Department 5 Seminar
Handling noise, uncertainty and their propagation

Michael Dietze[1]

1 - GFZ German Research Centre for Geosciences, Section 5.1 Geomorphology

Department 5 Seminar

Handling noise, uncertainty and their propagation

Some thoughts about terms and definitions

Signal	Time-dependent function or value that describes certain properties of an entity.
Noise	A scatter of values around a signal (red vs. white noise)

Some thoughts about terms and definitions

Signal	Time-dependent function or value that describes certain properties of an entity.
Noise	A scatter of values around a signal (red vs. white noise)
Uncertainty	Size of the value domain in which the true value is to be expected.
Error	Deviation of a measured value from the true value.
Precision	Reproducibility of a measured value (not its correctness).
Accuracy	Agreement between measured and true value.

Some thoughts about terms and definitions

Signal Time-dependent function or value that describes certain properties of an entity.

Noise A scatter of values around a signal (red vs. white noise)

Uncertainty Size of the value domain in which the true value is to be expected.

Error Deviation of a measured value from the true value.

Precision Reproducibility of a measured value (not its correctness).

Accuracy Agreement between measured and true value.

Propagation Rule to account for combined effects of individual uncertainties in connected systems.

$$\Delta A = \sqrt{\sum_{i=1}^{n} \left(\frac{\delta A}{\delta x_i}\right)^2 \cdot \Delta x_i^2}$$

Department 5 Seminar

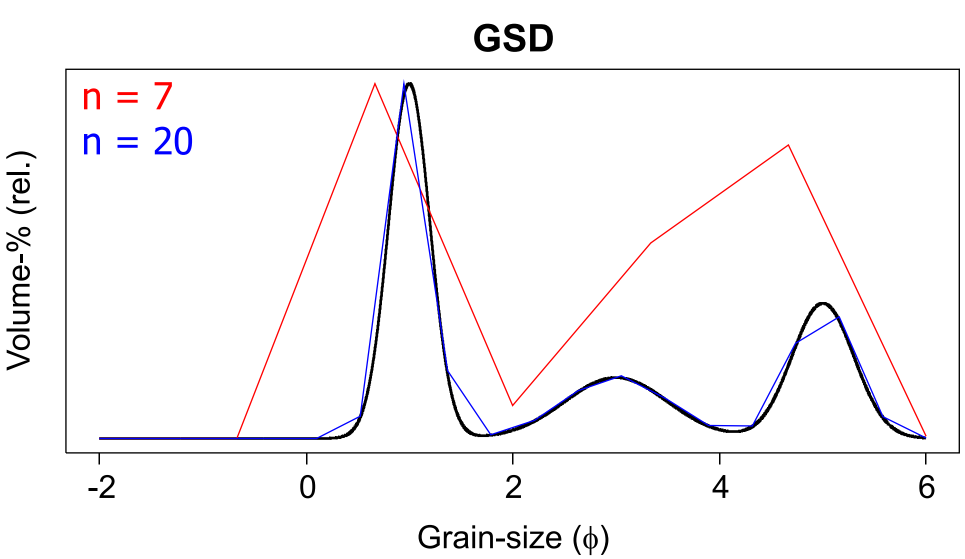
Handling noise, uncertainty and their propagation

Grain-size distribution unmixing and its role in understanding Earth surface dynamics

Michael Dietze[1]

1 - GFZ German Research Centre for Geosciences, Section 5.1 Geomorphology

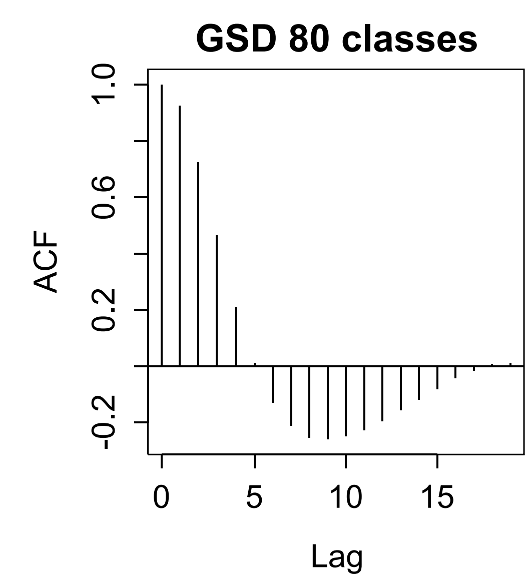
Grain-size data - a brief welcome



A continuous entity is described by discrete measured values.

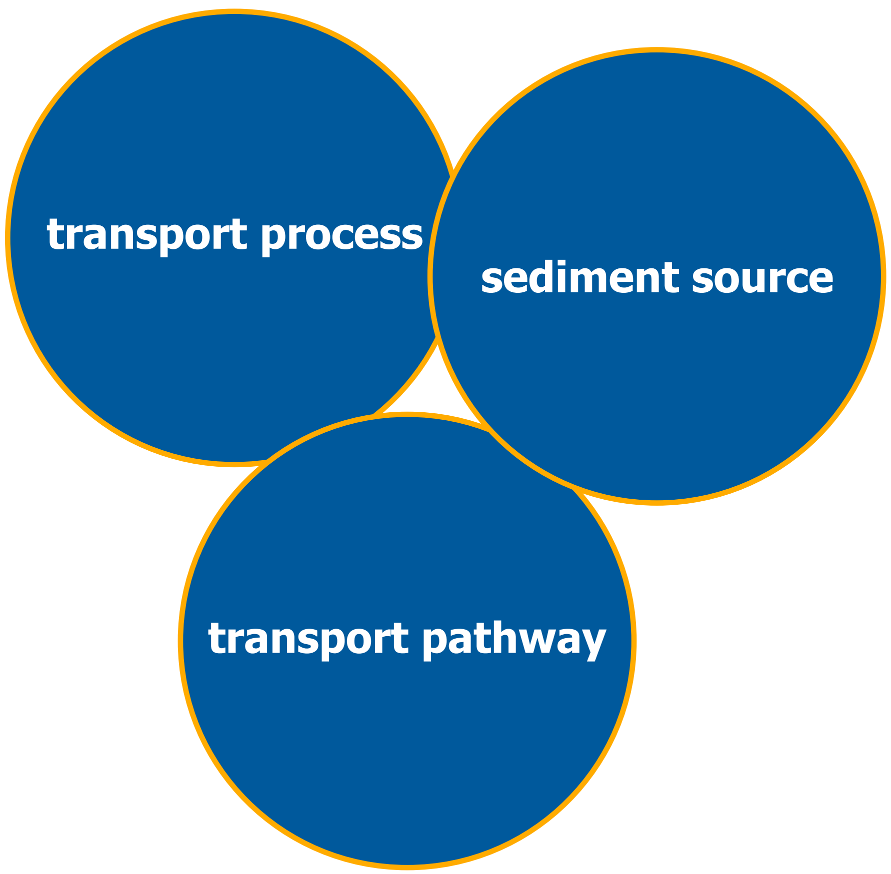
Redundancy, due to class autocorrelation

