn yr ardaloedd lle y mae $\sigma|S$ yn tueddu i newid (panel b). Ffafrir rhwng $\mu_\sigma|S = 5$ a 7 o gelloedd Voronoi, rhifau uchel o'u cymharu â'r rhag-fàs. Mae gwerthoedd $\gamma_\sigma|S$ yn tueddu i grynhoi oddeutu 0.5, 0.9 ac 1.5, yn wahanol iawn i'r rhagddwysedd.

O ran $\xi|S$, awgryma panel (a) fod $\xi|S$ prin yn amrywio â chyd-newidynnau, gyda gwerth o tua -0.1; mae golwg manwl yn dangos bod gwerthoedd yr haf fymryn yn is na gwerthoedd y gaeaf, gan awgrymu bod cynffon y dosbarthiad GP ychydig yn drymach (*heavier-tailed*) yn y gaeaf. Mae'r amrediad rhyngchwartel hefyd yn weddol gyson (panel b). Ffafrir rhwng $\mu_\xi|S = 1$ a 3 o gelloedd Voronoi, rhifau isel o'u cymharu â'r rhag-fàs. Mae gwerthoedd ymylol $\gamma_\xi|S$ yn tueddu i grynhoi oddeutu -0.1, yn wahanol iawn i'r rhagddwysedd.

Gan ddefnyddio'r cyd-ôl-ddosbarthiad $\Omega|S$, efelychwyd sampl o'r un maint â'r sampl gwreiddiol er mwyn cymharu eu priodweddau'n weledol yn Ffigur 5. Mae tueddiad gan y model a amcangyfrifwyd i grynhoi arsylwadau tua 200° yn rhywfaint mwy nag yn y data gwreiddiol, ond yn gyffredinol mae'r gymhariaeth yn dda.

Er mwyn dilysu'r model ymhellach, gellir cymharu amcangyfrifon ar gyfer cynffon dosraniad H_S brig dros bob dewis o gyd-newidynnau $\boldsymbol{\theta} \in \mathcal{D}$, a elwir yn ddosraniad **omni**; hynny yw, y dosraniad

$$F(h) = \Pr(H_S \leq h|S) = \int_{\boldsymbol{\theta}} \Pr(H_S \leq h|\boldsymbol{\theta}, S) f(\boldsymbol{\theta}|S)\, d\boldsymbol{\theta}. \quad (15)$$

