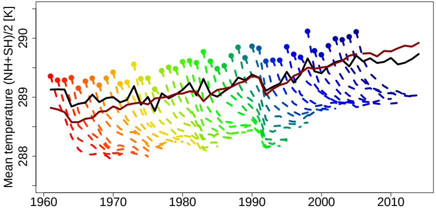


Decadal Prediction

Drift as a Particular Challenge for Verification

Henning Rust

Institut für Meteorologie, Freie Universität Berlin

7th Int. Verification Methods Workshop, Berlin, May 11th, 2017

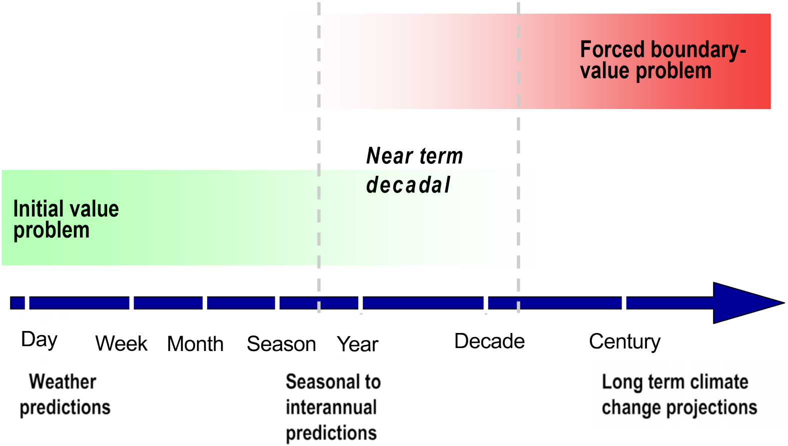
Decadal climate prediction . . .

. . . provides information about the future evolution of the statistics of regional climate from the output of a numerical model that has been initialized with observations . . .

[1]

[1] cited from Meehl et al. [2014]

Decadal climate prediction . . .

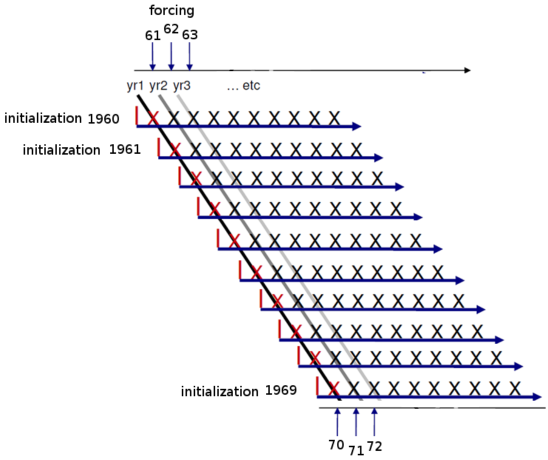


2

[2] taken from Boer et al. [2016]

Initialization with “observations”

Hindcast set: initialize every year, 10-yr hindcast each



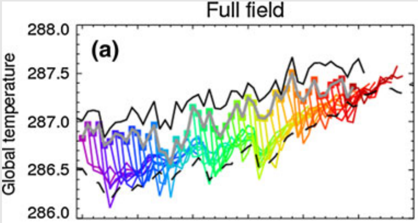
courtesy of Jens Grieger

Initialization with “observations”

Hindcast set: initialize every year, 10-yr hindcast each

Full-field

assimilate directly

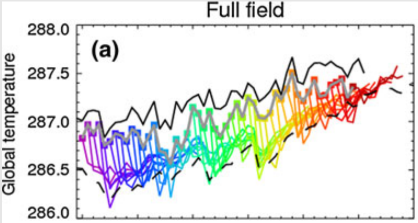


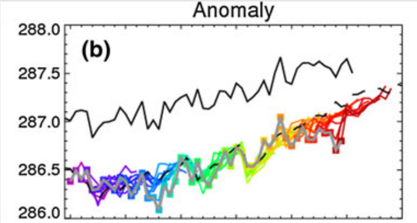
Annual mean global temperature,
taken from Smith et al. [2013]

Initialization with "observations"

Hindcast set: initialize every year, 10-yr hindcast each

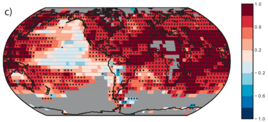
Full-field	Anomaly
assimilate directly	assimilate anomalies

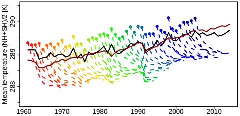




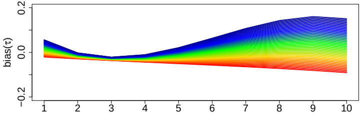
Annual mean global temperature,
taken from Smith et al. [2013]