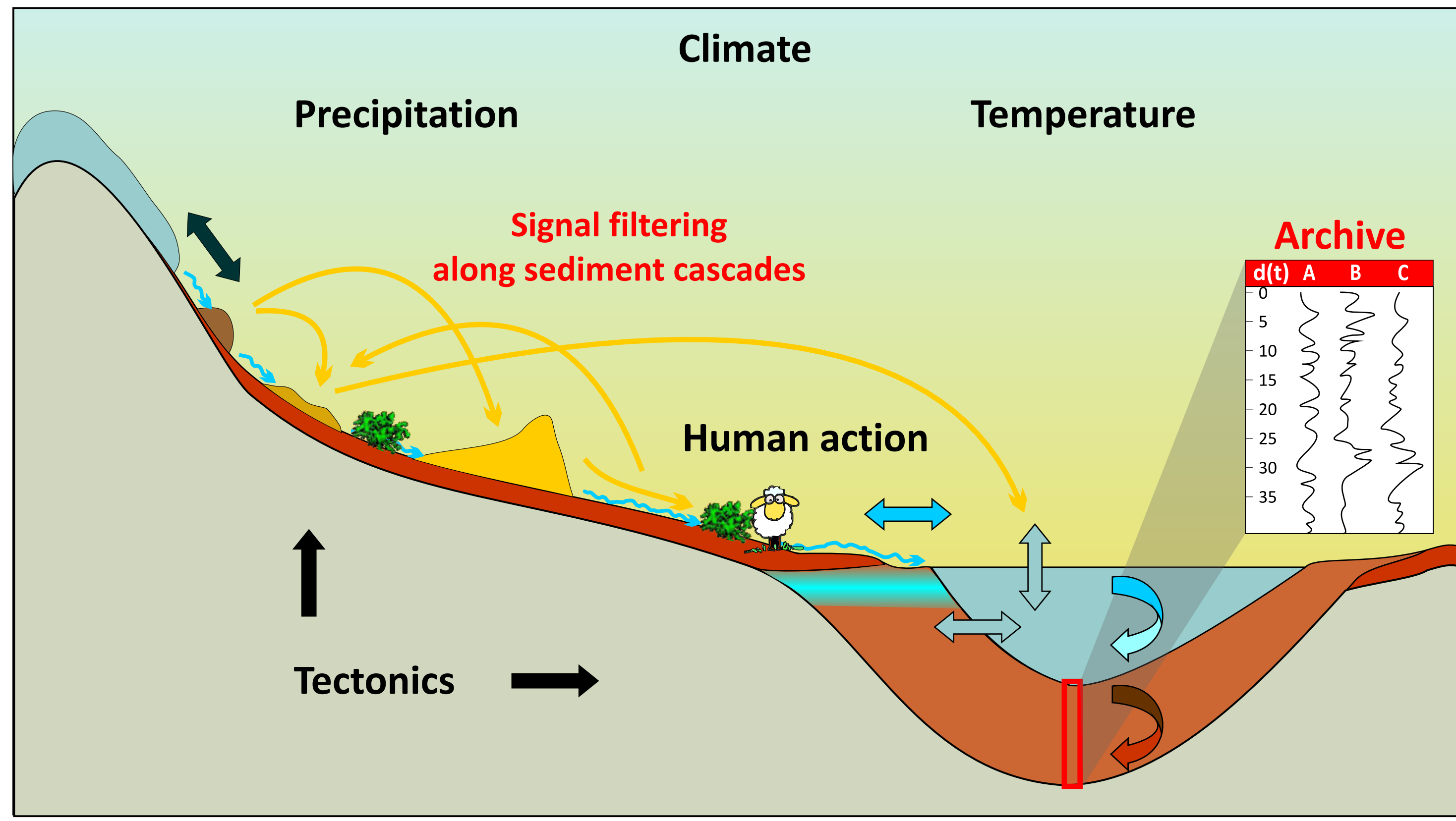
Uncertainty, (mainly) due to dis-cretisation



Grain-size data - a proxy for...

