Probabilistic climate reconstructions using Holocene pollen records from lake Holzmaar and Meerfelder Maar

12th International Palynological Congress IPC-XII 2008
September 5th 2008, Bonn, Germany

Christian Schölzel¹, Andreas Hense¹, Thomas Litt², and Norbert Kühl²

¹) Meteorological Institute at the University of Bonn, Germany
²) Institute for Palaeontology at the University of Bonn, Germany
Outline

Focus

1. Probabilistic palaeoclimate reconstructions based on pollen
2. Application to sediments from lake Holzmaar and Meerfelder Maar
State of the art
Palaeoenvironmental transfer functions

**Palaeoenvironmental transfer functions**

- Various "tools" for the development of transfer functions
- Historically focused on the palaeo archives

**Capabilities**

- Knowledge of the bio-geochemical processes forming the archives

**Further improvements/trends**

- Statistical formulation
- Probabilistic point of view
- Reconstruction of uncertainties
Historical background

Classical mutual climatic range

**Mutual climatic range (MCR)**
- Reduce to presence/absence
- Pollen and macro fossils
- Robust against no modern analogues

**Problems**
- Implicit assumption of uniform distributions
  - Not very close to plant physiology
  - Boundaries are the decisive factor
- Graphical definition of mutual climatic ranges causes overfitting
General statistical formulation

Definition of transfer functions

Definitions
The climate system is a stochastic system $\Rightarrow$ random variables

- Palaeo climate state $\vec{X}_0$ reconstruction e.g. $(T_{DJF}, T_{JJA}, P_{Ann})$
- Palaeo proxy variables $\vec{Y}_0$ fossil pollen spectra $\vec{y}_0$
- Climate state $\vec{X}$ recent climate data $X$
- Proxy variables $\vec{Y}$ recent proxy data $Y$

Classical regression techniques
Estimate $\hat{\mu}_{\vec{X}_0}$ with $\hat{\mu}_{\vec{X}_0} = E(\vec{X}_0 | \ldots)$
General statistical formulation
Definition of transfer functions

Definitions
The climate system is a stochastic system $\rightarrow$ random variables

- Palaeo climate state $\vec{X}_0$ reconstruction e.g. $(T_{DJF}, T_{JJA}, P_{Ann})$
- Palaeo proxy variables $\vec{Y}_0$ fossil pollen spectra $\vec{y}_0$
- Climate state $\vec{X}$ recent climate data $X$
- Proxy variables $\vec{Y}$ recent proxy data $Y$

Classical regression techniques
Estimate $\hat{\mu}_{\vec{X}_0}$ with $\mu_{\vec{X}_0} = E(\vec{X}_0 | \ldots)$
General statistical formulation
Definition of transfer functions

Definitions
The climate system is a stochastic system \( \implies \) random variables

- Palaeo climate state: \( \vec{X}_0 \) reconstruction e.g. \( (T_{DJF}, T_{JJA}, P_{Ann}) \)
- Palaeo proxy variables: \( \vec{Y}_0 \) fossil pollen spectra \( \vec{y}_0 \)
- Climate state: \( \vec{X} \) recent climate data \( X \)
- Proxy variables: \( \vec{Y} \) recent proxy data \( Y \)

Probabilistic climate reconstructions
Search for the conditional probability density function \( f_{\vec{X}_0 | \ldots (\vec{X}_0 | \ldots)} \)
Classification of transfer functions

Probabilistic reconstructions

- Distribution fit (with covariates):
  \[
  \tilde{X}_0 \mid \tilde{Y}_0, \tilde{X}, \tilde{Y} \overset{e.g.}{\sim} \mathcal{N}(\tilde{\mu}, \Sigma_{X_0})
  \]

- Bayesian approach with plug-in estimator

  \[
  f_{\tilde{X}_0 \mid \tilde{Y}_0, \tilde{X}, \tilde{Y}}(\tilde{X}_0 \mid \tilde{Y}_0, \tilde{X}, \tilde{Y}) \propto f_{\tilde{Y} \mid \tilde{X}}(\tilde{Y}_0 \mid \tilde{X}_0; \hat{\theta}) \pi_{\tilde{X}_0}(\tilde{X}_0)
  \]

- Bayesian approach (BHM) using MCMC integration

  \[
  f_{\tilde{X}_0 \mid \tilde{Y}_0, \tilde{X}, \tilde{Y}}(\tilde{X}_0 \mid \tilde{Y}_0, \tilde{X}, \tilde{Y}) \propto \int_{\mathcal{V}_{\theta}} f_{\tilde{Y} \mid \tilde{X}, \tilde{X}_0}(\tilde{Y}, \tilde{Y}_0 \mid \tilde{X}, \tilde{X}_0, \tilde{Y}) \pi_{\tilde{X}, \tilde{X}_0 \mid \tilde{X}}(\tilde{X}, \tilde{X}_0 \mid \tilde{Y}) \pi_{\theta}(\tilde{Y}) \, d\tilde{Y}
  \]

  "regression" & "calibration"
**Probabilistic approach**

**Bayesian indicator taxa model**

**Concept**
- Only use taxon present
- Replace MCR by proper probability distributions

**Bayesian indicator taxa model**
- For all taxa $i(\tilde{k}) \in \{1, \ldots, n_k\}$ : $y_{i(\tilde{k})} = 1$
  - in the fossil pollen spectrum (▷▷ appendix):

$$f_{\tilde{X}|Y_{i(1)}, \ldots, Y_{i(n_{\tilde{k}})}}(\tilde{x}_0 | 1, \ldots, 1) \propto \prod_{\tilde{k}=1}^{n_{\tilde{k}}} \frac{f_{\tilde{X}|y_{i(\tilde{k})}}(\tilde{x}_0 | 1)}{m_{\tilde{X}}(X)} \cdot \pi_{\tilde{x}_0}(\tilde{x}_0)$$