Verification of decadal prediction

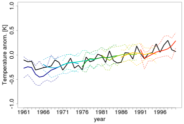




Drift

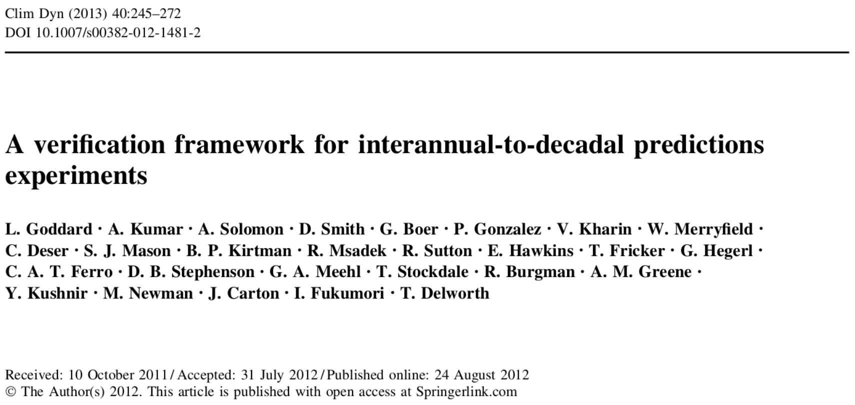


Drift adjustment



Re-Calibration

A framework

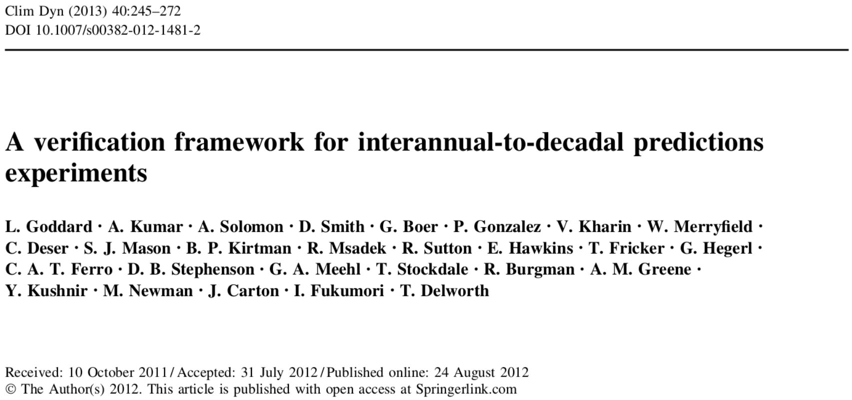

Clim Dyn (2013) 40:245–272
DOI 10.1007/s00382-012-1481-2

A verification framework for interannual-to-decadal predictions experiments

L. Goddard · A. Kumar · A. Solomon · D. Smith · G. Boer · P. Gonzalez · V. Kharin · W. Merryfield · C. Deser · S. J. Mason · B. P. Kirtman · R. Msadek · R. Sutton · E. Hawkins · T. Fricker · G. Hegerl · C. A. T. Ferro · D. B. Stephenson · G. A. Meehl · T. Stockdale · R. Burgman · A. M. Greene · Y. Kushnir · M. Newman · J. Carton · I. Fukumori · T. Delworth

Received: 10 October 2011 / Accepted: 31 July 2012 / Published online: 24 August 2012

A framework


Clim Dyn (2013) 40:245–272
DOI 10.1007/s00382-012-1481-2

A verification framework for interannual-to-decadal predictions experiments

L. Goddard · A. Kumar · A. Solomon · D. Smith · G. Boer · P. Gonzalez · V. Kharin · W. Merryfield · C. Deser · S. J. Mason · B. P. Kirtman · R. Msadek · R. Sutton · E. Hawkins · T. Fricker · G. Hegerl · C. A. T. Ferro · D. B. Stephenson · G. A. Meehl · T. Stockdale · R. Burgman · A. M. Greene · Y. Kushnir · M. Newman · J. Carton · I. Fukumori · T. Delworth

Received: 10 October 2011 / Accepted: 31 July 2012 / Published online: 24 August 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Q1: Predictions more accurate due to initialization?

Q2: Does ensemble spread appropriately represents uncertainty?

Verification: What can we expect?[3]

Temporal scales

annual/seasonal averages

- **1yr** lead-year 1
- **4yrs** lead-years 2-5, 6-9
- **8yrs** lead-years 2-9

more are preferable

[3] Decadal prediction verification framework Goddard et al. [2013]

Verification: What can we expect?[3]

Temporal scales	Spatial scales
annual/seasonal averages **1yr** lead-year 1 **4yrs** lead-years 2-5, 6-9 **8yrs** lead-years 2-9 more are preferable	scale of reference or larger, e.g. **Temp** 5°×5° **Precip** 2.5°×2.5° depends on study

[3]Decadal prediction verification framework Goddard et al. [2013]

Verifying ensemble predictions: Accuracy of mean

Ensemble mean

$$H_{j\tau} = \frac{1}{N_e} \sum_{i=1}^{N_e} H_{ij\tau}$$

$H_{ij\tau}$ ens. member i, initialization j, lead-year τ

Verifying ensemble predictions: Accuracy of mean

Ensemble mean

$$H_{j\tau} = \frac{1}{N_e} \sum_{i=1}^{N_e} H_{ij\tau}$$

$H_{ij\tau}$ ens. member i, initialization j, lead-year τ

Q1: More accurate due to initialization?

$$MSESS = 1 - \frac{MSE_H}{MSE_R}$$

historicals (no initialization but forcing) as reference

Verifying ensemble predictions: Spread

Ensemble Spread

$$\sigma^2_{H_j} = \frac{1}{N_e - 1} \sum_{i=1}^{N_e} \left(H_{ij\tau} - H_{j\tau} \right)^2$$

Verifying ensemble predictions: Spread

Ensemble Spread

$$\sigma^2_{H_j} = \frac{1}{N_e - 1} \sum_{i=1}^{N_e} \left(H_{ij\tau} - H_{j\tau}\right)^2$$

Q2: Does ensemble spread appropriately represents uncertainty?

$$CRPS(\mathcal{N}(\hat{H}_j, \overline{\sigma^2}_H), O_j)$$

Gaussian hindcast distribution[a]
conditional and unconditional bias adjusted ($\hat{H}_j$)

[a] Gneiting and Raftery [2007]

Verifying ensemble predictions: Spread

Ensemble Spread

$$\sigma^2_{H_j} = \frac{1}{N_e - 1} \sum_{i=1}^{N_e} \left(H_{ij\tau} - H_{j\tau}\right)^2$$

Q2: Does ensemble spread appropriately represents uncertainty?

$$CRPSS = 1 - \frac{\overline{CRPS_H}}{\overline{CRPS_R}}$$

MSE as variance of ref. fore-
cast, CRPSS=0 is optimal!

Verifying ensemble predictions: Spread

Ensemble Spread

$$\sigma^2_{H_j} = \frac{1}{N_e - 1} \sum_{i=1}^{N_e} \left(H_{ij\tau} - H_{j\tau}\right)^2$$

Q2: Does ensemble spread appropriately represents uncertainty?

$$CRPSS = 1 - \frac{\overline{CRPS_H}}{\overline{CRPS_R}}$$

MSE as variance of ref. forecast, CRPSS=0 is optimal!

Verifying ensemble predictions: Spread

Ensemble Spread

$$\sigma^2_{H_j} = \frac{1}{N_e - 1} \sum_{i=1}^{N_e} \left(H_{ij\tau} - H_{j\tau}\right)^2$$

Q2: Does ensemble spread appropriately represents uncertainty?

$$CRPSS = 1 - \frac{\overline{CRPS_H}}{\overline{CRPS_R}}$$

MSE as variance of ref. fore-cast, CRPSS=0 is optimal!

$$LESS = \ln\left(\frac{\overline{\sigma^2}_H}{MSE}\right)$$

Additionally, logarithmic ensemble spread score e.g. Kadow et al. [2014]