temporal buffers path dependency non-liniearity inheritance multiplicity

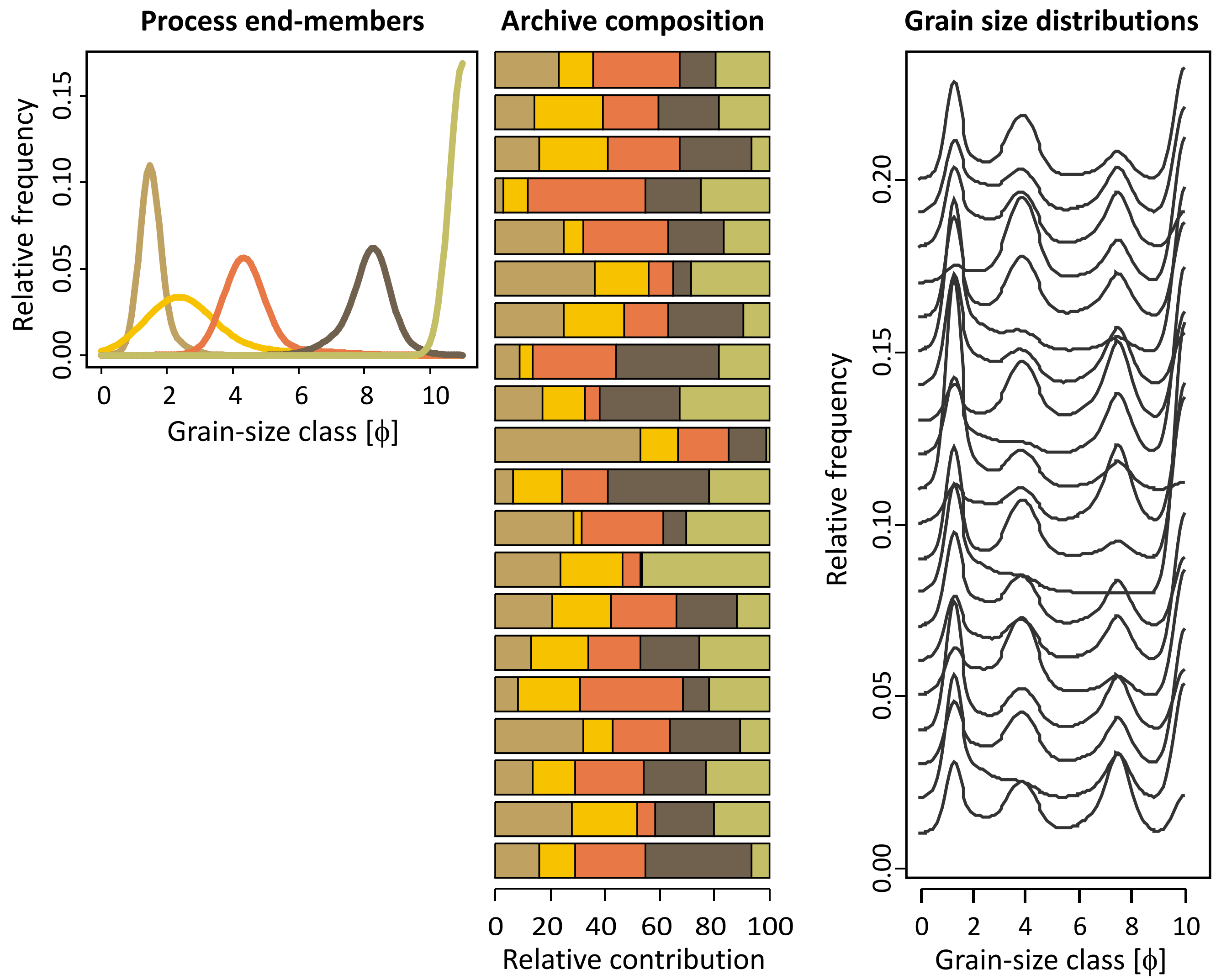




Glacier	Surface runoff	Soil water	Permafrost	Ground water	Lake water	Water cascade
Glacial deposits	Slope, fluvial, alluvial deposits	Aeolian deposits		Fluvio-lacustrine, littoral deposits	Limnic deposits	Sediment cascade

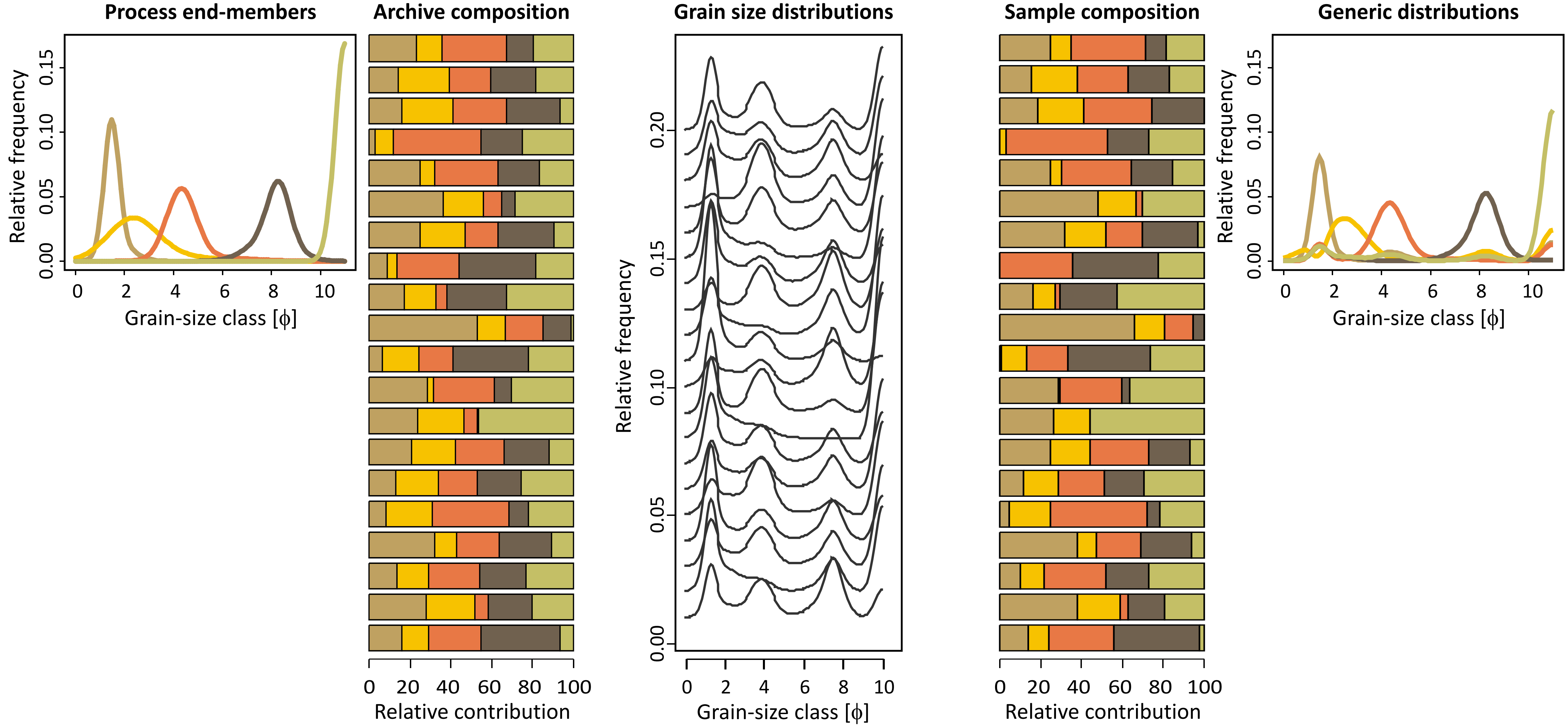
Concept of dynamic populations (Weltje & Prins, 2007)