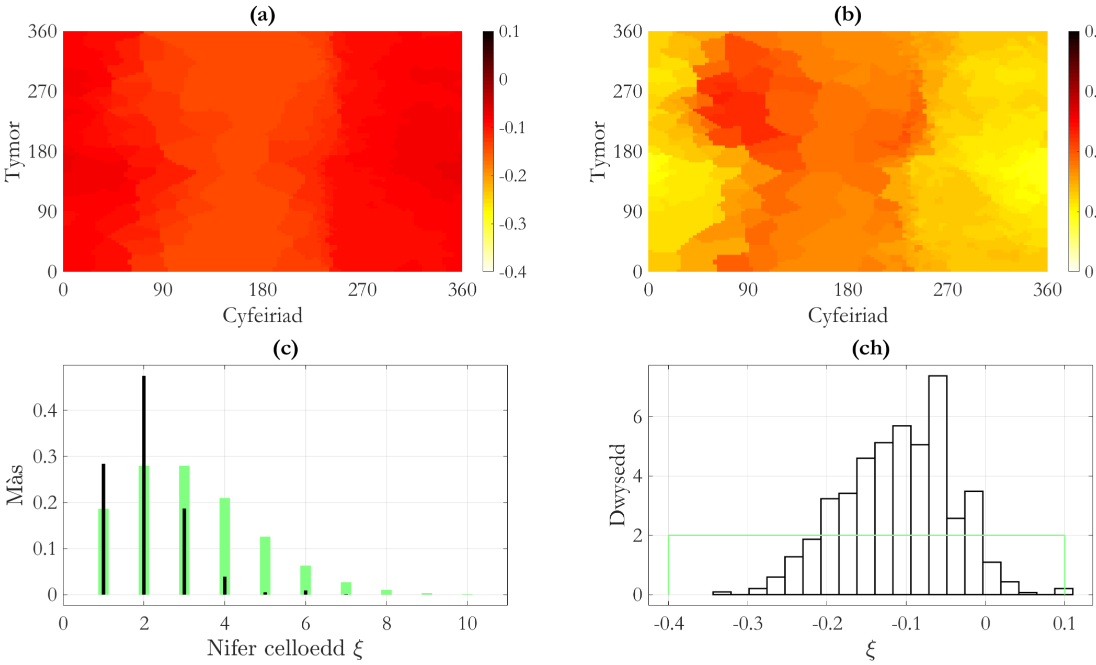


Ffigur 3: Priodweddau ôl-ddosraniad graddfa σ: (a) canolrif anunfan yr ôl-ddosbarthiad; (b) amrediad rhyngchwartel yr ôl-ddosbarthiad; (c) rhag-fàs (gwyrdd) ac ôl-fàs (du) y nifer μ_σ o gelloedd Voronoi; a (d) rhagddwysedd (gwyrdd) ac ôl-ddwysedd (du) gwerthoedd ymylol γ_σ y celloedd Voronoi. Defnyddir y 1000 iteriad olaf o'r gadwyn MCMC naid wrthdroadwy i amcangyfrif y mesurau a ddarlunnir. Cymer σ yr un unedau â H_S.

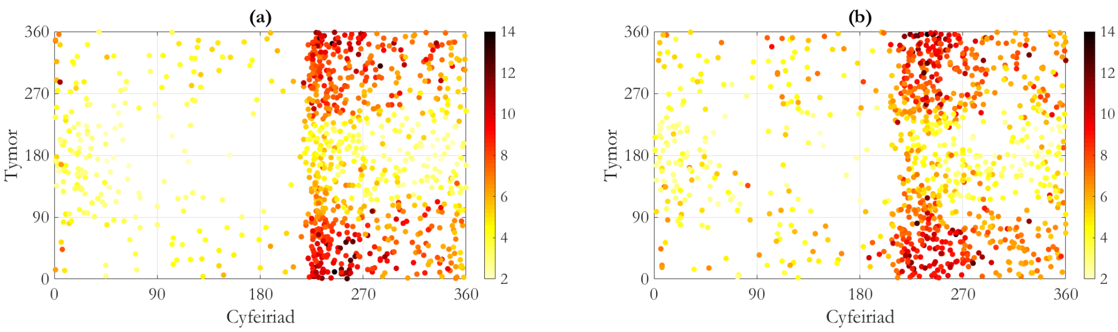
Mae'r dosraniad omni yn bwysig o ran meintioli gerwinder y storm yn ei chrynswth. Yn Ffigur 6(a), dangosir cynffon empirig dosraniad omni H_S brig mewn coch. Cynffonnau cyffelyb a grëwyd gan efelychiad yn ôl y model a amcangyfrifwyd yw'r llinellau llwydion. Mae'r cytundeb yn dda.

O ran peirianneg forol, un o'r casgliadau pwysicaf o fodelu H_S brig yw'r gwerth dychwelyd (Hafaliad (1) ac e.e. Jonathan et al., 2021) ar gyfer H_S brig yn cyfateb i gyfnod dychwelyd P sydd yn llawer hwy na chyfnod P_0 y data ei hun. Cyfrifir dosraniad y gwerth mwyaf mewn P mlynedd er enghraifft, gan gymryd bod cyfradd stormydd dros drothwy yn dilyn dosraniad Poisson. Wedyn, hawdd dangos (e.e. Jonathan ac Ewans, 2013) bod dosraniad uchafwerth H_S brig mewn P mlynedd, M_P, yn dilyn dosraniad gwerthoedd eithaf cyffredinoledig GEV (*generalized extreme value*)

$$\begin{aligned}\Pr(M_P \leq h|\boldsymbol{\theta}, S) &= \sum_{k=0}^{\infty} \frac{\exp\left(-\nu(\boldsymbol{\theta})P\right)\nu^k(\boldsymbol{\theta})P^k}{k!}(\Pr(H_S \leq h|\boldsymbol{\theta}, S))^k \qquad (16)\\ &= \exp\left(-\nu(\boldsymbol{\theta})P(1-\Pr(H_S \leq h|\boldsymbol{\theta}, S)\right)\end{aligned}$$



Ffigur 4: Priodweddau ôl-ddosraniad siâp ξ: (a) canolrif anunfan yr ôl-ddosbarthiad; (b) amrediad rhyngchwartel yr ôl-ddosbarthiad; (c) rhag-fàs (gwyrdd) ac ôl-fàs (du) y nifer μ_ξ o gelloedd Voronoi; a (d) rhagddwysedd (gwyrdd) ac ôl-ddwysedd (du) gwerthoedd ymylol γ_ξ y celloedd Voronoi. Defnyddir y 1000 iteriad olaf o'r gadwyn MCMC naid wrthdroadwy i amcangyfrif y mesurau a ddarlunnir.