Small ensembles and significance

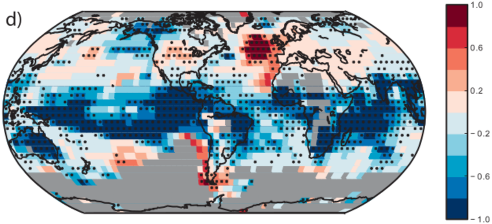
MiKlip ensembles

baseline0 3

baseline1 10

prototype 15 + 15

preop $\leq$ 15

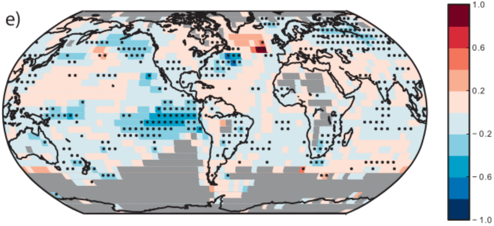


Goddard et al. [2013] suggest a bootstrap for significance

Small ensembles and significance

MiKlip ensembles

baseline0 3
baseline1 10
prototype 15 + 15
preop $\leq$ 15



Goddard et al. [2013] suggest a bootstrap for significance

Small ensembles and significance

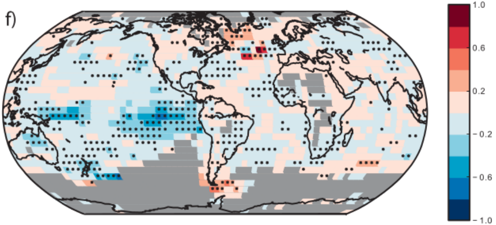
MiKlip ensembles

baseline0 3

baseline1 10

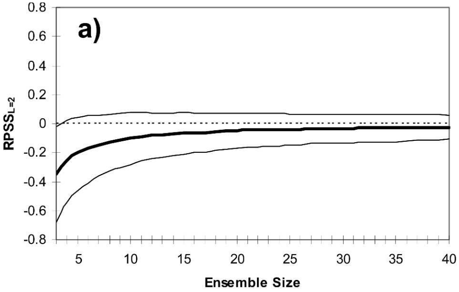
prototype 15 + 15

preop $\leq$ 15



Goddard et al. [2013] suggest a bootstrap for significance

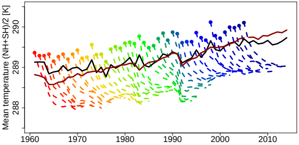
Small ensembles and biased scores



[4]

- bias corrected scores Ferro et al. [2008]
- application with RPS Kruschke et al. [2015]
- implemented in R-package `SpecsVerification` (Stefan Siegert)

[4]taken from Müller et al. [2005]



Mean temperature (NH+SH)/2 [K]
288
289
290
1960
1970
1980
1990
2000
2010

Drift

“Bias” or mean difference

$$ME = \frac{1}{N}\sum_{j=1}^{N} H_j - \frac{1}{N}\sum_{j=1}^{N} O_j$$

“Bias” or mean difference

$$ME = \frac{1}{N}\sum_{j=1}^{N} H_j - \frac{1}{N}\sum_{j=1}^{N} O_j := b$$

For decadal prediction and other cases

$$b = b(\tau, \mathbf{X}),$$

τ: forecast lead-time; $\mathbf{X}$ and climate state $\mathbf{X}$.

Sytematic error we belief we can compensate *a posteriori*

Drift

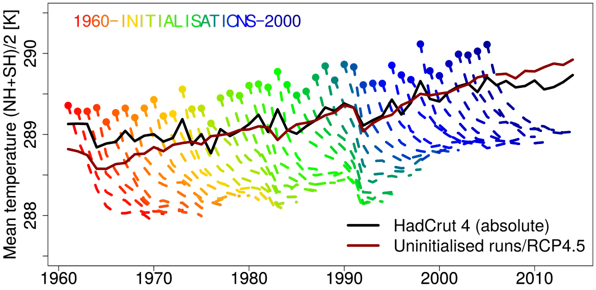
change in bias with forecast lead-time τ

$$D(\tau, \mathbf{X}) = \frac{\partial}{\partial \tau} b(\tau, \mathbf{X})$$

Drift

change in bias with forecast lead-time τ

$$D(\tau, \mathbf{X}) = \frac{\partial}{\partial \tau} b(\tau, \mathbf{X})$$

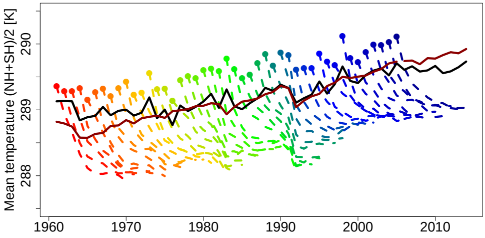


Quantifying drift

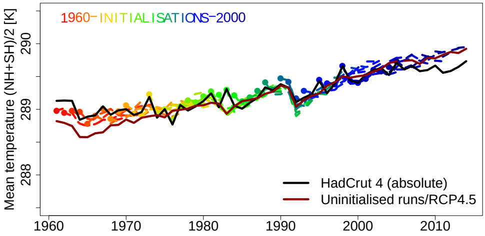
$$
\begin{aligned}
\|D(\tau, \mathbf{X})\| &= \sqrt{(D(\tau, \mathbf{X}))^2} \\
&= \sqrt{\left(\frac{\partial}{\partial \tau} b(\tau, \mathbf{X})\right)^2}
\end{aligned}
$$

Igor Kröner *Drift Quantification and Correction in Decadal Predictions of Climate Extremes Indices*, in preparation

Examples from MiKlip: Drift in global mean temperature



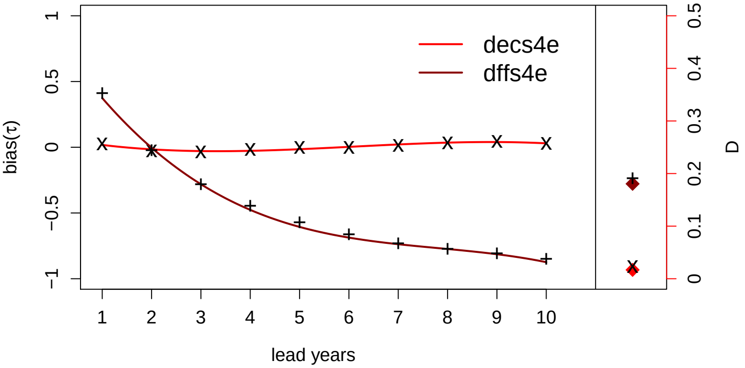
full-field



anomaly

courtesy of Igor Kröner

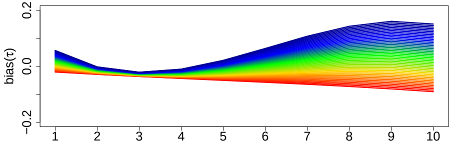
Examples from MiKlip: Drift in global mean temperature



red anomaly initialisation (baseline1, decs4e)

brown full-field initialisation (prototype, dffs4e)

courtesy of Igor Kröner



Drift Adjustment

Verification of forecasts with bias/drift

$$S(H(t,\tau), o(t))$$

Verification of forecasts with bias/drift

$$S(H(t,\tau), o(t))$$

How good is the forecast once we have compensated for systematic errors?