Grain-size data - a proxy for...



from process to record

Grain-size data - a proxy for...

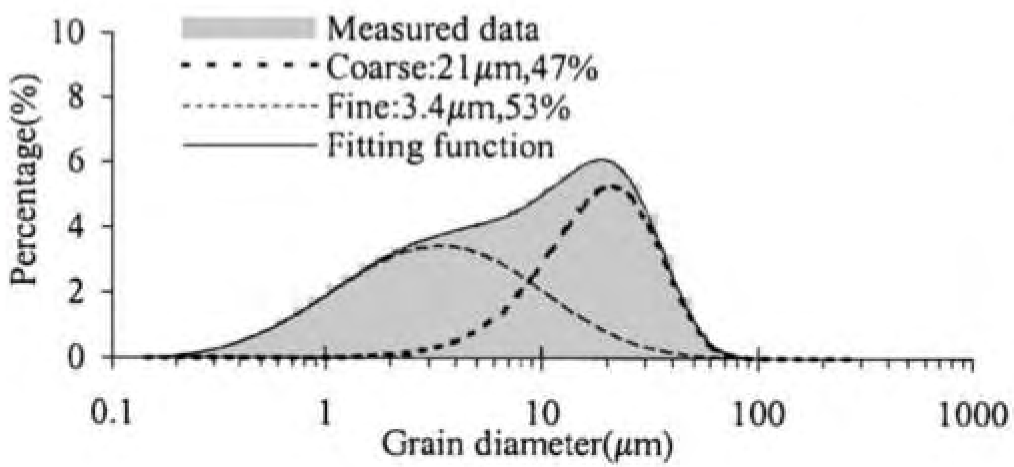


from **process** to **record**

from **record** to **process**

How to unmix grain-size data

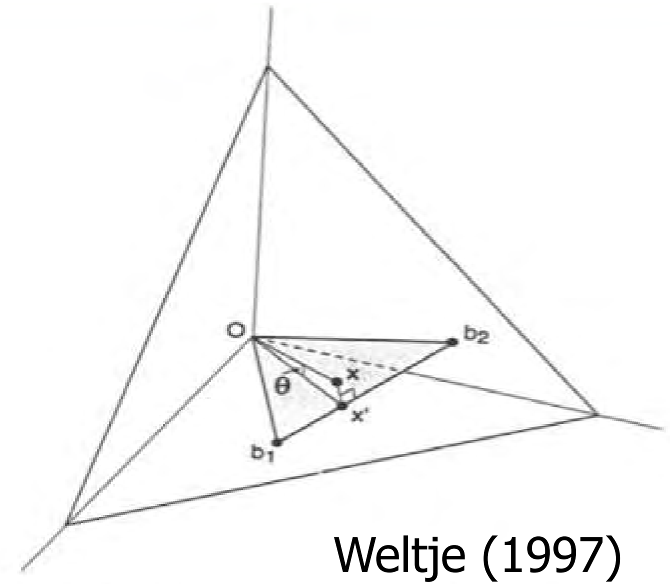
Finite mixture modelling



Sun et al. (2002)

Somewhat "straightforward"
Parametric description possible
Only local fitting
Strong influence of assumptions

End-member modelling



Weltje (1997)

Somewhat "vague"
No reduction of grain-size classes
Global fitting
Good constraints on (the few) parameters

EMMA - End-member modelling analysis (FORTRAN > Matlab > R)

Rescaling of the data matrix X to constant sum c

```
X  <- X / apply(X, 1, sum) * c
```

Weight transformation with quantile range l_w to get matrix W

```
qts <- function(X, lw) quantile(X,
  c(lw, 1-lw), type = 5)
ls  <- t(apply(X, 2, qts, lw = lw))
W   <- t((t(X) - ls[,1]) / (
  ls[,2] - ls[,1]))
```

Similarity Matrix A calculation from W based on outer product

```
A <- t(W) %*% W
```

Eigenspace extraction to get vectors V and cumulative scores L_v.

```
EIG <- eigen(A)
V   <- EIG$vectors[,order(seq(ncol(A), 1, -1))]
Vf  <- V[,order(seq(ncol(A), 1, -1))]
L   <- EIG$values[order(seq(ncol(A), 1, -1))]
Lv  <- cumsum(sort(L/sum(L), decreasing = TRUE))
```

Rotation of the eigenvector matrix V_f to get rotated Matrix V_r

```
Vr <- do.call(rotation, list(Vf[,1:q]))
```

Extract and **sort** factor loadings V_q, **rescale** (V_{qr}) and **normalise** (V_{qn})

```
Vq <- Vr$loadings[,order(seq(q, 1, -1))]
Vqr <- t(t(t(Vq) / apply(Vq, 2, sum)) * c)
Vqn <- t((Vqr - apply(Vqr, 1, min)) / (
  apply(Vqr, 1, max) - apply(Vqr, 1, min)))
```

