

Sigríður Sif Gylfadóttir

Samanburður á brúunaraðferðum

## Útdráttur

Fjallað er um samanburð á fjórum brúunaraðferðum í kortagerð: `griddata linear`, `griddata cubic`, `gsmooth2` og `kriging`. Aðferðirnar eru stuttlega ræddar og þeim beitt á meðalhitagögn frá 1961-2000. Skekkjur hinna ólíku aðferða eru bornar saman og mat lagt á hvaða aðferð henti best til að búa til meðalhitakort af landinu.

## 1 Inngangur

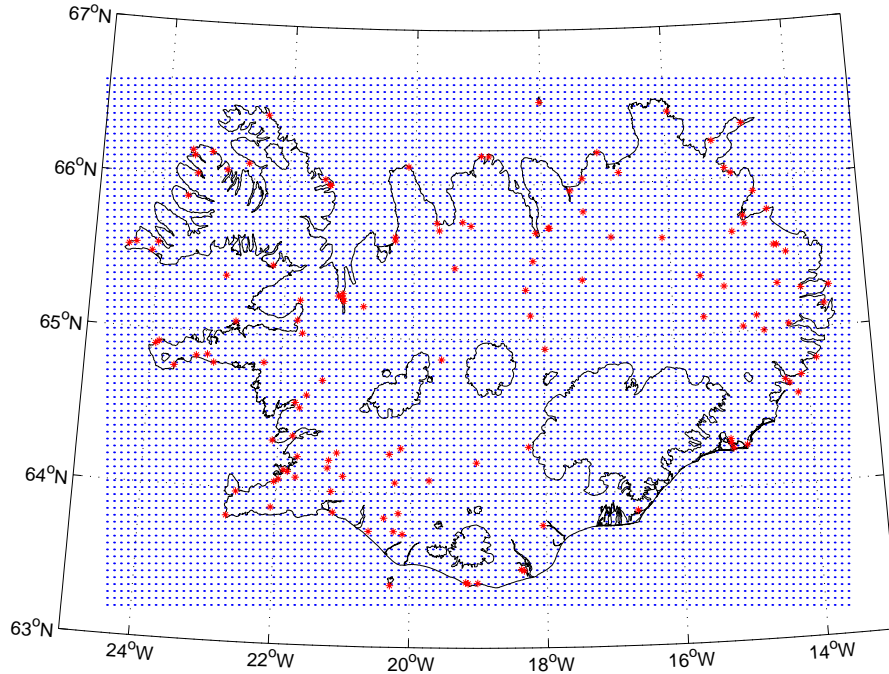
Veður- og jarðeðlisfræðilegar mælingar á yfirborði jarðar eru sjaldnast gerðar með reglulegu millibili, heldur þar sem aðstæður leyfa. Til þess að hægt sé að vinna úr tvívíðum, ójafnt dreifðum mælingum er oft nauðsynlegt að færa mælingar yfir á reglulegt net, og til þess fyrirfinnast ýmsar aðferðir. Hér á eftir fara niðurstöður á samanburði á brúunaraðferðum í kortagerð.

Lítið er á fjórar aðferðir: `kriging`, sem er tölfræðilegs eðlis, og þrjár margliðubrúanir af mismunandi stigum. Gögnin sem notuð eru við samanburðinn eru hitamælingar á Íslandi. Frávik meðalmánaðarhitastigs frá meðalmánaðarhitastigi 1961-90 á mönnuðum veðurathugunarstöðvum vítt og breitt um landið er brúað yfir á net að upplausn  $5 \times 5 \text{ km}^2$ . Aðferðirnar eru bornar saman fyrir 4 sumarmánuði og 6 vetrarmánuði á tímabilinu 1960-2000. Netíð sem brúað er á er sýnt á mynd 1 ásamt veðurstöðvum. Mismunur er á milli mánaða hvaða stöðvar eru virkar. Þessi vinna er þáttur í stærra verkefni sem snýst um gerð meðalhitakorta af Íslandi.

## 2 Lýsing á aðferðum

Reikniritið `griddata` fylgir með Matlab og finnst í fjórum tilbrigðum, `linear`, `cubic`, `nearest neighbor`, og `v4` sem er gamla Matlab 4 aðferðin. Öll tilbrigðin, að undanskildu hinu síðastnefnda, teljast til rúmfræðilegra brúunaraðferða. Þær byggja á skiptingu svæðis í delaunay þríhyrninganet, þar sem hornpunktar þríhyrninga eru gagnapunktur. Gildi í brúarpunkti er reiknað út frá hornpunktum þríhyrningsins sem hann lendir í. Reikniritið `linear` brúar línulega innan hvers þríhyrnings, þannig að innan þríhyrninga er brúunin samfelld, en ósamfelld á mörkum þeirra. Reikniritið `cubic` framkvæmir þriðja stigs brúun innan þríhyrninga, og verður flöturinn sem fæst samfelldur og með samfellda afleiðu (e. *smooth*). Reikniritið `nearest neighbor` metur gildið í brúunarpunkti sem gildið í nálægasta hornpunkti delaunay þríhyrnings, og er ómöguleg til kortlagningar þegar gagnasafnið er lítið og dreift yfir stórt svæði. Aðferðin sem kom með útgáfu 4 af Matlab, `v4`, notar tvívíða splæsibrúun (e. *minimum curvature*, *biharmonic spline*) og vegur áhrif þekktra punkta með fallinu

$$g = r^{-2} (\log(r) - 1)$$



Mynd 1: Netíð sem brúað er yfir á. Bláir punktar: hnútpunktar netsins, rauðar stjörnur: stöðvar sem notaðar eru. Mismunur er á milli mánaða hvaða stöðvar eru virkar.

þar sem  $r$  er fjarlægð frá brúunarpunkti. Aðferðin notar allt gagnasafnið (ólíkt fyrirnefndum aðferðum) og getur því verið seinvirk og minnisfrik. Hún telst til algebrulegra brúunaradferða, en þær einkennast af því að áhrif mælinga eru vegin með ákveðnu falli, gjarnan háð fjarlægð.

Reikniritið `gsmooth2` finnur staðbundið, samfellt nálgunarfall fyrir punktasetnið í tveimur víddum með aðferð minnstu fervika. Gildi nálgunarinnar í ákveðnum punkti  $x_i, y_i$  er reiknað með því að fella annars stigs margliðu að gagnapunktunum með vægisfallinu

$$w_i = \exp\left(-\frac{(x - x_i)^2 + (y - y_i)^2}{s^2}\right).$$

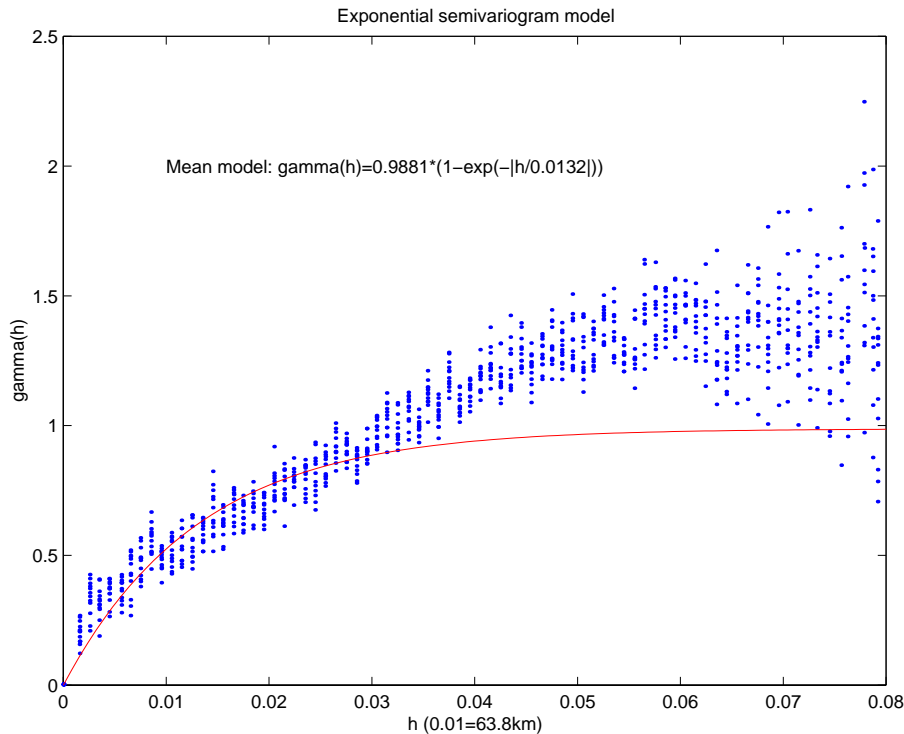
Gagnapunktur nærri punktinum þar sem fallið er reiknað hafa mest vægi en punktar sem eru fjær en 3-4 sinnum  $s$  frá punktinum hafa hverfandi áhrif á gildi nálgunarfallsins. Sérstök annars stigs margliða er fundin með aðferð minnstu fervika fyrir hvern punkt þar sem nálgunarfallið er reiknað, því er `gsmooth2` tiltölulega frekt á reikniafl tölvunnar. Nálgunarfallið er samfellt og diffranlegt sökum þess að vægi einstakra gagnapunkta breytist á samfelldan hátt með  $x, y$ . Hægt er að gefa upp hámarksfjarlægð í gagnapunkta sem tekið er tillit til þegar nálgunarfallið er reiknað og er hún gefin upp sem margfeldi af  $s$ . Þetta flýttir svolítið fyrir reikningum þegar gagnapunktar eru margir en hafa hverfandi áhrif á niðurstöðurnar ef fjarlægðin frá brúunarpunkti er meiri en ca.  $4 \cdot s$ . Fallið `gsmooth2` fyrir Matlab er skrifað af Tómasi Jóhannessyni.