Verification of forecasts with bias/drift

$$S(\hat{H}(t,\tau), o(t))$$

$$\hat{H}(t,\tau) = H(t,\tau) - \hat{b}(t,\tau)$$

Verification of forecasts with bias/drift

$$S(\hat{H}(t,\tau), o(t))$$

$$\hat{H}(t,\tau) = H(t,\tau) - \hat{b}(t,\tau)$$

Recommendation[a] full-field and anomalies

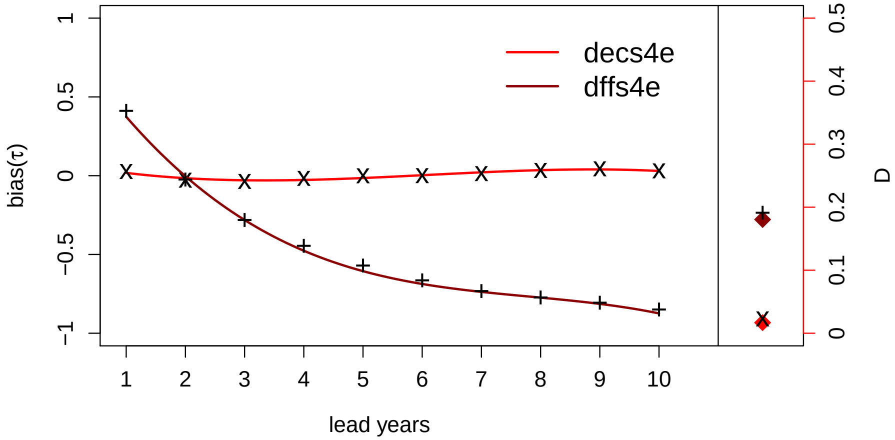
assume $b(\tau, \mathbf{X}(t)) = b(\tau, t)$

$$\hat{b}_{ICPO}(t,\tau) = \frac{1}{\#\{t_i \backslash t\}} \sum_{t_i \backslash t} (H(t_i,\tau) - o(t_i)) \approx ME(\tau)$$

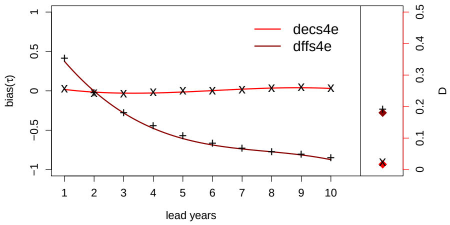
yr of forecast to be verified left out (ICPO 2011)
lead-years τ are treated individually

[a] Boer et al. [2016]

Back to the drift . . .



Back to the drift . . .



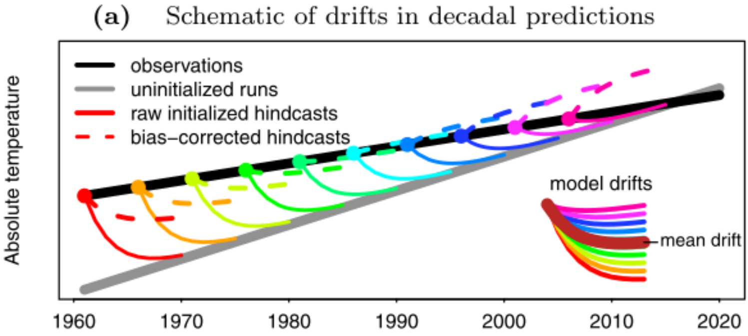
Drift ist smooth in τ

$b(\tau, \mathbf{X}(t)) = b(\tau)$, smooth/parametric form in τ

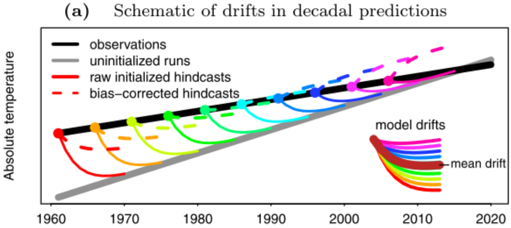
$$\hat{b}_{Gan}(t, \tau) = a_0 + a_1\,\tau + a_2\,\tau^2 + a_3\,\tau^3$$

third order polynomial in τ Gangstø et al. [2013]
(exponential Pattantyús-Ábrahám et al. [2016])

Back to the drift . . .



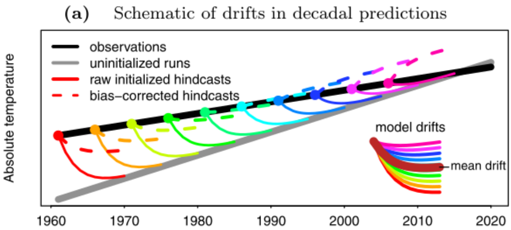
Back to the drift . . .



Drift might change with climate (t)[5]

[5] Kharin et al. [2012]Fučkar et al. [2014]

Back to the drift . . .

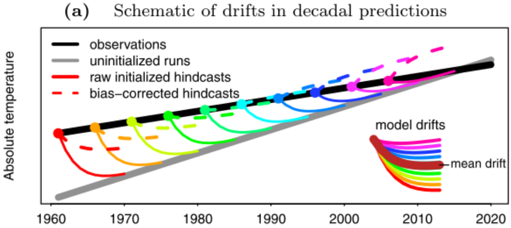


Drift might change with climate (t)[5]

$$\hat{b}(t,\tau) = a_0 + a_1\,\tau + a_2\,\tau^2 + a_3\,\tau^3$$

[5] Kharin et al. [2012]Fučkar et al. [2014]

Back to the drift . . .

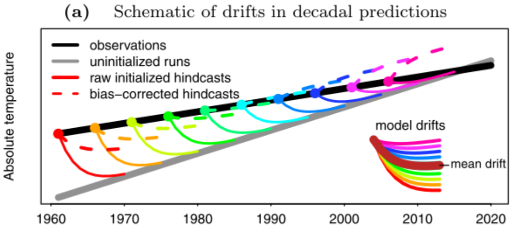


Drift might change with climate (t)[5]

$$\hat{b}(t,\tau) = a_0(t) + a_1(t)\,\tau + a_2(t)\,\tau^2 + a_3(t)\,\tau^3$$

[5] Kharin et al. [2012]Fučkar et al. [2014]

Back to the drift . . .