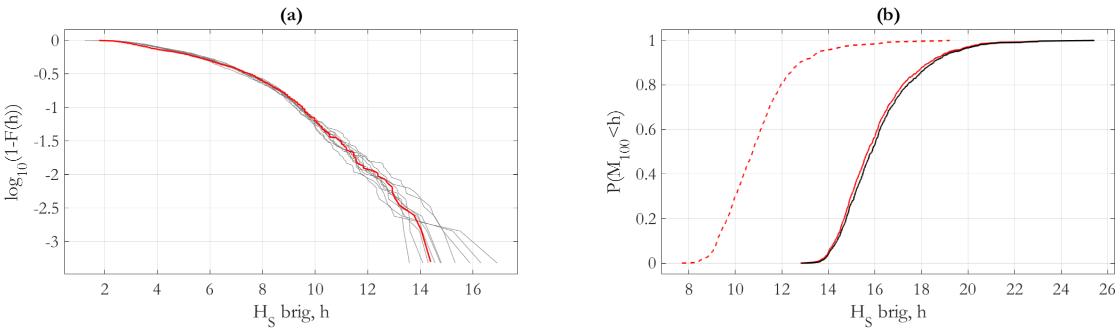


Ffigur 5: H_S brig dros drothwy gwerth eithaf (mewn metrau): arsylwadau gwreiddiol (a) ac efelychiad yn ôl y model (b) yn erbyn cyfeiriad a thymor.

ac o ganlyniad

$$\Pr(M_P \leq h|S) = \int_{\boldsymbol{\theta}} \Pr(M_P \leq h|\boldsymbol{\theta}, S) f(\boldsymbol{\theta}|S)\ d\boldsymbol{\theta} \tag{17}$$

sydd yn hawdd i'w gyfrifo. Yn Ffigur 6(b), ceir amcangyfrif (mewn du) o ddosraniad uchafwerth omni can mlynedd $M_{100}|S$. Gwelir bod canolrif yr ôl-ddosraniad tua 16m,



Ffigur 6: Gwerthoedd eithaf H_S brig dros drothwy (mewn metrau): (a) cymharu cynffon dosraniad omni H_S brig yn empirig o'r data (coch) ac o efelychiadau yn ôl y model (llwyd); (b) dosraniad uchafwerth omni can mlynedd $M_{100}|S$ (du) ac uchafwerthoedd can mlynedd wedi eu cyfyngu i gyfeiriadau yn y cyfwng $(200°, 300°]$ (coch solet), ac i gyfeiriadau a thymhorau yn y cyfyngau $(200°, 300°] \times (135°, 225°]$ (coch toredig).

ond bod tebygolrwydd bach fod gwerthoedd H_S brig dros 20m. Goblygiadau hyn yw bod y tonnau unigol mwyaf a welir mewn canrif yn debygol o fod ymhell dros 25m o uchder o'r cafn i'r brig. Er diddordeb, mae'r ffigur hefyd yn dangos ôl-ddosraniad uchafwerthoedd can mlynedd wedi eu cyfyngu i gyfeiriadau yn y cyfwng $(200°, 300°]$ mewn coch solet; mae'r gwahaniaeth â'r dosraniad omni yn fach, sy'n golygu mai'r cyfwng hwn sy'n dominyddu cynffon y dosraniad omni. Fel cyferbyniad, cynrychiola'r llinell goch doredig ôl-ddosraniad $M_{100}|S$ wedi ei gyfyngu i gyfeiriadau a thymhorau yn y cyfyngau $(200°, 300°] \times (135°, 225°]$. Mae effaith cyfyngu i dymor llai garw yn amlwg: llawer diogelach yw teithio a gweithio ym Môr y Gogledd yn yr haf! Noder y gellid defnyddio ystadegau megis **dargyfeiriad Kullback-Leibler** ac ystadegyn **Kolmogorov-Smirnov** i werthuso'r cytundeb rhwng y gwahanol amcangyfrifon o gynffon dosraniad.