Aðferðin kriging er tölfraeðileg aðferð til brúunar yfir á reglulegt net. Slíkar aðferðir finna brúunargildi sem vegið meðaltal þekktra gilda og lágmarka skekkjuna (e. *mean square error*) í brúuninni. Að sumu leyti líkjast þessar aðferðir þeim algebrulegu, en í stað þess að nota fyrirfram skilgreind stærðfræðileg föll til að veða áhrif mæligilda nota þær tölfraeðilega eiginleika gagnasafnsins.

Reikniritið **kriging** notar nálgun á dreifniriti (e. *variogram*) gagnanna sem vægisfall. Dreifniritið geymir upplýsingar um dreifni, eða breytileika, gagnanna í rúmi, bæði á smáum og stórum kvarða (sjá skilgreiningu í Kitanidis (1997) bls. 30-32). Fyrst er hrátt dreifnirit reiknað út frá gögnunum og valið líkan síðan fellt að og notað í brúuninni. Til þess að ekki þyrfti að finna nýtt líkan í hvert skipti sem brúun var gerð var svokallað veðurfræðilegt dreifnirit (e. *climatological variogram*) reiknað. Mæligögn hvers mánaðar á tímabilinu 1961-2001 voru tekin saman og líkan af dreifniriti fundið. Þannig fékkst meðalferill hvers mánaðar ársins. Mánaðarferlarnir voru svo líkir að hentugra þótti að reikna einn meðalferil í stað mánaðarlegra ferla. Þá voru reiknuð dreifnirit hvers mánaðar lögð saman og veldisvísifallið

$$\gamma(h) = A \left( 1 - e^{-\left| \frac{h}{B} \right|} \right)$$

felldu á hráu gögnunum og stuðlarnir  $A$  og  $B$  metnir með ólínulegu aðhvarfi (e. *non-linear regression*), niðurstöður eru sýndar á mynd 2. Stuðullinn  $A$  kallast sylla og



Mynd 2: Veðurfræðilegt dreifnirit. Bláir punktar: hrátt dreifnirit, rauð lína: veldisvísifallslíkan.

$B$  fylgnislengd. Þegar fjarlægð á milli mælipunkta er stærri en fylgnislengdin hefur dreifniritið náð föstu gildi, syllugildi. Það þýðir að mælingar í þessum punktum

eru nokkurn veginn óháðar. Allir punktar í meiri fjarlægð en fylgnislengd frá brúunarpunkti fá sama vægi og ef enginn punktur lendir innan fylgnislengdar verður brúunargildið meðalgildi mælinga.

Kriging reikniritið sem notað var er skrifað eftir forskrift í Kitanidis (1997) bls. 65-71. Það leysir línulegt jöfnuhneppi til að finna vægið sem hver mælipunktur fær í ákvörðun á brúunargildi. Til að spara reikninga er hér notað kriging “með hreyfanlegu nágreinni” (e. *kriging with moving neighborhood*) þannig að aðeins 3 nálægustu mælipunktar eru notaðir til brúunarinnar. Auk þess að fækka reikningum gerir þessi aðferð það að verkum að brúunargildið verður aðeins háð nálægum mælingum, en brúunin verður fyrir vikið ekki eins samfelld.

### 3 Forrit og gögn

Öll forrit og gagnaskrár er að finna á faldi í möppunni

```
/disk3/urs/ursv1/siggasif/interp2D/.
```

Úrvinnsluforritin eru skrifuð í Matlab, en auk þess voru skrifaðar tvær skeljarstefjur fyrir Unix til að sækja upplýsingar í gagnagrunninn.

#### 3.1 Skeljarstefjur

Til að ná í gögnin voru skrifaðar tvær skeljarstefjur.

**meantemp** Nær í upplýsingar um stöðvarnúmer, staðsetningu (lengd og breidd) og meðalhita valins mánaðar úr gagnagrunninum. Skráin er keyrð með eftirfarandi skipun

```
meantemp <ár> <mán> > <úttaksskrá>.
```

**semiklima** Nær í upplýsingar um stöðvarnúmer, staðsetningu, meðalhita mánaðar og ártal fyrir árabilið 1961-2001. Skráin er keyrð með

```
semiklima <mán> > <úttaksskrá>.
```

#### 3.2 Inntaksskrár í Matlab

Inntaksskrárnar sem Matlab notar eru hér útlistaðar. Þannig sleppur notandi við að keyra skeljarstefjurnar, nema hann óski sjálfur eftir því.

**medaltol.dat** Skrá sem geymir 30 ára mánaðarmedaltöl hitastigs á mönnum veðurathugunarstöðvum og er á forminu

```

0          janúar   febrúar   ...
stöðvarnúmer meðaltal meðaltal ...
:          meðaltal ...
:

```

`data/meantemp` Mappa sem geymir hitastigsskrár (fengnar með `meantemp`) og heita `<x>.dat`, þar sem `x` er nafn mánaðar auk árs, t.d. `jan60.dat`. Formið á skránum er

```

stöðvarnúmer lengd breidd meðalhiti.

```

`data/beta` Hér eru geymdar upplýsingar um dreifnisfall mánaða. Skrárnar bera nafnið `beta_<month>.dat` þar sem `month` er enskt heiti mánaðarins. Auk þess er þar að finna meðalferil allra mánaða `beta.dat`.

`data/semi` Mappa sem geymir upplýsingar um meðalhita á öllum stöðvum á árunum 1961 til 2001, fyrir hvern mánuð ársins. Skrárnar bera nöfnin `<month>.dat` þar sem `month` er enskt heiti mánaðar. Þær eru á forminu

```

stöðvarnúmer lengd breidd meðalhiti ártal

```

### 3.3 Matlab stefjur

Listi yfir Matlab stefjur sem notaðar voru. Lýsing á virkni þeirra og hvernig kalla skuli á þau er að finna fremst í sjálfum skránum (sem allar enda á `.m`). Hana má einnig fá með því að skrifa `help <nafn stefju>` í skipanaglugga Matlab.

`init` Les inn mánaðarskrá og skilar af sér staðsetningu (`x,y`) í Lambert hnitum, hitafrávik, `fravik`, fylkjunum XI og YI sem innihalda hnit hnútpunkta netsins sem brúað er á, dreifniritlíkani, `beta` og `range (s)` fyrir `gsmooth2`. Nafn mánaðarskrár og mánaðar þarf að gefa.

`kriging` Brúar yfir á reglulegt net með aðferð kriging aðferðinni. Staðsetningar verða að vera í Lambert hnitum.

`krigConf` Brúar yfir á reglulegt net með kriging aðferðinni og gefur einnig vikmörk (e. *confidence interval*). Staðsetningar verða að vera í Lambert hnitum.

`gsmooth2` Brúar með því að fella 2.stigs flöt að mælingum.

`errcalc` Reiknar út skekkju og vikmörk í brúun. Notar `krigConf`

`errcalc1` Reiknar út skekkju í brúun. Notar `kriging`

<code>cut</code>	Skert út gildi utan strandlínunar og á hálendinu.
<code>xfigure</code>	Teiknar kort af brúun með fylltum jafngildislínum.
<code>SemiClim</code>	Reiknar út veðurfarsfræðilegt dreifnirit. Notar til þess <code>rawsemi</code> .
<code>rawsemi</code>	Reiknar hrátt dreifnirit gagna.
<code>nlreg_e</code>	Fellir veldisvísis líkan að gögnum.
<code>nlreg_g</code>	Fellir Gauss líkan að gögnum.
<code>nlreg_s</code>	Fellir spherical líkan að gögnum.
<code>f4plot</code>	Forrit sem teiknar kort með niðurstöðum aðferðanna fjögurra og reiknar skekkju í brúun.