Factor scores matrix M_q calculation from W (non-negative least quares)

```
Mq <- matrix(nrow = nrow(X), ncol = q)
for (i in 1:nrow(X)) {
  Mq[i,] = nnls(Vqn, as.vector(t(W[i,])))$X
}
```

Dataset modelling W_m as inner product of M_q and V_{qn}^T

```
Wm <- Mq %*% t(Vqn)
```

Rescaling by calculating scaling factors s and use them with l_s to get V_{qsn}

```
s <- (c - sum(ls[,1])) / apply(Vqn *
  unname(ls[,2] - ls[,1]), 2, sum)
for(i in 1:q) Vqs[,i] <- t(s[i] * t(Vqn[,i]) *
(ls[,2] - ls[,1]) + ls[,1])
  Vqsn <- t(t(Vqs)/ apply(Vqs, 2, sum)) * c
```

Rescaling of factor scores to get M_{qs}

```
Mqs <- t(t(Mq) / s) / apply(
  t(t(Mq) / s), 1, sum)
```

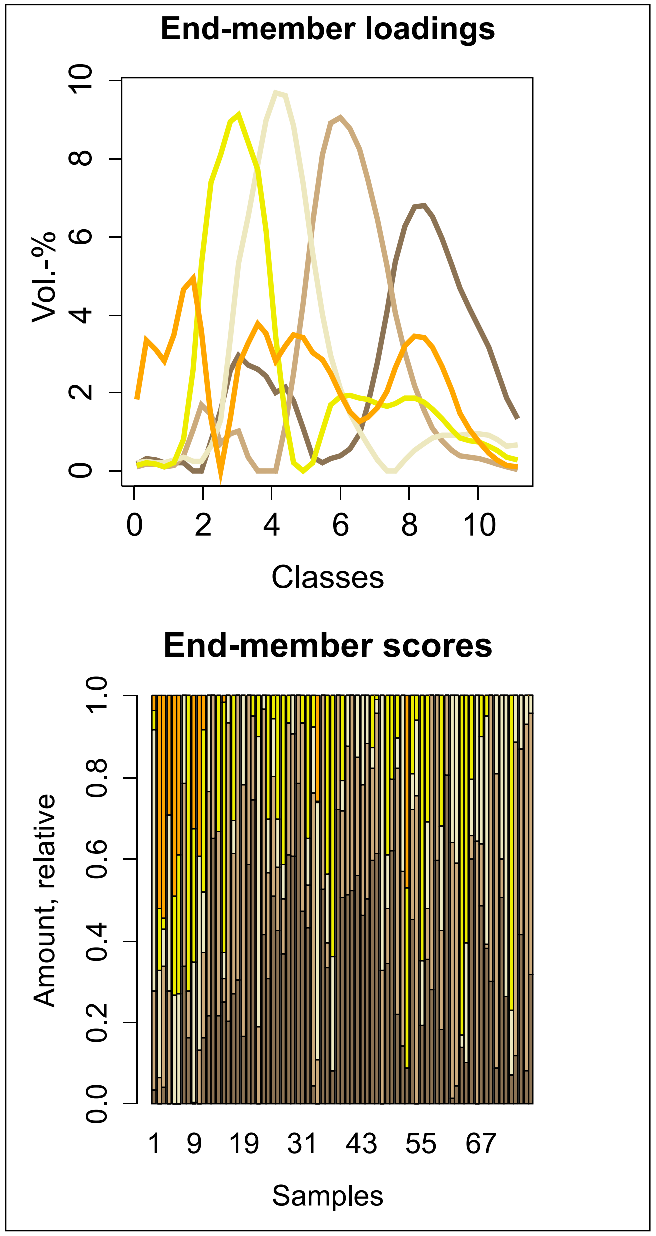
Model values to get matrix X_m

```
Xm <- Mqs %*% t(Vqs)
```

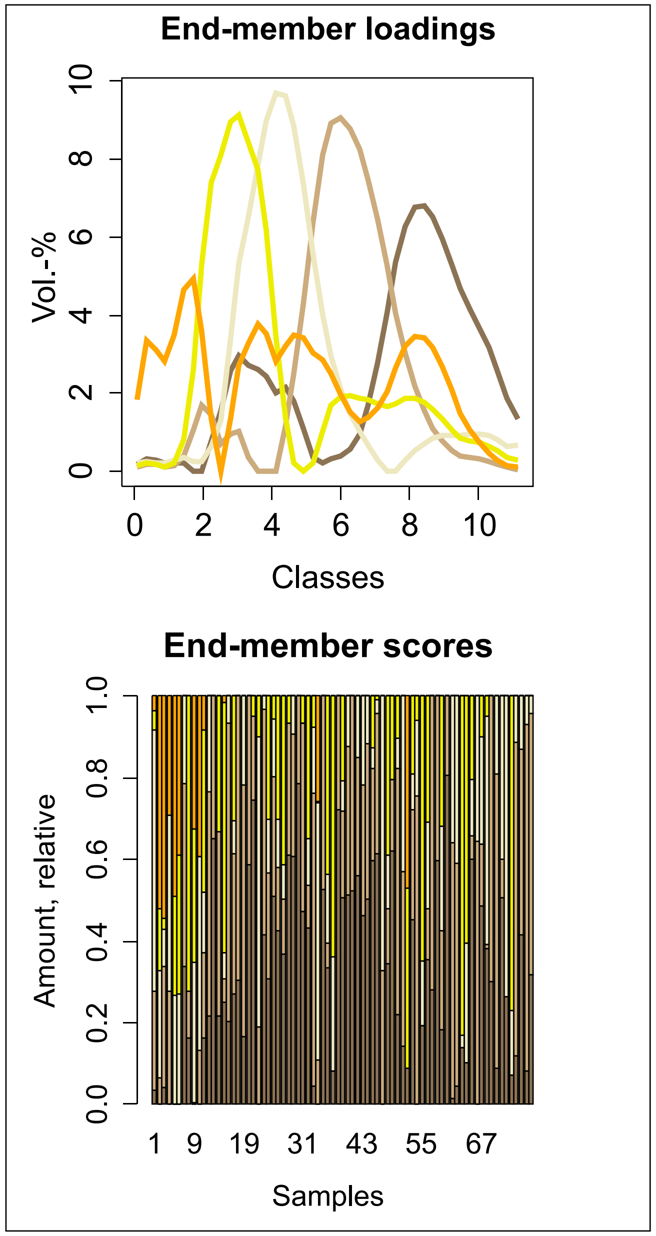
Model evaluation: scores explained variance ($M_{qs\ var}$), absolute model errors (E_m, E_n), explained data variance (R_m, R_n), overlapping modes (ol), mode classes ($modes$)