Drift might change with climate (t)[5]

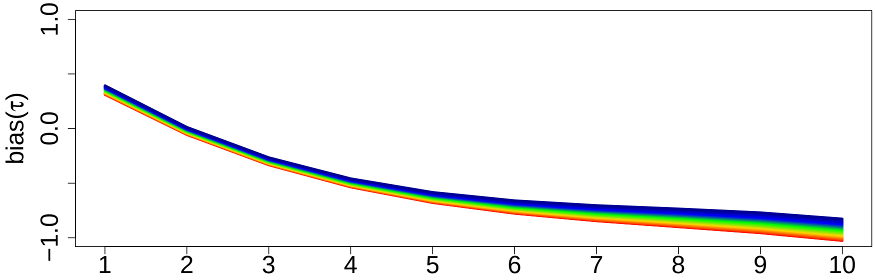
$$\hat{b}_{Kru}(t,\tau) = (b_0 + b_1\, t) + (b_2 + b_3\, t)\,\tau + (b_4 + b_5\, t)\,\tau^2 + (b_6 + b_7\, t)\,\tau^3$$

Kruschke et al. [2015]

[5] Kharin et al. [2012]Fučkar et al. [2014]

Drift . . .

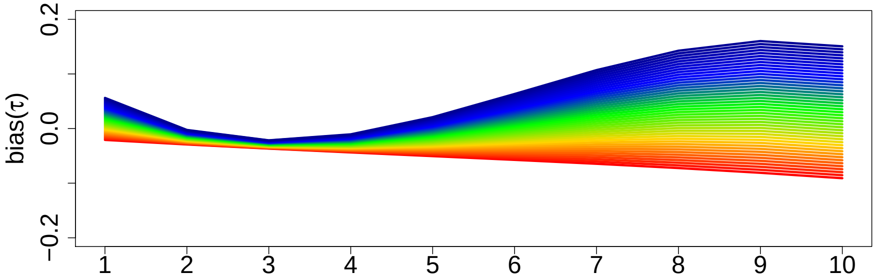
$$\hat{b}_{Kru}(t,\tau) = (b_0 + b_1\,t) + (b_2 + b_3\,t)\,\tau + (b_4 + b_5\,t)\,\tau^2 + (b_6 + b_7\,t)\,\tau^3$$



tas, full-field initialisation,
red (early init, 1960) to blue (late init, 2004)

Drift . . .

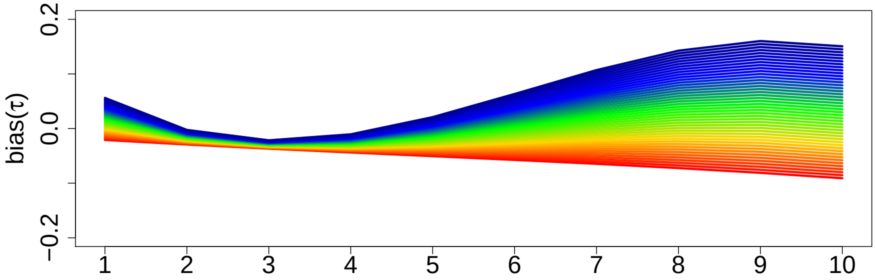
$$\hat{b}_{Kru}(t,\tau) = (b_0 + b_1\,t) + (b_2 + b_3\,t)\,\tau + (b_4 + b_5\,t)\,\tau^2 + (b_6 + b_7\,t)\,\tau^3$$



tas, anomaly initialisation,
red (early init, 1960) to blue (late init, 2004)

Drift . . .

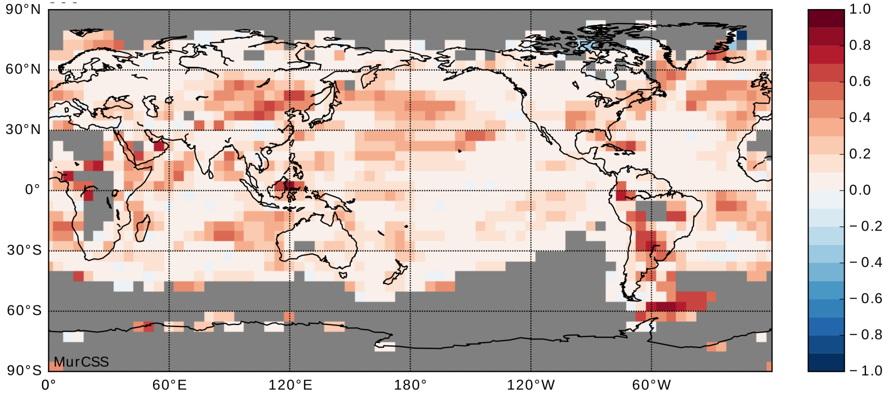
$$\hat{b}_{Kru}(t,\tau) = (b_0 + b_1\,t) + (b_2 + b_3\,t)\,\tau + (b_4 + b_5\,t)\,\tau^2 + (b_6 + b_7\,t)\,\tau^3$$



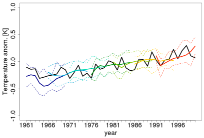
tas, anomaly initialisation,
red (early init, 1960) to blue (late init, 2004)

That seems complex! Does this help?

Parametric drift adjustment vs ICPO

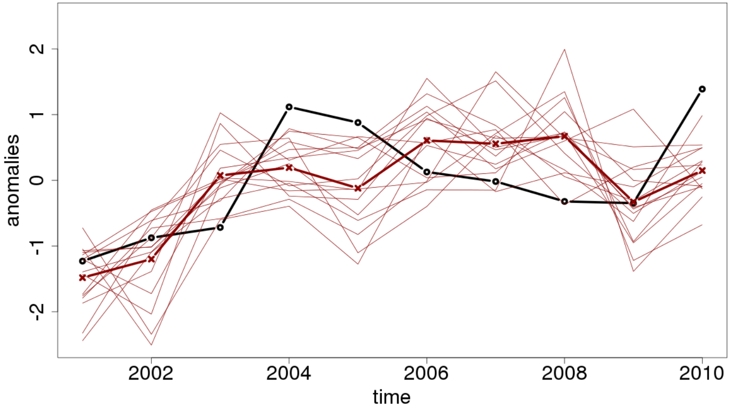


tas, full-field, MSESS, polynomial vs ICPO, yr2-5



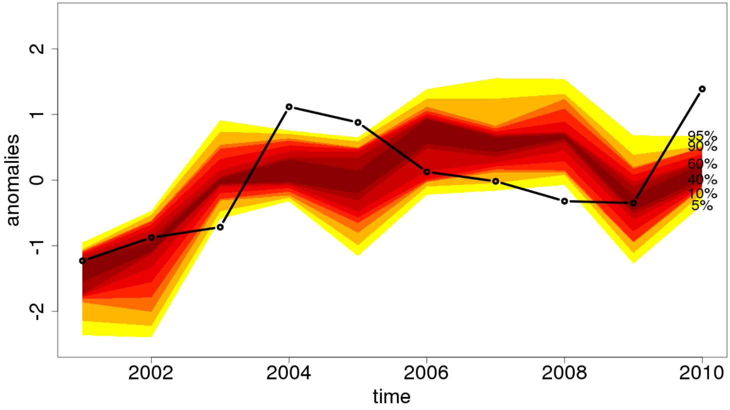
Re-calibration

Probabilistic forecast



all figures curtesy of Alexander Pasternack

Probabilistic forecast



all figures curtesy of Alexander Pasternack

Re-calibration method for decadal predictions

$$f_i(t,\tau) = \mu(t,\tau) + \epsilon_i(t,\tau)$$

[6] $\mu(t,\tau)$: ensemble mean, $i = 1 \ldots M$ member, $t =$init. year, $\tau =$ lead year