### 3.4 Skipanaröð

Til þess að keyra forritin eru eftirfarandi skipanir skrifaðar í skipanaglugga Matlab. Hér er sýnt dæmi fyrir kringing, en ef önnur aðferð er notuð er viðeigandi skipun sett í línu 2. Þær eru einnig tilgreindar.

```
[x,y,dT,XI,YI,beta,range] =init('month','filename');
```

```
[Z,S2] =kriging(x,y,dT,XI,YI,beta,minpnt);
```

eða

```
[Z,S2,conf] =krigConf(x,y,dT,XI,YI,beta,minpnt,stdev);
```

þar sem `minpnt` er fjöldi punkta sem nota á í brúuninni og `stdev` er staðalfrávik gagnanna fengið með

```
stdev=std(dT);
```

Því næst er ritað

```
Z=cut(XI,YI,Z);
```

```
xfigure(XI,YI,Z);
```

Ef reikna á skekkju í `kriging` er skrifað

```
[rms,minerr,maxerr] =errcalc1(x,y,dT,range,beta,minpnt,method);
```

eða ef nota á `krigConf`

```
[rms,minerr,maxerr,zo,conf_int,inside] =  
errrcalc(x,y,dT,range,beta,minpnt,method);
```

þar sem `method` er aðferðin sem notuð er, í þessu tilfelli `'kriging'`. Mikilvægt er að nafn aðferðar sé innan einfaldra gæsalappa.

Aðrar aðferðir er kallað á með eftirfarandi skipunum

```
Z=gsmooth2(x,y,dT,range,XI,YI);  
  
Z=griddata(x,y,dT,XI,YI,'linear');  
  
Z=griddata(x,y,dT,XI,YI,'cubic');
```

Þegar reikna á veðurfarsfræðilegt dreifnirit er kallað á

```
[stH,stSV]=SemiClim('month');
```

þar sem `month` er enskt nafn mánaðar. Til að finna jöfnu dreifnifallsins er skrifað

```
beta=nlreg_x(stH,stSV);
```

þar sem  $x$  tilgreinir líkanið sem óskað er eftir, en þau eru:

- $x = s$  : Veldisvísifall (e. exponential model).
- $x = g$  : Gaussfall (e. gaussian model).
- $x = s$  : Kúlufall (e. spherical model).

Til að teikna mynd af niðurstöðum aðferðanna fjögurra og reikna skekkjuna er skrifað

```
f4plot;
```

en fyrst þarf að laga reikniritið og gefa mánaðarnafnið sem um ræðir (á ensku) og nafn skránnar sem geymir hitaupplýsingar mánaðar.

## 4 Niðurstöður

### 4.1 Skekkjumat

Skekkjan er metin á eftirfarandi hátt. Eitt gildi er plokkað út, og síðan er valið brúunarreiknirit keyrt á eftirstandandi mæligildi, til að finna spágildi í útplokkaða punktinum. Þetta er gert fyrir hvern og einn þekktan punkt. Því næst er skekkjan reiknuð skv.

$$\sigma_{rms} = \sqrt{\frac{\sum_{i=1}^N (\hat{z}_i - z_i)^2}{N}}$$

þar sem  $N$  er fjöldi þekktra punkta,  $\hat{z}_i$  brúunargildi og  $z_i$  þekkt gildi. Auk þess er hámarksskekkjan reiknuð

$$\max = \max(|\hat{z}_i - z_i|).$$



Aðferð	Ár	Mán	$\sigma_{rms}$	max	Ár	Mán	$\sigma_{rms}$	max
kriging	1960	jún	0.345	0.951	1965	Jan	<b>0.245</b>	0.634
gsmooth2		61 st.	0.510	2.440		67 st.	0.248	<b>0.523</b>
griddata linear			<b>0.325</b>	<b>0.889</b>			0.254	0.553
griddata cubic			0.334	0.929			0.264	0.569
kriging	1969	ágú	0.330	<b>0.881</b>	1972	Des	<b>0.304</b>	<b>0.708</b>
gsmooth2		72 st.	0.388	0.961		74 st.	0.353	1.139
griddata linear			<b>0.325</b>	1.427			0.339	0.922
griddata cubic			0.341	1.385			0.351	0.966
kriging	1978	mar	<b>0.275</b>	0.665	1983	Júl	<b>0.317</b>	1.053
gsmooth2		78 st.	0.307	0.765		75 st.	0.363	1.130
griddata linear			0.276	<b>0.594</b>			0.342	1.063
griddata cubic			0.290	0.630			0.354	<b>1.050</b>
kriging	1989	feb	<b>0.334</b>	<b>0.758</b>	1992	Ágú	0.244	0.779
gsmooth2		80 st.	0.352	1.027		70 st.	<b>0.239</b>	0.819
griddata linear			0.354	0.778			0.243	<b>0.713</b>
griddata cubic			0.377	0.867			0.255	0.751
kriging	1996	jan	0.337	0.942	2000	Des	0.359	<b>0.818</b>
gsmooth2		57 st.	0.360	1.036		48 st.	0.506	1.971
griddata linear			0.329	<b>0.792</b>			<b>0.346</b>	0.883
griddata cubic			<b>0.329</b>	0.807			0.347	0.903

Tafla 1: Skekkja í kringing, gsmooth2, griddata linear og griddata cubic. Fyrir alla mánuði er  $s = 0.01$  (gsmooth2).

Aðferð	$\sigma_{rms}$	max
kriging	0.309	0.819
gsmooth2	0.363	1.181
griddata linear	0.313	0.861
griddata cubic	0.324	0.886

Tafla 2: Meðalskekkja í kringing, gsmooth2, griddata linear og griddata cubic

Niðurstöður er að finna í töflu 1 og eru minnstu gildi feitletruð.

Þar sem að gögnin fyrir einstaka mánuði geta verið misgóð tölfræðilega séð, gefur það réttari mynd ef meðalskekkja allra mánaða er reiknuð (tafla 2). Kringing hefur þar vinninginn hvað varðar  $\sigma_{rms}$ , þó svo að lítil munur sé á kringing og báðum griddata aðferðunum.

Villudreifingin er sýn á myndum 3-6. Kort voru teiknuð fyrir hvern mánuð og hverja aðferð og eru þau sýnd á myndum 7-16.

## 4.2 Umræða

Kringing hefur oftast minnstu skekkju og stendur jafnfætis linear hvað varðar minnstu hámarksskekkju. Auk þess eru meðalskekkja og meðalhámarksskekkja

minnstar. Fast á hæla kriging fylgir **linear**. Þrátt fyrir litla skekkju er aðferðin óhentug annars vegar vegna þess hversu ósamfelld hún er og hins vegar vegna þess að hún bryggjar ekki. Það þýðir að hún brúar ekki út fyrir þá línu sem tengir útpunkta svæðisins. Meðfram strandlínu landsins verður kortið því autt. Það jákvæða við **linear** er hversu einföld aðferðin er, svo hægt er að vera nokkuð viss um að hún spinnur ekki upp fyrirbæri sem gögnin ekki gefa tilefni til. Útkoman úr **cubic** er mjög svipuð og úr **linear**, enda náskyldar aðferðir. Skekkjan er þó yfirleitt örlítið hærri. Aðferðin **cubic** gefur samfelldari brúun sem lítur betur út á mynd. Þriðja stigs brúanir eru viðkvæmar fyrir óreglulegri dreifingu mælinga, en það á einmitt við um gagnasafnið sem hér er notað. Það gildir einnig um **cubic** að ekki er brúað út fyrir endimörk gagnasafns. Aðferðin **gsmooth2** kemur almennt ekki vel út, hún gerir of mikla einföldun vegna fátæktar punktásafns. Aðferðin tekur líklega of mikið tillit til punkta í mikilli fjarlægð, enda þarf a.m.k. 6 punkta til að finna 2.gráðu margliðu á fleti.