- $f_{\tilde{X}|Y_k}(\tilde{x}_0 | 1)$: Taxon-specific likelihood function
- $\pi_{\tilde{x}_0}(\tilde{x}_0)$: Prior distribution of the climate state vector
- $m_{\tilde{X}}(X)$: Marginal distribution of the climate state vector
Probabilistic approach

Taxon specific likelihood functions $f_{\tilde{X}|Y_k}$

- Recent spatial distribution and climatology
- Precipitation component not Gaussian
- Solved by a copula approach
Proxy data

Pollen spectrum (Holzmaar)
Climate reconstruction

Bivariate prior and posterior (Holzmaar)

Climate reconstruction

The Bivariate posterior probability density functions of winter temperature and annual precipitation for three sample layers indicate

- Variations in the expectation value as well as
- Fluctuations in the variance ... \( \text{Var}(\vec{X}_0) \), not \( \text{Var}(\hat{\mu}_{\vec{X}_0}) \)
Climate reconstruction
Marginal posterior distributions (Holzmaar)
Summary

- Probabilistic approaches based on classical methods
- Non-homogeneous reconstruction of the variance
- Multivariate, non-normally distributed climate state vectors

Note

- Analogous development for biomisation methods

Outlook

- Bayesian hierarchical model, spatial dependence
- Unconditional climate distribution
- Maybe presence/absence information too weak . . .
References


Appendix

7 General statistical formulation
   - “Regression” and “Calibration”
   - Probabilistic indicator taxa model
   - Copula approach

8 Conditional/weighted climate-taxon relations
   - Simple test on kernel density estimators

9 Meerfelder Maar
   - Pollen spectrum
   - Local climate reconstruction

➤➤ Main part
“Regression” and “Calibration”
A more probabilistic point of view

Reconstruction
Conditional probability density for the palaeo climate state $\tilde{X}_0$

$$f_{\tilde{X}_0|\tilde{Y}_0,\tilde{X},\tilde{Y}}(\tilde{X}_0|\tilde{Y}_0, \mathbf{X}, \mathbf{Y}) = \int_{\mathcal{V}_\theta} f_{\tilde{X}_0|\tilde{Y}_0, \theta}(\tilde{X}_0|\tilde{Y}_0, \tilde{\theta}) \pi_{\tilde{\theta}|\tilde{X}, \tilde{Y}}(\tilde{\theta}|\mathbf{X}, \mathbf{Y}) \, d\tilde{\theta}$$

with parameter space $\tilde{\theta} \in \mathcal{V}_\theta$
Probabilistic indicator taxa model

Derivations

\[
f_{Y_1, \ldots, Y_{n_k} | \bar{X}}(y_1, \ldots, y_{n_k} | \bar{x}) = f_{Y_1 | \bar{X}}(y_1 | \bar{x}) \cdot \prod_{k=2}^{n_k} f_{Y_k | Y_1, \ldots, Y_{k-1}, \bar{X}}(y_k | y_k, \ldots, y_{k-1}, \bar{x}) \]

\[
= \prod_{k=1}^{n_k} f_{Y_k | \bar{X}}(y_k | \bar{x})
\]

\[
= \prod_{k=1}^{n_k} \frac{f_{\bar{X} | Y_k}(\bar{x} | y_k) \cdot \pi_{Y_k}(y_k)}{m_{\bar{X}}(\bar{x})}
\]

\[
f_{\bar{X} | \bar{Y}}(\bar{x} | y_1, \ldots, y_{n_k}) = \frac{f_{\bar{Y} | \bar{X}}(\bar{y_1}, \ldots, \bar{y_{n_k}} | \bar{x}) \cdot \pi_{\bar{X}}(\bar{x})}{m_{\bar{Y}}(y_1, \ldots, y_{n_k})}
\]

\[
= \frac{\pi_{\bar{X}}(\bar{x})}{m_{\bar{Y}}(y_1, \ldots, y_{n_k})} \cdot \prod_{k=1}^{n_k} \frac{f_{\bar{X} | Y_k}(\bar{x} | y_k) \cdot \pi_{Y_k}(y_k)}{m_{\bar{X}}(\bar{x})}
\]

\[
f_{\bar{X} | Y_{i(1)}, \ldots, Y_{i(n_k)}}(\bar{x} | 1, \ldots, 1, C) \propto \pi_{\bar{X}}(\bar{x}) \cdot \prod_{\bar{k}=1}^{n_k} \frac{f_{\bar{X} | Y_{i(\bar{k})}}(\bar{x} | 1, C)}{m_{\bar{X}}(\bar{x})}
\]
Copula approach
A graphical representation

\[ f_{\mathcal{N}(0,\Sigma)}^{-1} \]
\[ C_\vec{X} \]
\[ f_\vec{X} \]
Conditional climate-taxon relations
Kernel density estimators

Betula_tree Betula_tree (Be_m)
KDE (Quantiles*)
Conditional density*

Carpinus betulus (Cabe_m)
KDE (Quantiles*)
Conditional density*
Proxy data

Pollen spectrum (Meerfelder Maar)

Holocene pollen spectrum from lake Meerfelder Maar
Climate reconstruction
Marginal posterior distributions (Meerfelder Maar)
1 Introduction

2 Background
   - Palaeoenvironmental transfer functions
   - Mutual climatic range approach

3 Methods
   - General statistical formulation
   - Probabilistic indicator taxa model

4 Data
   - Fossil pollen spectra

5 Results
   - Local climate reconstructions

6 Conclusion

►► Appendix