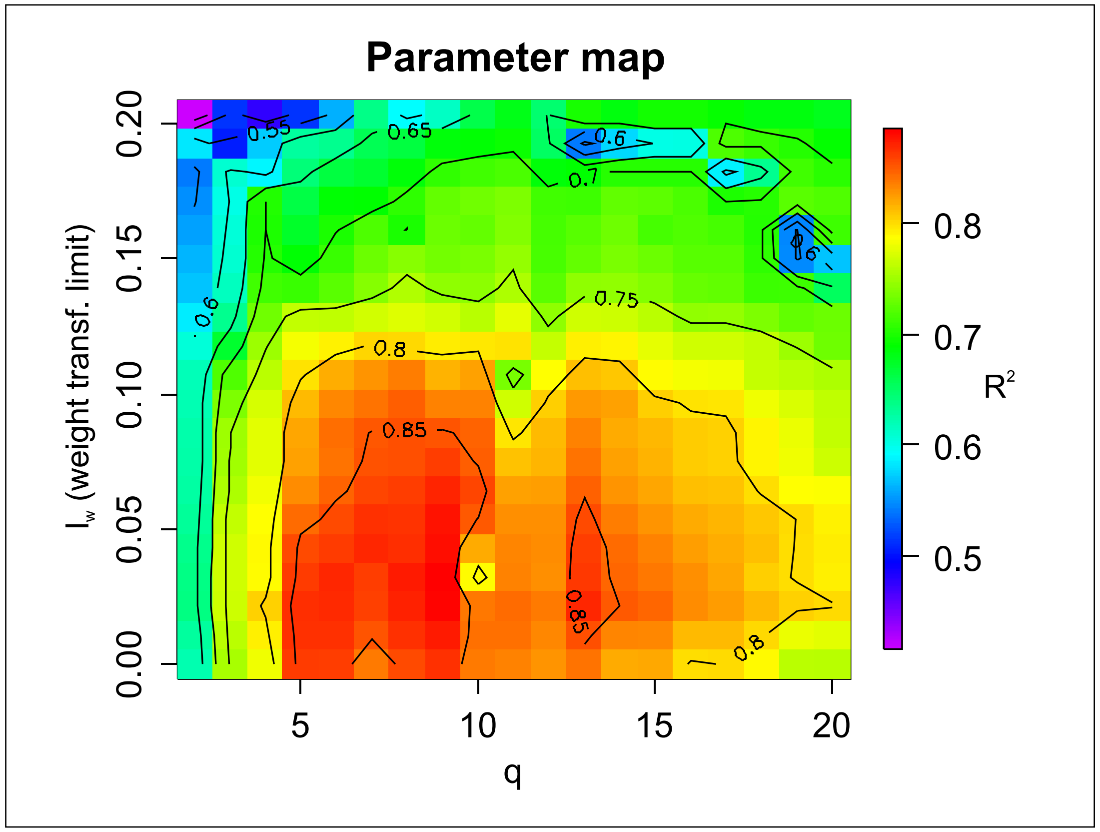
Dietze & Dietze (2013)