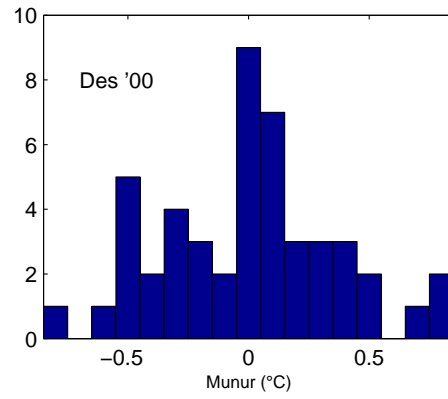
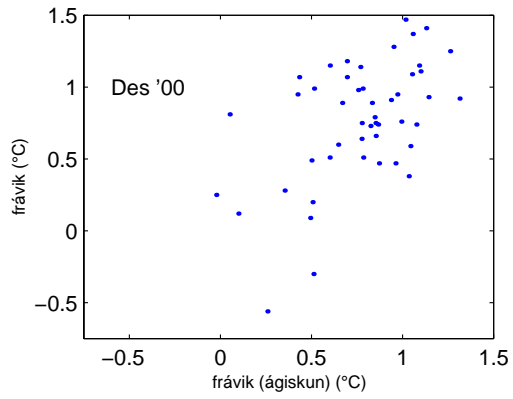
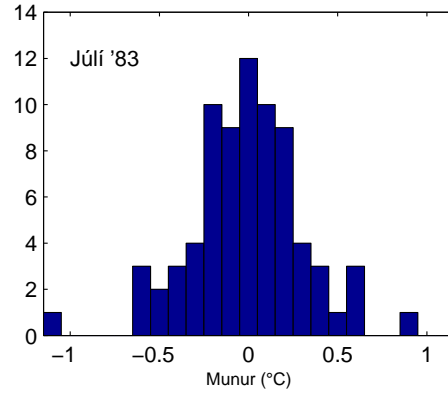
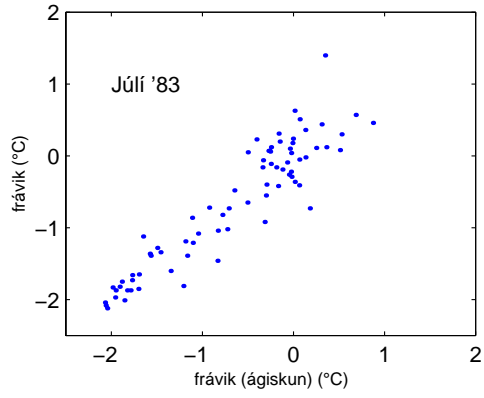
Þess ber að geta að í brúuninni eru allar mælistöðvar notaðar, einnig þær sem staðsettar eru á hálendinu. Á myndunum hafa niðurstöður fyrir hálendið einfaldlega verið klipptar út.

## 5 Lokaorð

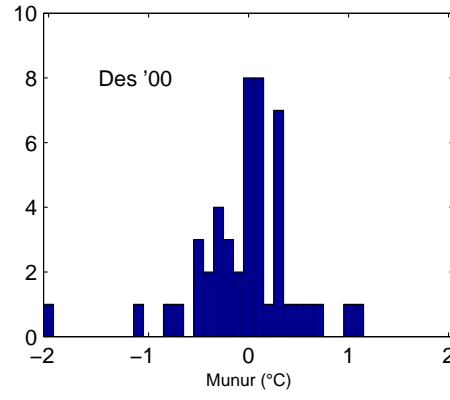
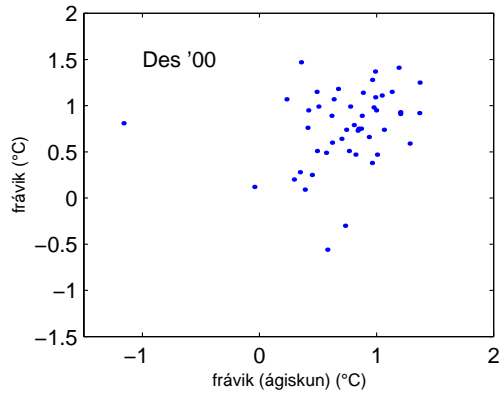
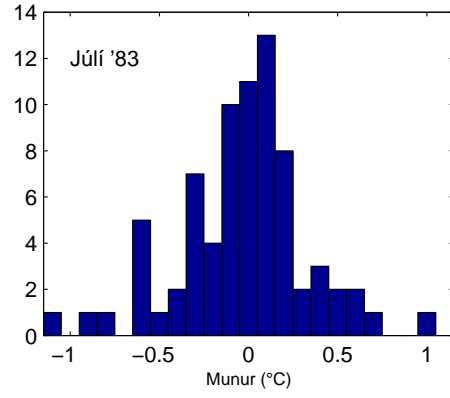
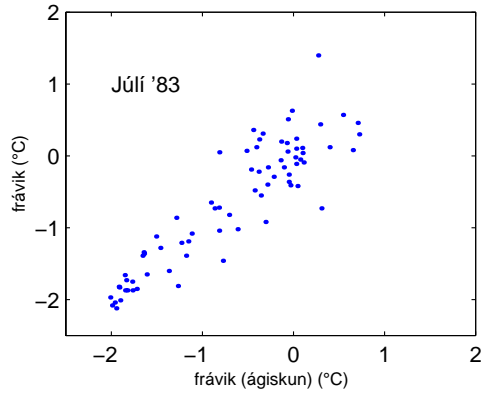
Byggt á þessum samanburði koma **kriging** og **linear** helst til greina, til úrvinnslu á þeim gögnum sem hér eru notuð. Aðferðin **gsmooth2** hentar best þegar um þéttar mælingar er að ræða, en því er ekki að skipta með hitamælingar á Íslandi. Almennt þykir **kriging** nákvæmust til kortlagningar óreglulegra, tvívíðra gagna. Sú aðferð gefur almennt litla skekkju, minnstu meðalskekkju og minnstu hámarksskekkju, svo bilið sem skekkjan lendir á er minnst. Hún hentar því mjög vel fyrir þau not sem hér ræðir um.

Auk þeirra rita sem að framan er vitnað í reynast eftirfarandi góð til glöggvunar á umfjöllunarefninu: Davis (1986); Gerald and Wheatley (1999); Hanselman and Littlefield (1997); Gustavsson (1981); Middleton (2000).

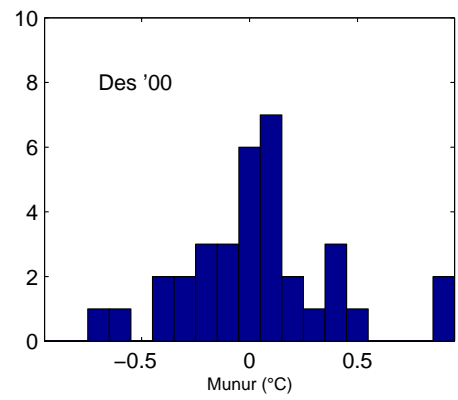
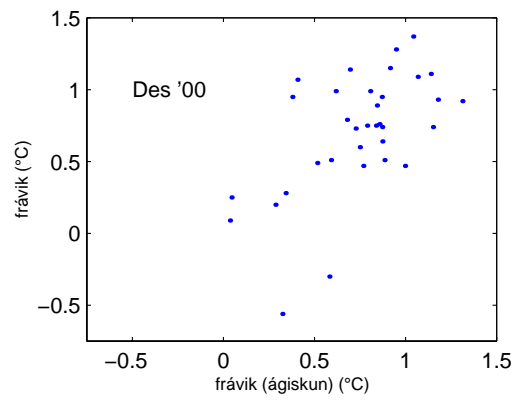
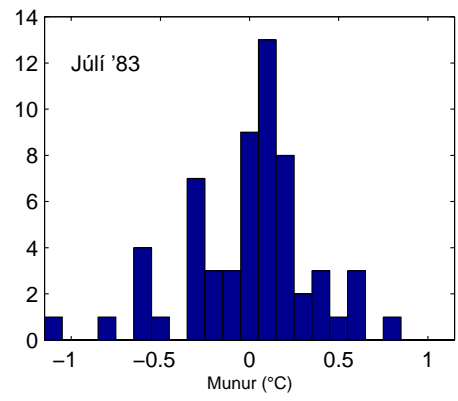
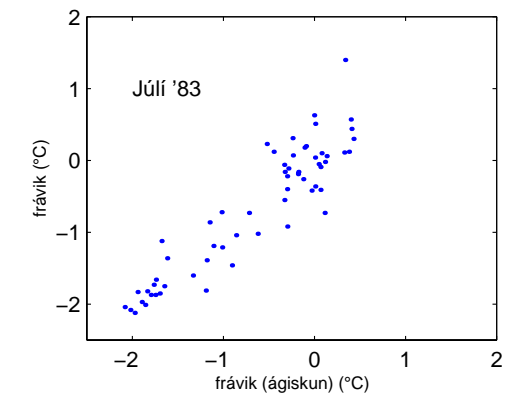
Dr. Philippe Crochet eru færðar bestu þakkir fyrir fyrir góð ráð og ábendingar.