6 Trafodaeth a chasgliadau

Mae nodweddu eithafon prosesau amgylcheddol anunfan yn her i'r gwyddonydd. Yn aml, mae nifer o gyd-newidynnau posib ar gael i esbonio amrywiant ym mhriodweddau gwerthoedd eithaf. Rhaid modelu'r amrywiant hwn yn ofalus os mai'r bwriad yw amcangyfrif gwerthoedd dychwelyd a chwantilau eithafol tebyg ar gyfer allbynnau'r broses. Yn y gwaith hwn, cyflwynir methodoleg ystadegol er mwyn modelu gwerthoedd eithaf gerwinder stormydd morol anunfan. Cyfunir model Pareto cyffredinoledig ar gyfer gerwinder brig stormydd morol dros drothwy â chynrychioliad Voronoi ar gyfer amrywiad paramedrau'r model gwerthoedd eithaf gyda chyfeiriad a thymor y storm. Defnyddir rhesymu Bayesaidd MCMC naid wrthdroadwy i amcangyfrif priodweddau'r cynrychioliad Voronoi. Dilysir bod efelychiadau yn ôl y model a amcangyfrifwyd yn cyfateb yn dda i'r data gwreiddiol, a defnyddir y model i amcangyfrif dosraniad uchafwerthoedd gerwinder stormydd am gyfnodau dychwelyd sy'n llawer hwy na chyfnod y data gwreiddiol.

Yn y gwaith presennol, er mwyn symlrwydd, defnyddir rhesymu Bayesaidd i amcangyfrif dim ond paramedrau'r model Pareto cyffredinoledig ar gyfer gwerthoedd brigau dros drothwy. Defnyddir amcangyfrifon empirig ar gyfer y trothwy gwerthoedd eithaf a chyfradd brigau dros drothwy yn arbennig. Mewn gwaith tebyg â chyd-newidynnau un dimensiwn gan Zanini et al. (2020), cynhwysir amcangyfrif y trothwy a'r gyfradd yn y rhesymu Bayesaidd hefyd.

Er mai cyd-newidynnau mewn dau ddimensiwm a drafodir yn yr erthygl bresennol, mae'r fethodoleg yn ymestyn yn ddidrafferth i gyd-newidynnau mewn dimensiynau uwch. Y prif reswm am hyn yw bod baich cyfrifiannol gwireddu ymraniad Voronoi yn graddio'n dda â dimensiwn. Y prif gymhlethdodau mewn dimensiynau uwch yw (1) pennu mesur addas o bellter (`pell` yn nodiant Adran 3.2), a (2) dewis neu amcangyfrif gwerthoedd addas ar gyfer yr hyperbaramedrau (y sonnir amdanynt e.e. ym mharagraff olaf Adran 4), i sicrhau bod yr amcangyfrif Bayesaidd yn mynd yn ei flaen yn hwylus.

O'i gymharu â chynrychioliadau eraill megis sbleiniau, mae cynrychioliad cyson fesul rhan Voronoi yn syml iawn. Mae'r symlrwydd hwn yn sicr yn anfanteisiol pan fydd effeithiau cyd-newidynnau yn tueddu i fod yn amlwg a hawdd eu hamcangyfrif, neu'n newid yn raddol â chyfesurynnau. Mae'r cynrychioliad Voronoi yn ennill ei blwyf mewn cymwysiadau lle y mae (1) amcangyfrif yr amrywiad cyd-newidynnol yn anodd (e.e. yng nghyswllt modelu gwerthoedd eithaf), (2) angen cynrychioliad syml i'w werthuso, a (3) angen cynrychioliad amlddimensiynol sy'n gyfrifiannol hwylus ei amcangyfrif.

Cydnabyddiaeth

Hoffai'r awdur gydnabod cymorth amryw o gyd-ymchwilwyr, ym Mhrifysgol Caerhirfryn a chwmni Shell yn arbennig. Hoffai ddiolch i Dr Tudur Davies o Brifysgol Aberystwyth, ac i Eryl a Robin, am gyngor wrth baratoi'r erthygl. Hoffai ddiolch ymhellach i ddau arfarnwr, golygyddion iaith a therminolegydd am eu hamser a'u sylwadau craff ar fersiwn blaenorol o'r erthygl.