[6]e.g. Weigel et al. [2008]

Re-calibration method for decadal predictions

$$f_i(t,\tau) = \mu(t,\tau) + \epsilon_i(t,\tau)$$

[6] $\mu(t,\tau)$: ensemble mean, $i = 1 \dots M$ member, $t =$init. year, $\tau =$ lead year

Re-calibrated ensemble

$$f_i^{Cal}(t,\tau) = \mu(t,\tau) + \epsilon_i(t,\tau)$$

[6] e.g. Weigel et al. [2008]

Re-calibration method for decadal predictions

$$f_i(t, \tau) = \mu(t, \tau) + \epsilon_i(t, \tau)$$

[6] $\mu(t, \tau)$: ensemble mean, $i = 1 \dots M$ member, $t =$init. year, $\tau =$ lead year

Re-calibrated ensemble

$$f_i^{Cal}(t, \tau) = \alpha(t, \tau) + \mu(t, \tau) + \epsilon_i(t, \tau)$$

1) α: bias and drift,

[6]e.g. Weigel et al. [2008]

Re-calibration method for decadal predictions

$$f_i(t,\tau) = \mu(t,\tau) + \epsilon_i(t,\tau)$$

[6] $\mu(t,\tau)$: ensemble mean, $i = 1 \ldots M$ member, $t =$init. year, $\tau =$ lead year

Re-calibrated ensemble

$$f_i^{Cal}(t,\tau) = \alpha(t,\tau) + \beta(t,\tau)\mu(t,\tau) + \epsilon_i(t,\tau)$$

1) α: bias and drift, 2) β: conditional bias,

[6] e.g. Weigel et al. [2008]

Re-calibration method for decadal predictions

$$f_i(t,\tau) = \mu(t,\tau) + \epsilon_i(t,\tau)$$

[6] $\mu(t,\tau)$: ensemble mean, $i = 1 \ldots M$ member, $t =$init. year, $\tau =$ lead year

Re-calibrated ensemble

$$f_i^{Cal}(t,\tau) = \alpha(t,\tau) + \beta(t,\tau)\mu(t,\tau) + \gamma(t,\tau)\epsilon_i(t,\tau)$$

1) α: bias and drift, 2) β: conditional bias, 3) γ: spread

[6] e.g. Weigel et al. [2008]

Re-calibration method for decadal predictions

$$f_i(t,\tau) = \mu(t,\tau) + \epsilon_i(t,\tau)$$

[6] $\mu(t,\tau)$: ensemble mean, $i = 1 \ldots M$ member, t =init. year, τ = lead year

Re-calibrated ensemble

$$f_i^{Cal}(t,\tau) = \alpha(t,\tau) + \beta(t,\tau)\mu(t,\tau) + \gamma(t,\tau)\epsilon_i(t,\tau)$$

1) α: bias and drift, 2) β: conditional bias, 3) γ: spread

find $\alpha(t,\tau)$, $\beta(t,\tau)$, $\gamma(t,\tau)$ such that ensemble is calibrated with maximum sharpness

[6] e.g. Weigel et al. [2008]

A model for α, β, and γ

a first go [a]:

$$\alpha(t,\tau) = (a_0 + a_1 t) + (a_2 + a_3 t)\tau + (a_4 + a_5 t)\tau^2 + (a_6 + a_7 t)\tau^3$$

---

[a] Pasternack et al. [2017]

A model for α, β, and γ

a first go [a]:

$$\alpha(t,\tau) = (a_0 + a_1 t) + (a_2 + a_3 t)\tau + (a_4 + a_5 t)\tau^2 + (a_6 + a_7 t)\tau^3$$
$$\beta(t,\tau) = (b_0 + b_1 t) + (b_2 + b_3 t)\tau + (b_4 + b_5 t)\tau^2 + (b_6 + b_7 t)\tau^3$$
$$\gamma(t,\tau) = (c_0 + c_1 t) + (c_2 + c_3 t)\tau + (c_4 + c_5 t)\tau^2 + (c_6 + c_7 t)\tau^3$$

[a] Pasternack et al. [2017]

A model for α, β, and γ

a first go [a]:

$$\alpha(t,\tau) = (a_0 + a_1 t) + (a_2 + a_3 t)\tau + (a_4 + a_5 t)\tau^2 + (a_6 + a_7 t)\tau^3$$
$$\beta(t,\tau) = (b_0 + b_1 t) + (b_2 + b_3 t)\tau + (b_4 + b_5 t)\tau^2 + (b_6 + b_7 t)\tau^3$$
$$\gamma(t,\tau) = (c_0 + c_1 t) + (c_2 + c_3 t)\tau + (c_4 + c_5 t)\tau^2 + (c_6 + c_7 t)\tau^3$$

Find parameters . . .

. . . by minimizing scores:

- CRPS Gneiting et al. [2005]
- ignorance score

details

[a] Pasternack et al. [2017]

A model for α, β, and γ

a first go [a]:

$$\alpha(t,\tau) = (a_0 + a_1 t) + (a_2 + a_3 t)\tau + (a_4 + a_5 t)\tau^2 + (a_6 + a_7 t)\tau^3$$
$$\beta(t,\tau) = (b_0 + b_1 t) + (b_2 + b_3 t)\tau + (b_4 + b_5 t)\tau^2 + (b_6 + b_7 t)\tau^3$$
$$\gamma(t,\tau) = (c_0 + c_1 t) + (c_2 + c_3 t)\tau + (c_4 + c_5 t)\tau^2 + (c_6 + c_7 t)\tau^3$$