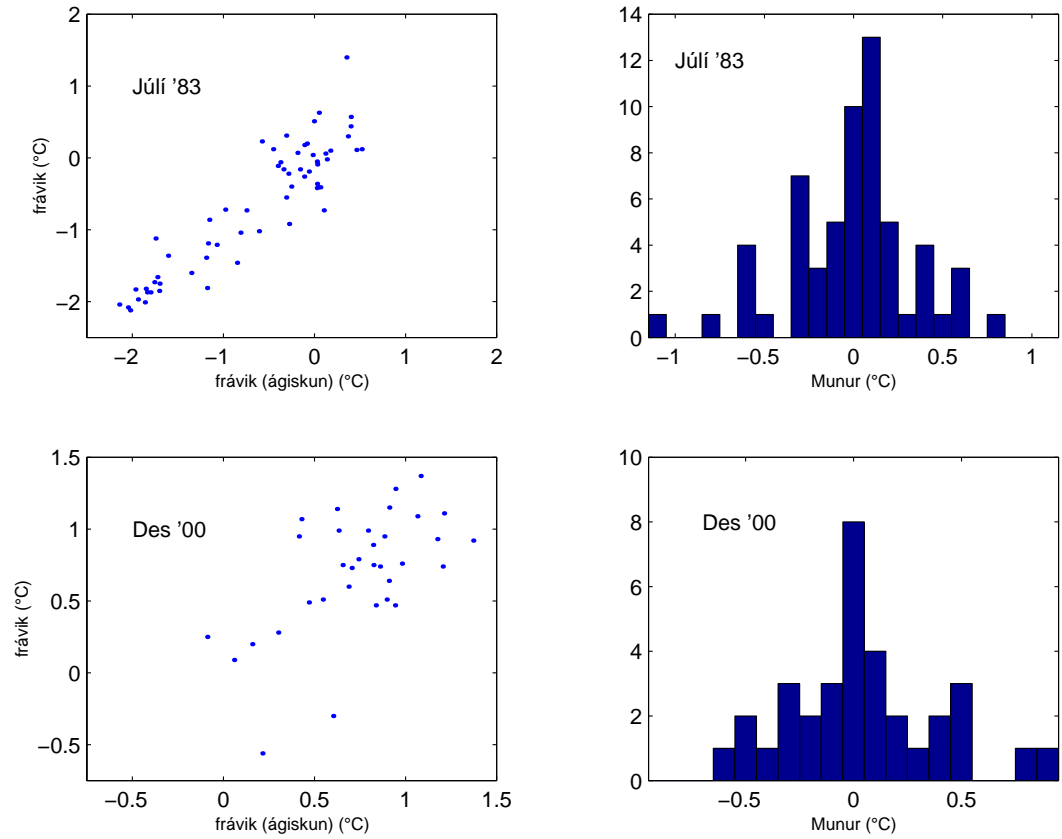
Mynd 3: Skekkjudreifing í kringing.



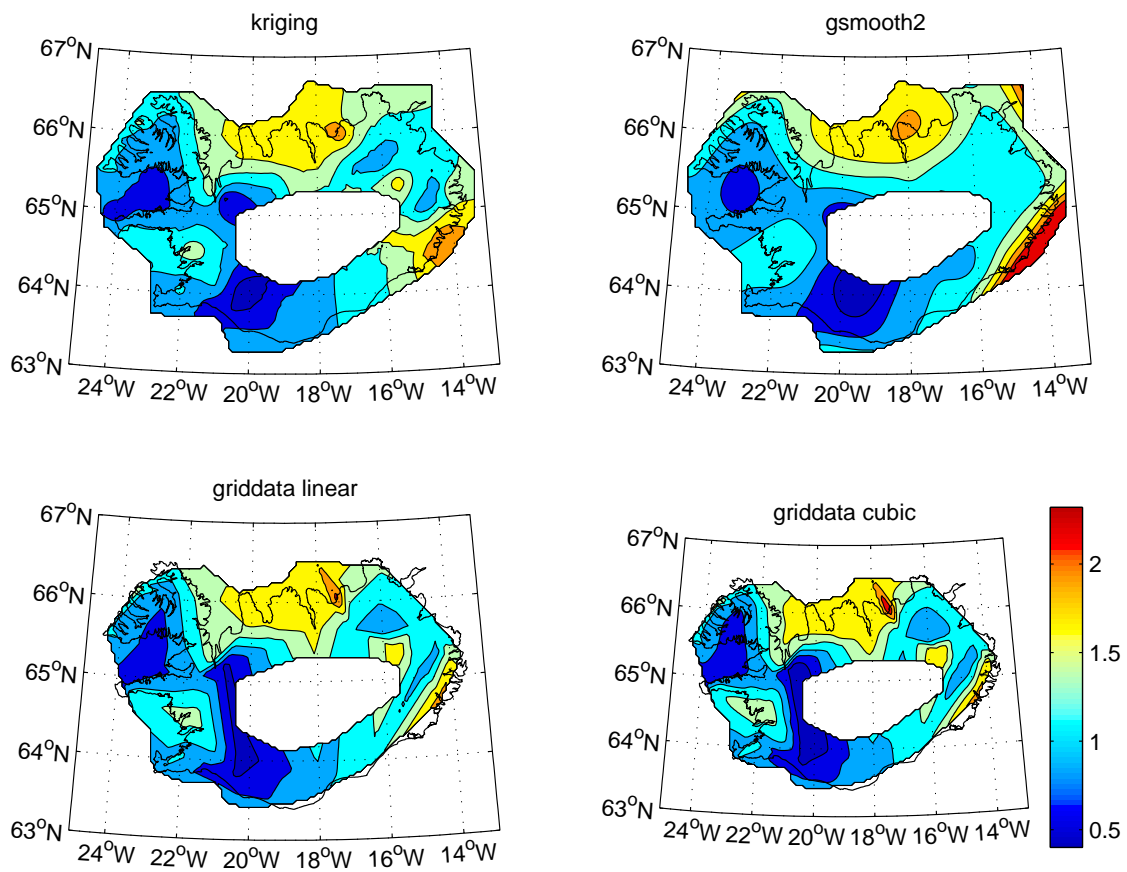
Mynd 4: Skekkjudreifing í `gsmooth2`.



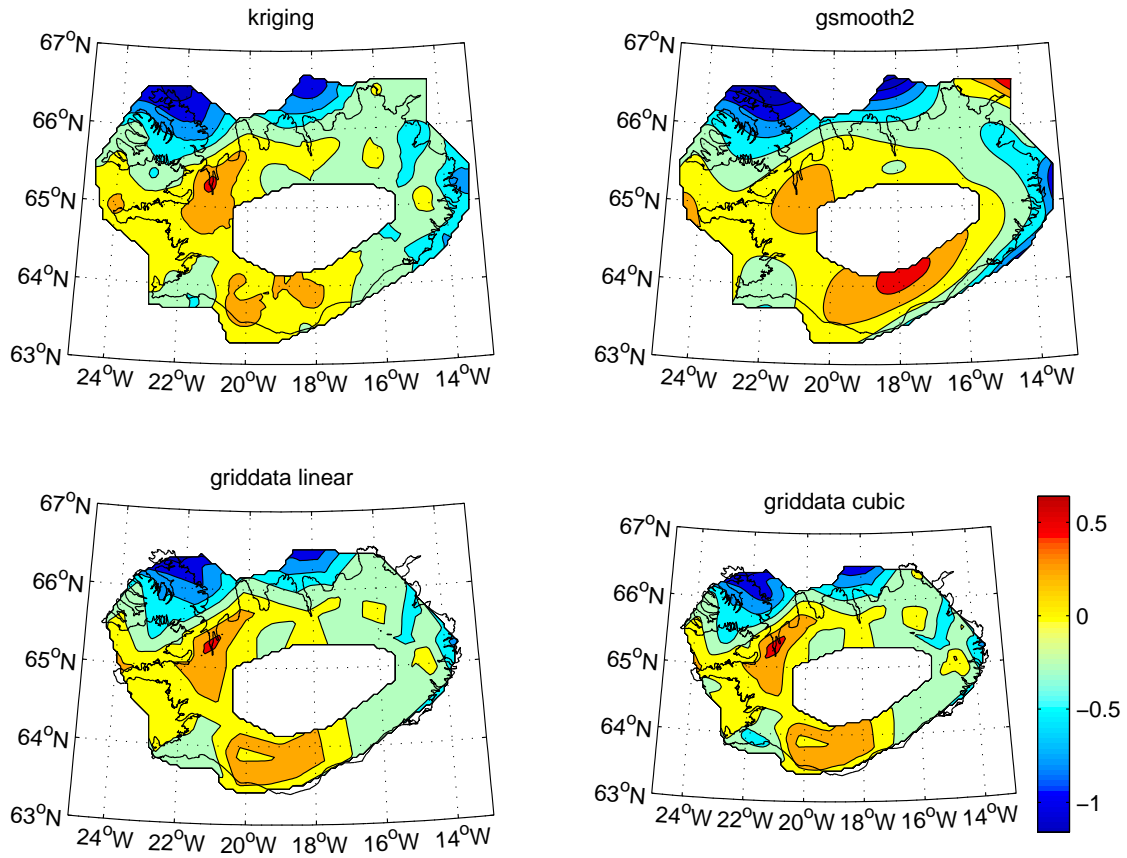
Mynd 5: Skekkjudreifing í griddata linear.



Mynd 6: Skekkjudreifing í griddata cubic.