From EMMA to robust EMMA

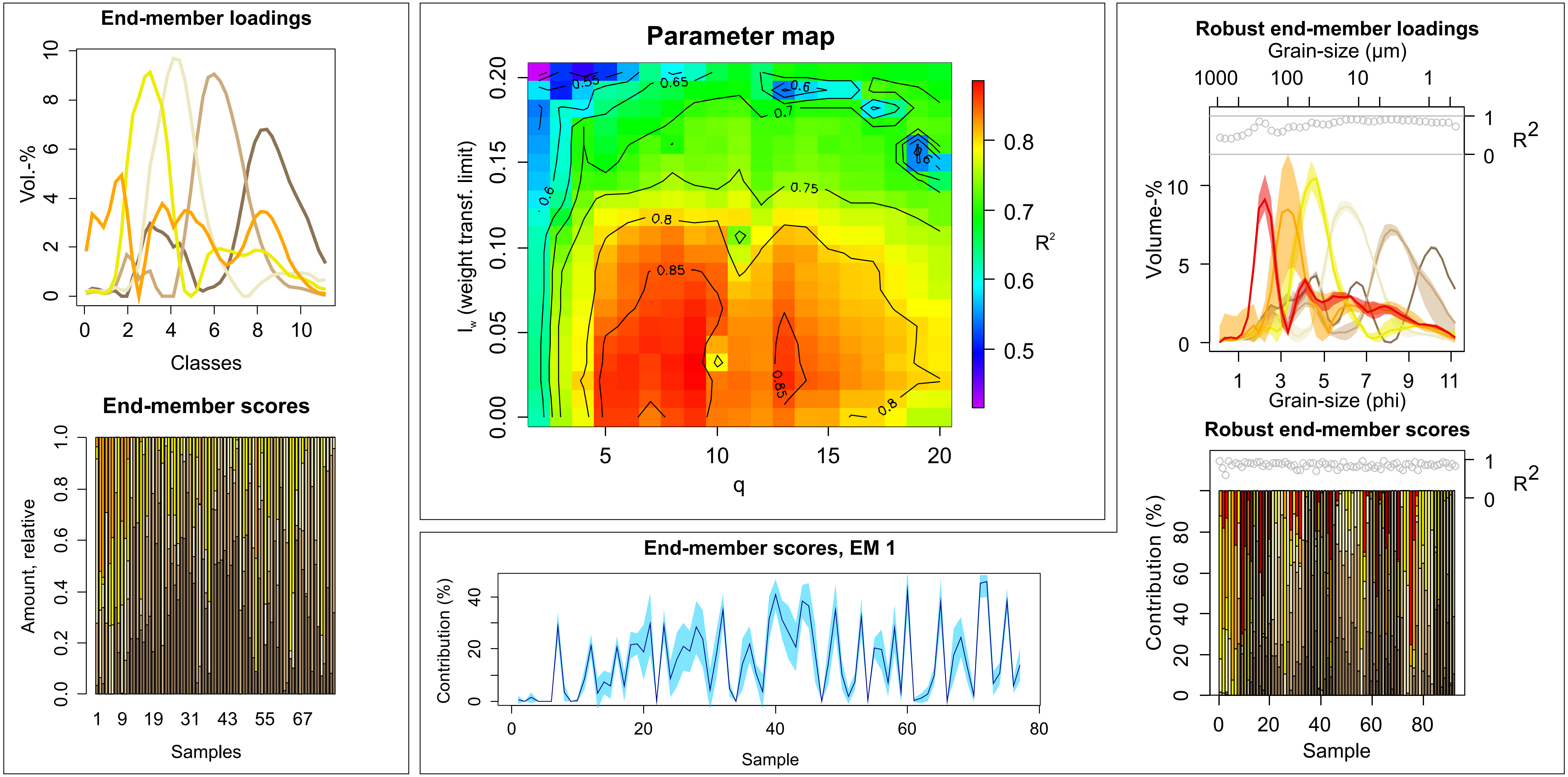


From EMMA to robust EMMA

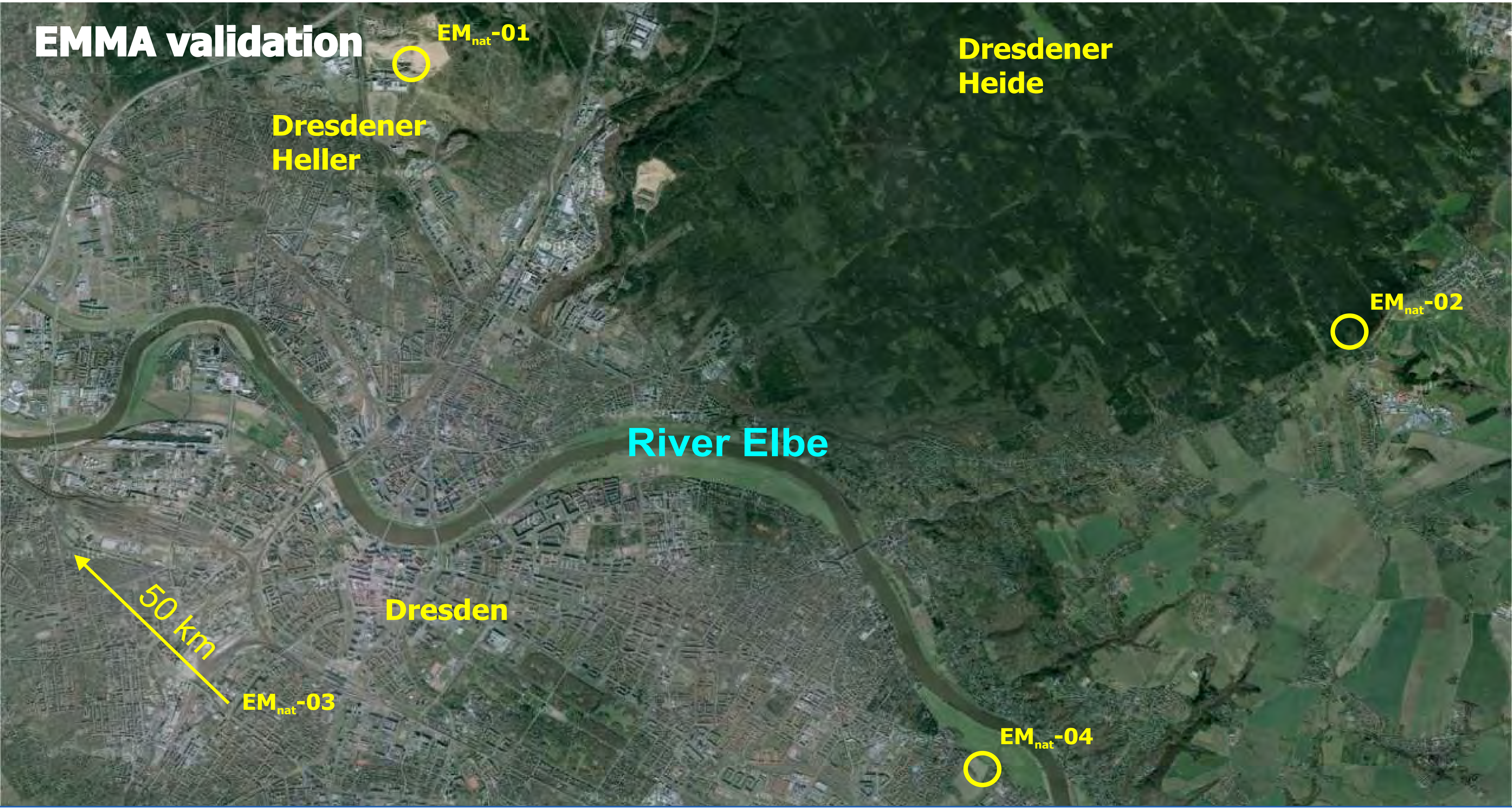




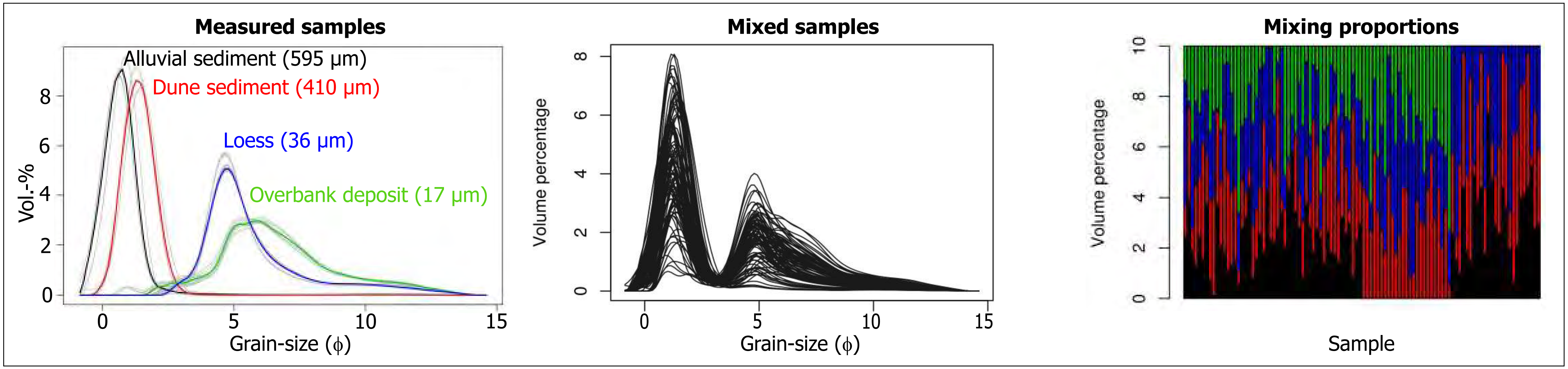