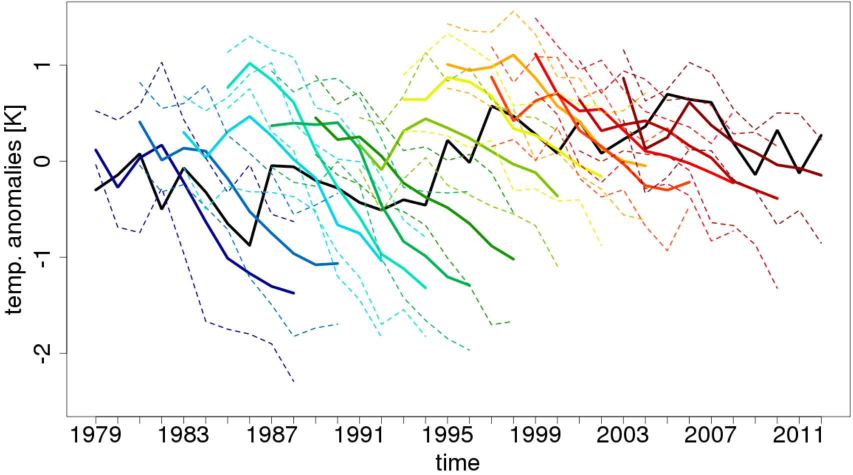
Find parameters ...

... by minimizing scores:

- CRPS Gneiting et al. [2005] details
- ignorance score

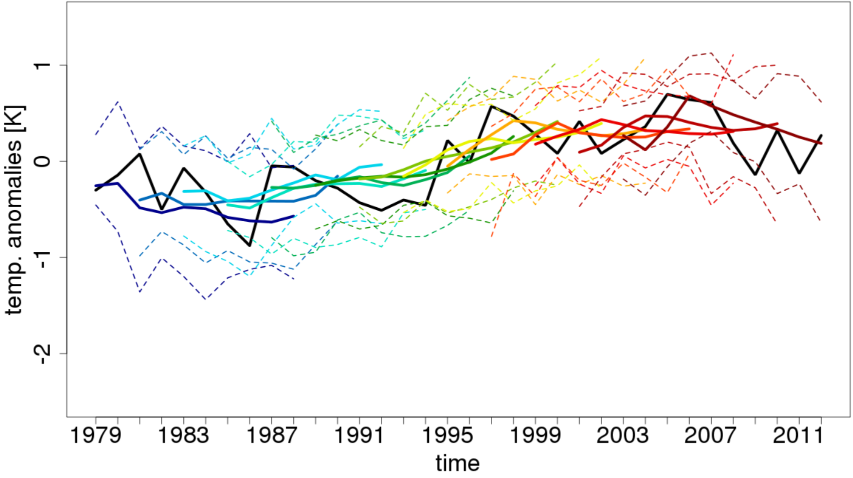
[a] Pasternack et al. [2017]

Re-calibration for MiKlip



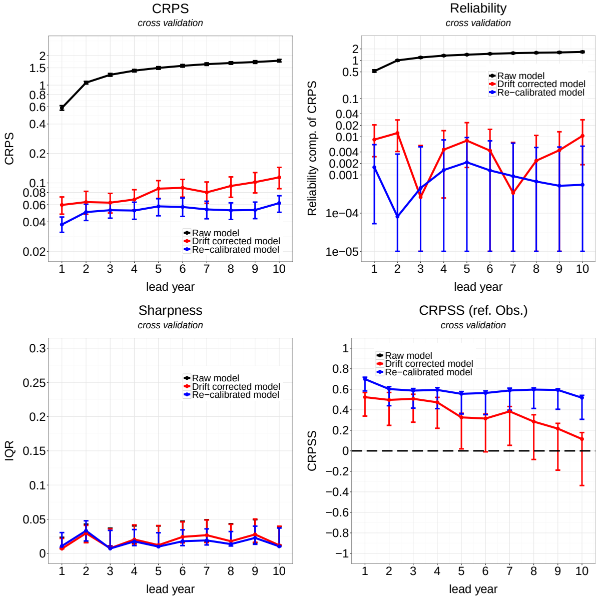
tas, global mean, full-field

Re-calibration for MiKlip



tas, global mean, full-field

Re-calibration for MiKlip – Verification

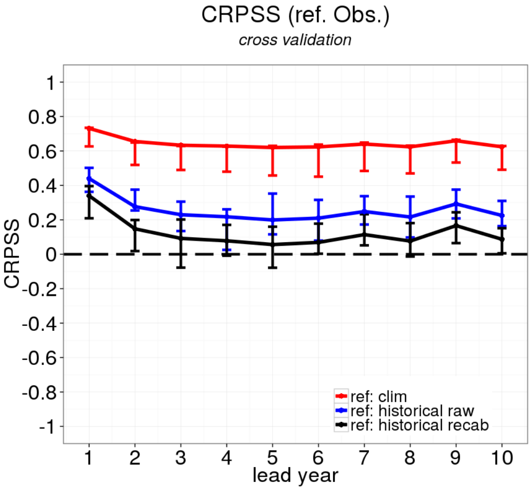


Re-calibration for MiKlip – Verification

Re-calibrating the historical simulations (reference):

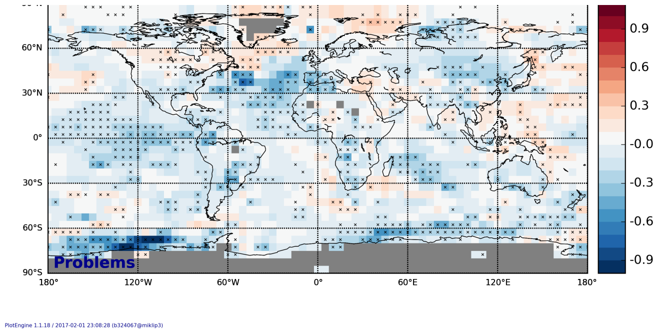
Re-calibration for MiKlip – Verification

Re-calibrating the historical simulations (reference):



Verifying grid-cells with CRPSS

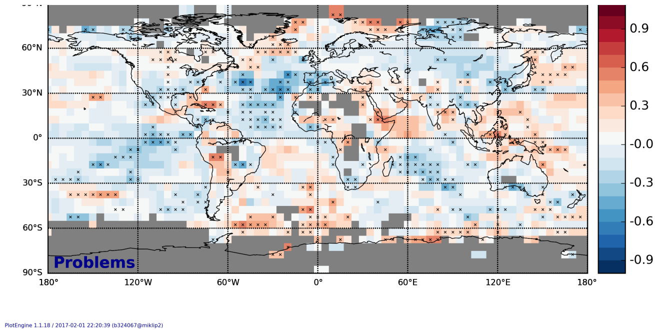
mean surface temperature, lead-year 7-10, vs historical