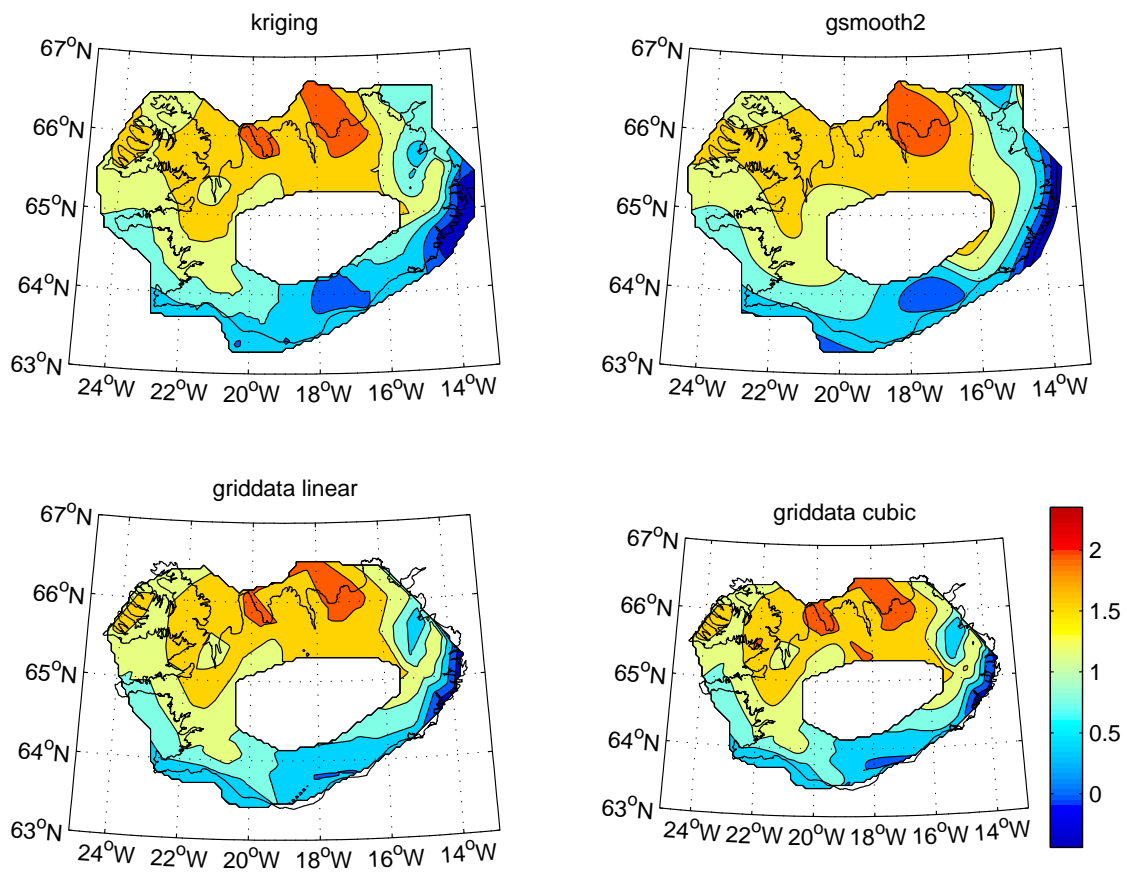


Mynd 7: Júní 1960.

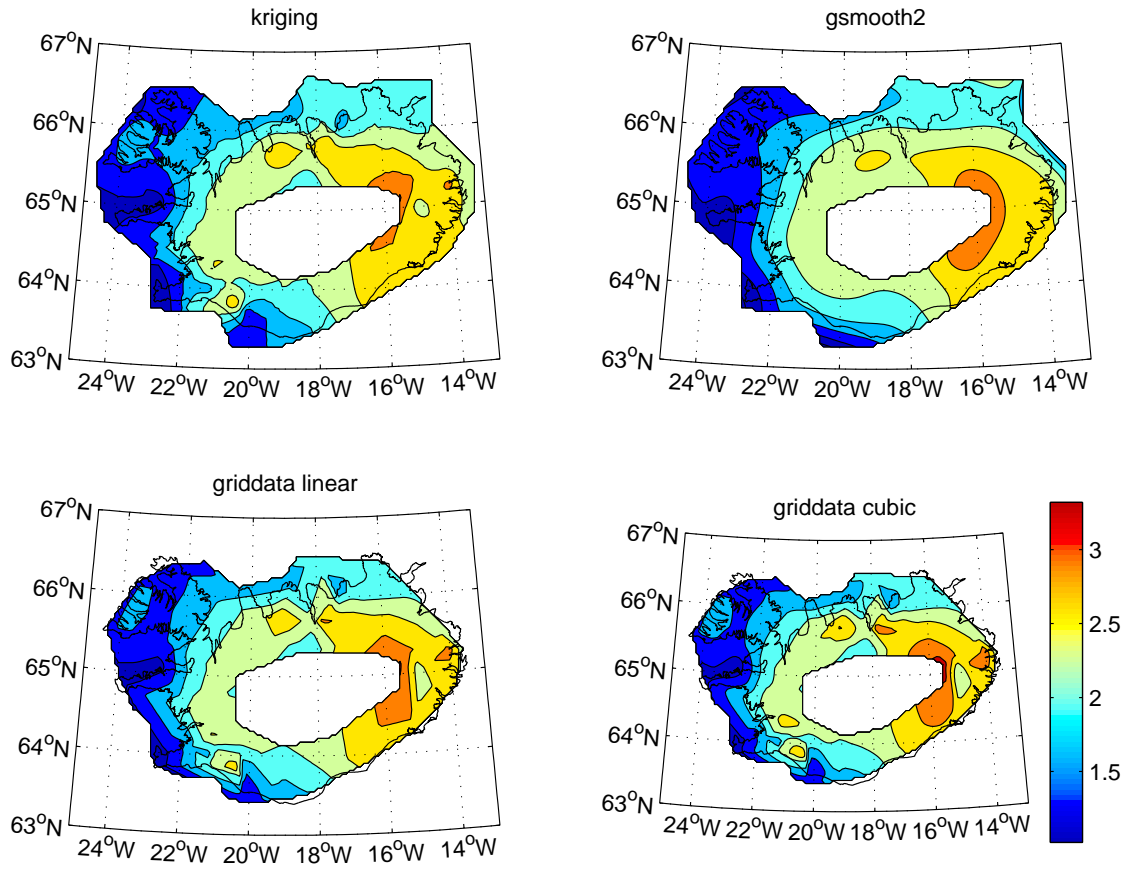


Mynd 8: Janúar 1965.