Llyfryddiaeth

Bao, Y., Song, Z., Qiao, F., 2020. FIO-ESM version 2.0: Model description and evaluation. Journal of Geophysical Research: Oceans 125, e2019JC016036.

Battjes, J.A., Groenendijk, H.W., 2000. Wave height distributions on shallow foreshores. Coastal Eng. 40, 161–182.

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., 2004. Statistics of extremes: theory and applications. Wiley, Chichester, UK.

Bloemendaal, N., Haigh, I.D., de Moel, H., Muis, S., Haarsma, R.J., Aerts, J.C., 2020. Generation of a global synthetic tropical cyclone hazard dataset using storm. Scientific Data 7, 1–12.

Bodin, T., Sambridge, M., Gallagher, K., 2009. A self-parametrizing partition model approach to tomographic inverse problems. Inverse Problems 25, 055009.

Coles, S., 2001. An introduction to statistical modelling of extreme values. Springer.

Coles, S.G., Powell, E.A., 1996. Bayesian methods in extreme value modelling: a review and new developments. International Statistics Review 64, 119–136.

Davison, A.C., Padoan, S.A., Ribatet, M., 2012. Statistical modelling of spatial extremes. Statist. Sci. 27, 161–186.

Ewans, K.C., Jonathan, P., 2008. The effect of directionality on northern North Sea extreme wave design criteria. J. Offshore. Arct. Eng. 130, 041604:1–041604:8.

Gamerman, D., Lopes, H.F., 2006. Markov chain Monte Carlo: stochastic simulation for Bayesian inference. Chapman and Hall / CRC, Boca Raton, USA.

Green, P., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 82, 711–732.

Heffernan, J.E., Tawn, J.A., 2004. A conditional approach for multivariate extreme values. J. R. Statist. Soc. B 66, 497–546.

ISO19901-1, 2015. Petroleum and natural gas industries. Specific requirements for offshore structures. Part 1: Metocean design and operating considerations. First edition. International Standards Organisation.

Jonathan, P., Ewans, K.C., 2008. On modelling seasonality of extreme waves, in: Proc. 27th International Conf. on Offshore Mechanics and Arctic Engineering, 4-8 June, Estoril, Portugal.

Jonathan, P., Ewans, K.C., 2011. Modelling the seasonality of extreme waves in the Gulf of Mexico. ASME J. Offshore Mech. Arct. Eng. 133:021104.

Jonathan, P., Ewans, K.C., 2013. Statistical modelling of extreme ocean environments with implications for marine design : a review. Ocean Eng. 62, 91–109.

Jonathan, P., Randell, D., Wadsworth, J., Tawn, J., 2021. Uncertainties in return values from extreme value analysis of peaks over threshold using the generalized Pareto distribution. Ocean Eng. 220, 107725.

Kinsman, B., 2012. Wind waves: Their generation and propagation on the ocean surface. Dover, New York.

NORSOK N-006, 2015. NORSOK Standard N-006:2015: Assessment of structural integrity for existing offshore load-bearing structures. NORSOK, Norway.

Northrop, P., Attalides, N., Jonathan, P., 2017. Cross-validatory extreme value threshold selection and uncertainty with application to ocean storm severity. J. Roy. Statist. Soc. C 66, 93–120.

Pickands, J., 1975. Statistical inference using extreme order statistics. Ann. Statist. 3, 119–131.

Randell, D., Feld, G., Ewans, K., Jonathan, P., 2015. Distributions of return values for ocean wave characteristics in the South China Sea using directional-seasonal extreme value analysis. Environmetrics 26, 442–450.

Randell, D., Turnbull, K., Ewans, K., Jonathan, P., 2016. Bayesian inference for non-stationary marginal extremes. Environmetrics 27, 439–450.

Reistad, M., Breivik, O., Haakenstad, H., Aarnes, O.J., Furevik, B.R., Bidlot, J.R.,

2011. A high-resolution hindcast of wind and waves for the North Sea, the Norwegian Sea, and the Barents Sea. J. Geophys. Res. 116, 1–18.

Zanini, E., Eastoe, E., Jones, M., Randell, D., Jonathan, P., 2020. Covariate representations for non-stationary extremes. Environmetrics e2624 .