ICPO	$\alpha_{\text{ICPO}}(t, \tau)$
param. drift correction	$\alpha(t, \tau)$
plus cond. bias	$\beta(t, \tau)$
plus ensemble spread	$\gamma(t, \tau)$

Verifying grid-cells with CRPSS

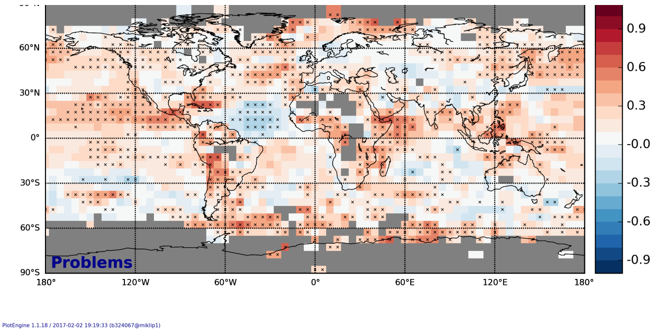
mean surface temperature, lead-year 7-10, vs historical



ICPO	$\alpha_{\text{ICPO}}(t, \tau)$
param. drift correction	$\alpha(t, \tau)$
plus cond. bias	$\beta(t, \tau)$
plus ensemble spread	$\gamma(t, \tau)$

Verifying grid-cells with CRPSS

mean surface temperature, lead-year 7-10, vs historical



ICPO $\alpha_{\mathrm{ICPO}}(t,\tau)$

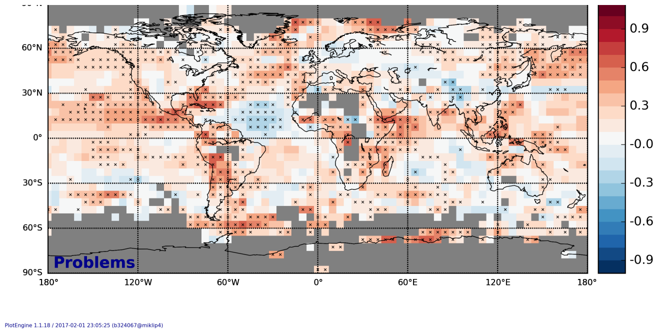
param. drift correction $\alpha(t,\tau)$

plus cond. bias $\beta(t,\tau)$

plus ensemble spread $\gamma(t,\tau)$

Verifying grid-cells with CRPSS

mean surface temperature, lead-year 7-10, vs historical



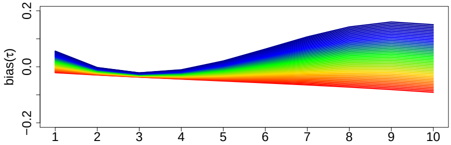
ICPO	$\alpha_{\mathrm{ICPO}}(t,\tau)$
param. drift correction	$\alpha(t,\tau)$
plus cond. bias	$\beta(t,\tau)$
plus ensemble spread	$\gamma(t,\tau)$

Summary

Verification of decadal predictions

- framework Goddard et al. [2013]
  - ensemble mean accuracy: MSESS
  - ensemble spread: CRPS based σ^2_{ens} vs MSE
  - consider: LESS
- (multi-)annual averages
- score corrections for small ensembles
- drift issue

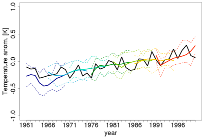
Summary



Drift adjustment

- $b = b(\tau, \mathbf{X}(t))$
- drift: full-field > anomaly initialisation
- parametric post-processing helps
- drift depending on climate not just on time Fučkar et al. [2014]

Summary



Re-calibration

- $f_i^{Cal} = \alpha(t,\tau) + \beta(t,\tau)\mu(t,\tau) + \gamma(t,\tau)\epsilon_i(t,\tau)$
- parametrics form for α, β, γ
- minimize CRPS/IGN to estimate parameters
- improves calibration, does not reduce sharpness (leave-10-yrs out cross validation)

Open issues

- grid-cell wise ’climate trend’ estimation
- model selection
- parameter uncertainty

References

G. J. Boer et al. The decadal climate prediction project (DCPP) contribution to CMIP6. *Geosci. Model Dev.*, 9:3751–77, 2016.

C. A. T. Ferro et al. On the effect of ensemble size on the discrete and continuous ranked probability scores. *Meteor. Appl.*, 15:19 – 24, 2008.

N. S. Fučkar et al. A posteriori adjustment of near-term climate predictions: Accounting for the drift dependence on the initial conditions. *Geophys. Res. Lett.*, 41(14):5200–5207, 2014.

R. Gangstø et al. Methodological aspects of the validation of decadal predictions. *Clim. Res.*, 55: 181–200, 2013.

T. Gneiting and A. E. Raftery. Strictly proper scoring rules, prediction, and estimation. *J. Amer. Statist. Assoc.*, 102(477):359–378, 2007.

T. Gneiting et al. Calibrated probabilistic forecasting using ensemble model output statistics and minimum crps estimation. *Month. Weather Rev.*, 133:1098–1118, 2005.

L. Goddard et al. A verification framework for interannual-to-decadal predictions experiments. *Climate Dynamics*, 40:245–272, 2013.

C. Kadow et al. Evaluation of forecasts by accuracy and spread in the miklip decadal climate prediction system. *Met. Z.*, 01 2014.

V. V. Kharin et al. Statistical adjustment of decadal predictions in a changing climate. *Geophys. Res. Lett.*, 39:L19705, 2012.

T. Kruschke et al. Probabilistic evaluation of decadal predictions for northern hemisphere winter storms. *Meteorol. Z.*, 2015.

G. A. Meehl et al. Decadal climate prediction: An update from the trenches. *Bull. Amer. Meteorol. Soc.*, 95(2):243–267, 2014.

W. A. Müller et al. A debiased ranked probability skill score to evaluate probabilistic ensemble forecasts with small ensemble sizes. *J. Clim.*, 18(10): 1513–1523, 2005.

A. Pasternack et al. Decadal forecast calibration – a parametric strategy accounting for drift, conditional bias and ensemble spread. *in preparation*, 2017.

M. Pattantyús-Ábrahám et al. Bias and drift of the medium-range decadal climate prediction system (MiKlip) validated by european radiosonde data. *Meteorologische Zeitschrift*, pages 709–720, 2016.

D. M. Smith, R. Eade, and H. Pohlmann. A comparison of full-field and anomaly initialization for seasonal to decadal climate prediction. *Climate Dynamics*, 41(11):3325–3338, 2013.

A. P. Weigel, M. A. Liniger, and C. Appenzeller. Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? *Quart. J. Royal Meteor. Soc.*, 134(630): 241–260, 2008.