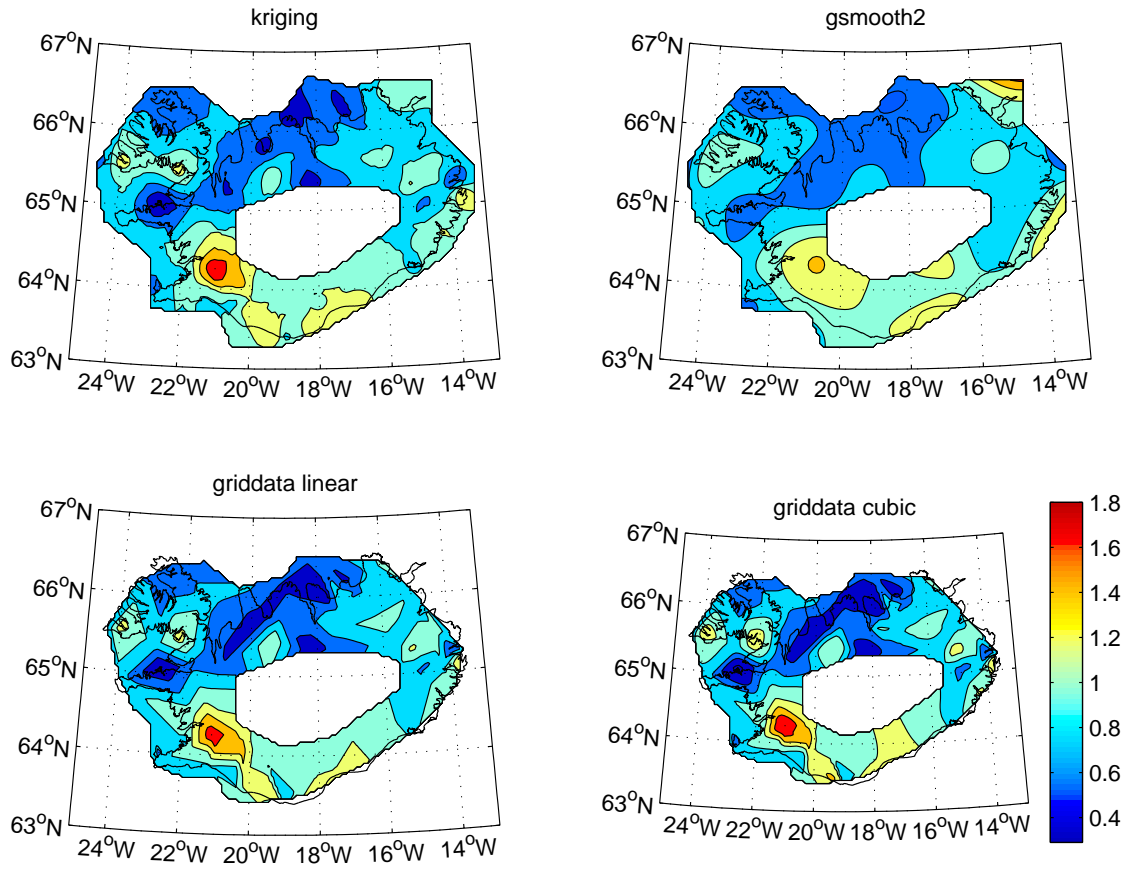




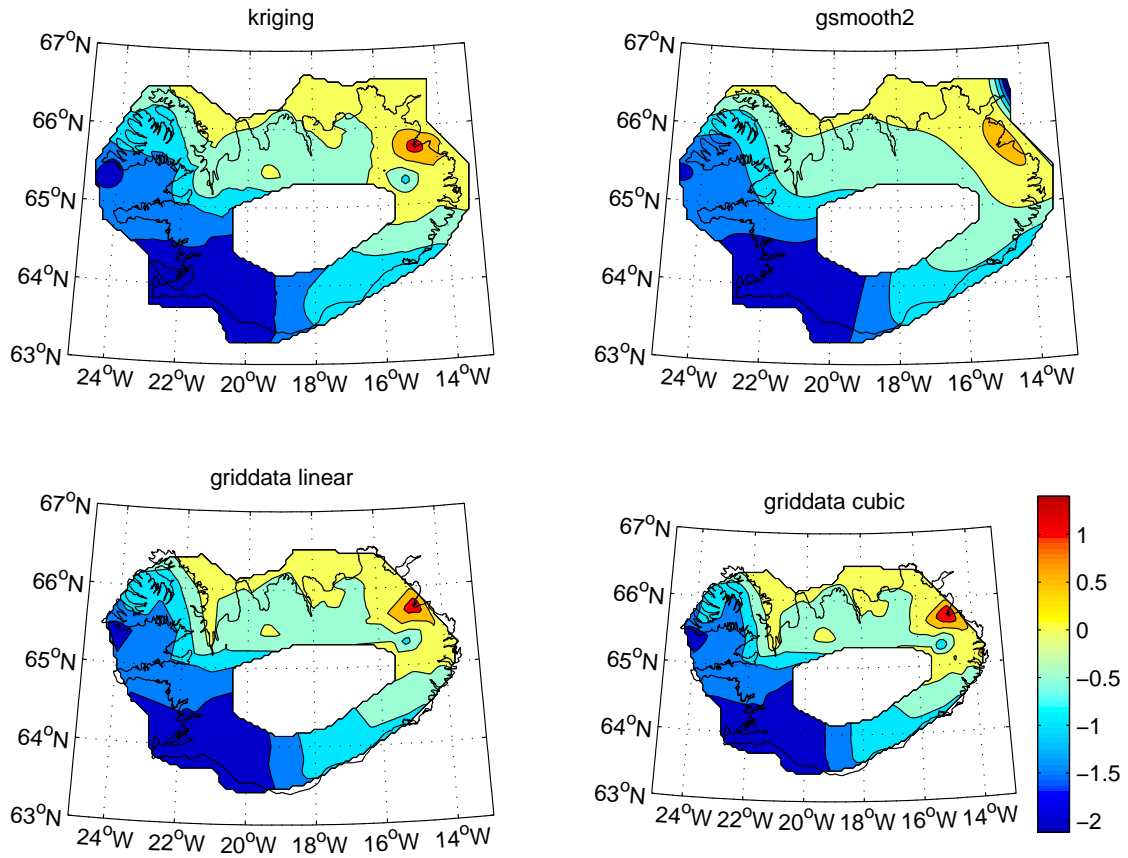
Mynd 9: Ágúst 1969.



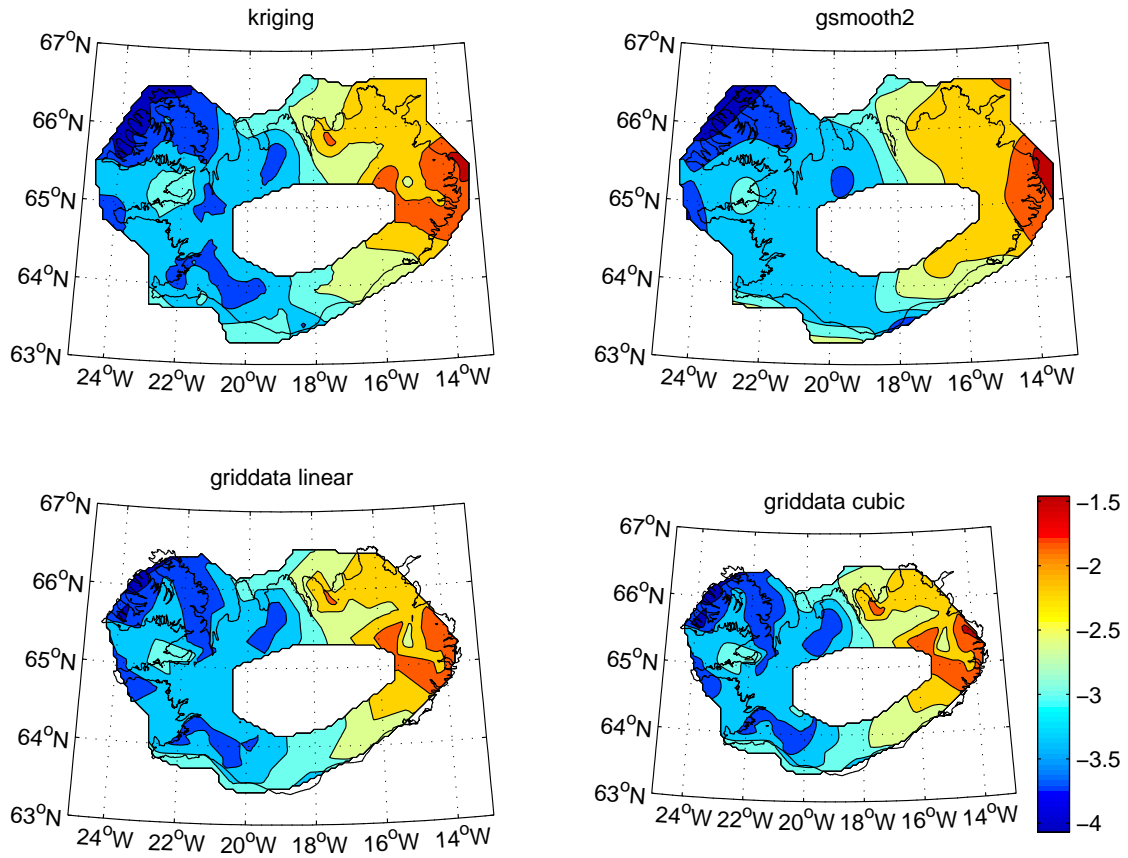
Mynd 10: Desember 1972.



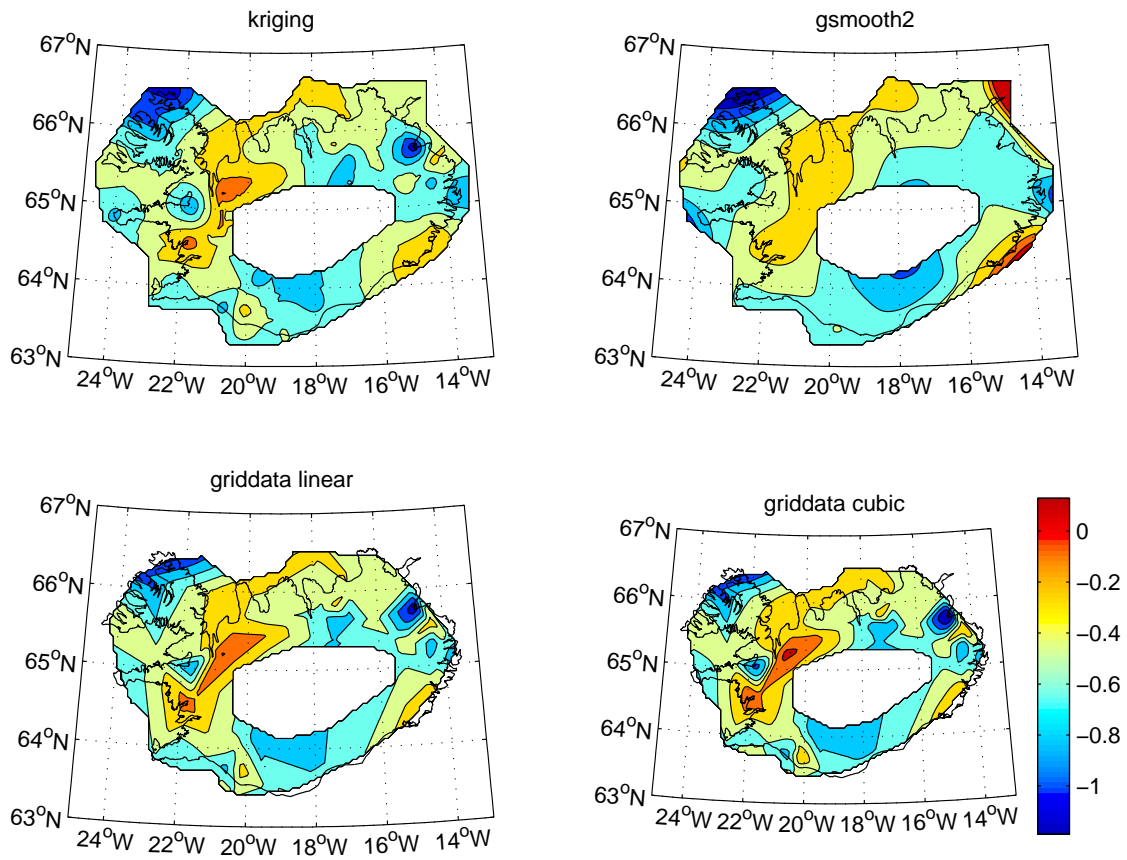
Mynd 11: Mars 1978.



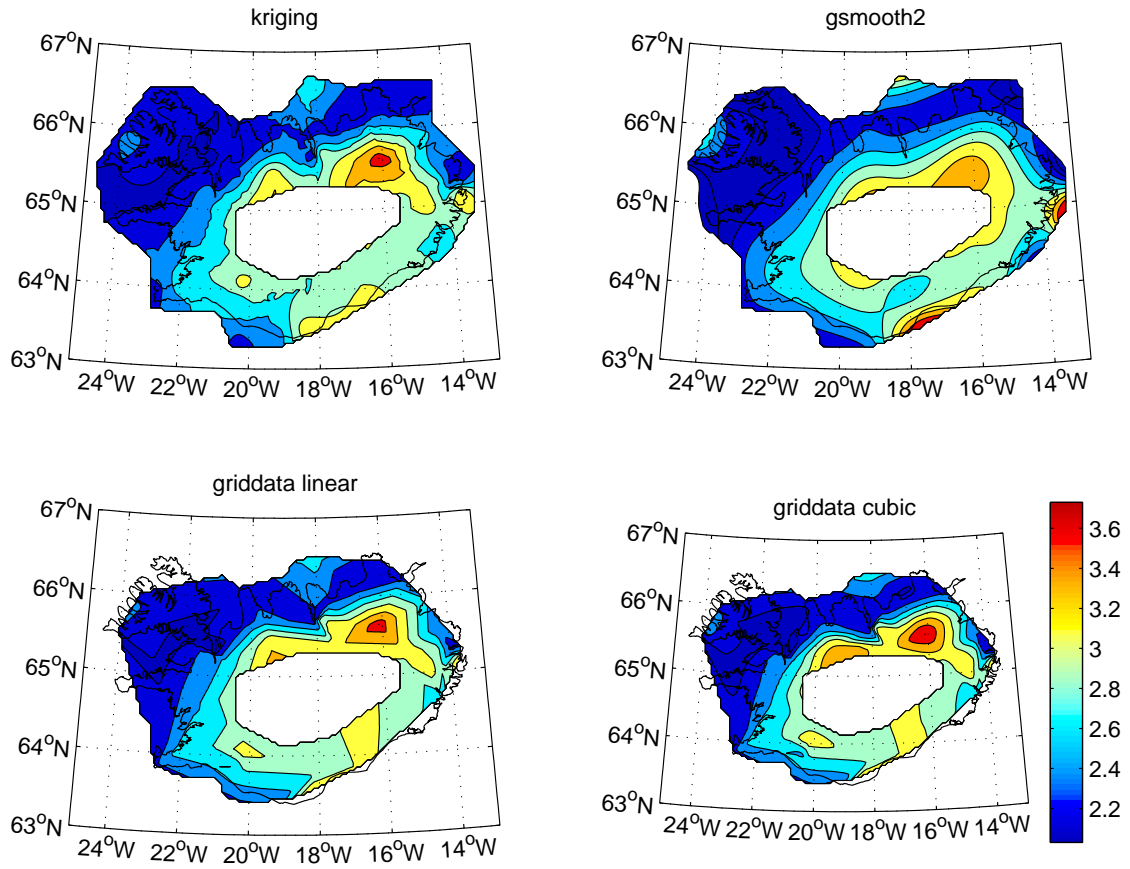
Mynd 12: Júlí 1983.



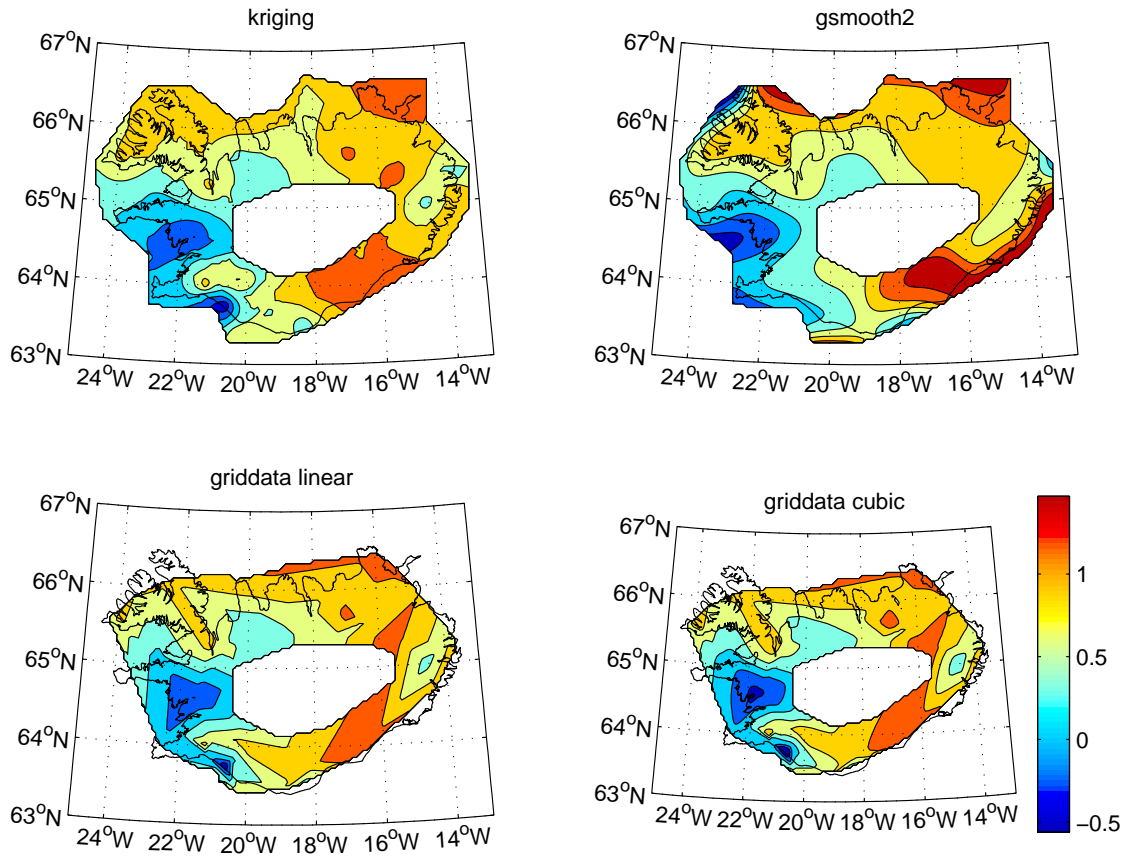
Mynd 13: Febrúar 1989.



Mynd 14: Ágúst 1992.



Mynd 15: Janúar 1996.



Mynd 16: Desember 2000.



## Heimildir

- Davis, J. C. (1986). *Statistics and Data Analysis in Geology*. John Wiley & Sons, Inc., New York.
- Gerald, C. F. and Wheatley, P. O. (1999). *Applied Numerical Analysis*. Addison Wesley Longman, Inc., Reading.
- Gustavsson, N. (1981). A review of methods for objective analysis. In L. Bengtsson, M. G. and Källén, E., editors, *Dynamic Meteorology : Data Assimilation Methods*, pages 17–76. Springer–Verlag New York, Inc., New York.
- Hanselman, D. and Littlefield, B. (1997). *The Student Edition of MATLAB : version 5, user's guide*. Prentice-Hall, Inc., New Jersey.
- Kitanidis, P. K. (1997). *Introduction to Geostatistics : Applications in Hydrogeology*. Cambridge University Press, Cambridge.
- Middleton, G. (2000). *Data Analysis in the Earth Sciences using Matlab*. Prentice-Hall, Inc., New Jersey.