From EMMA to robust EMMA



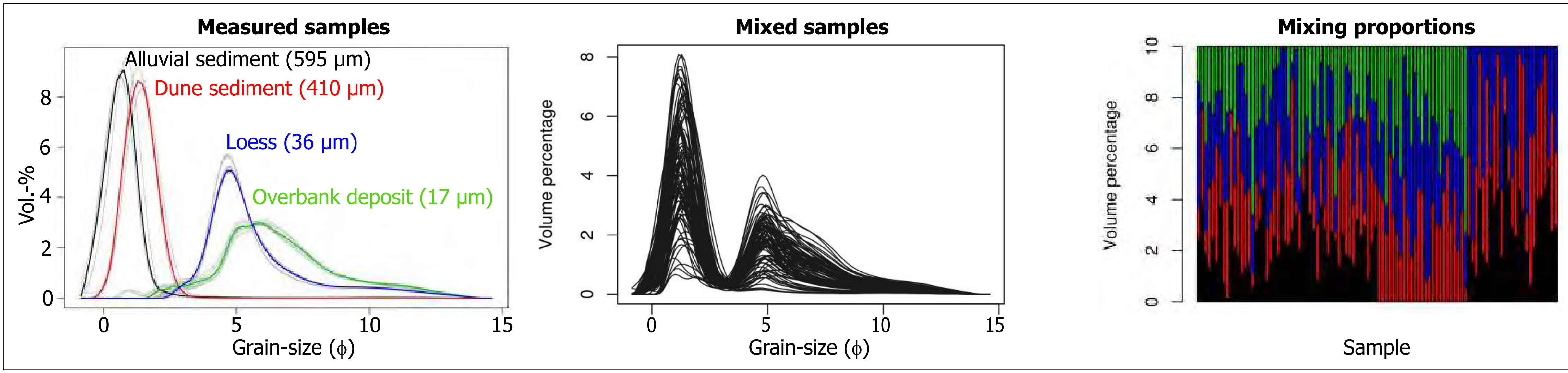
EMMA validation

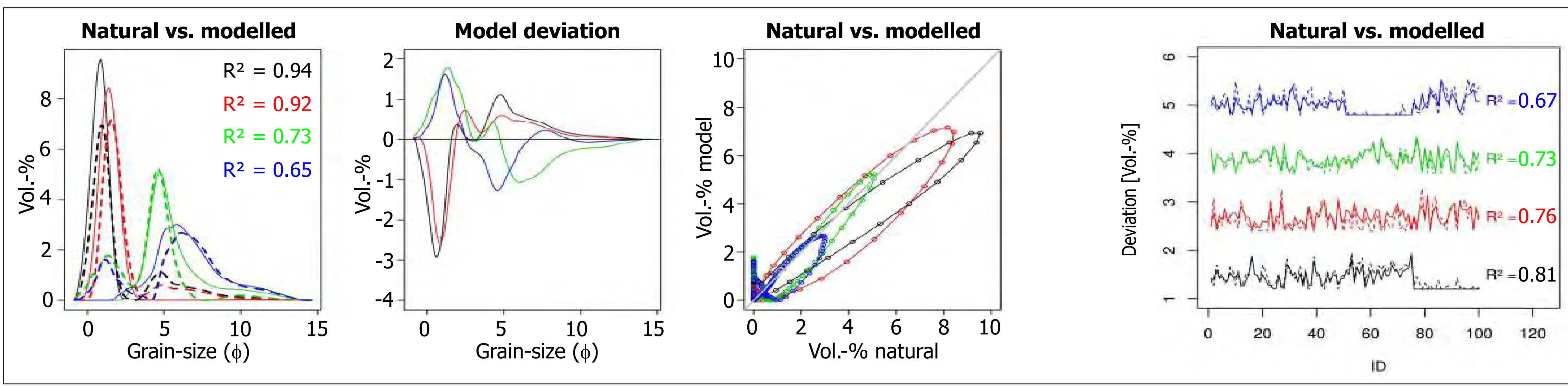


Mixing natural process end-members

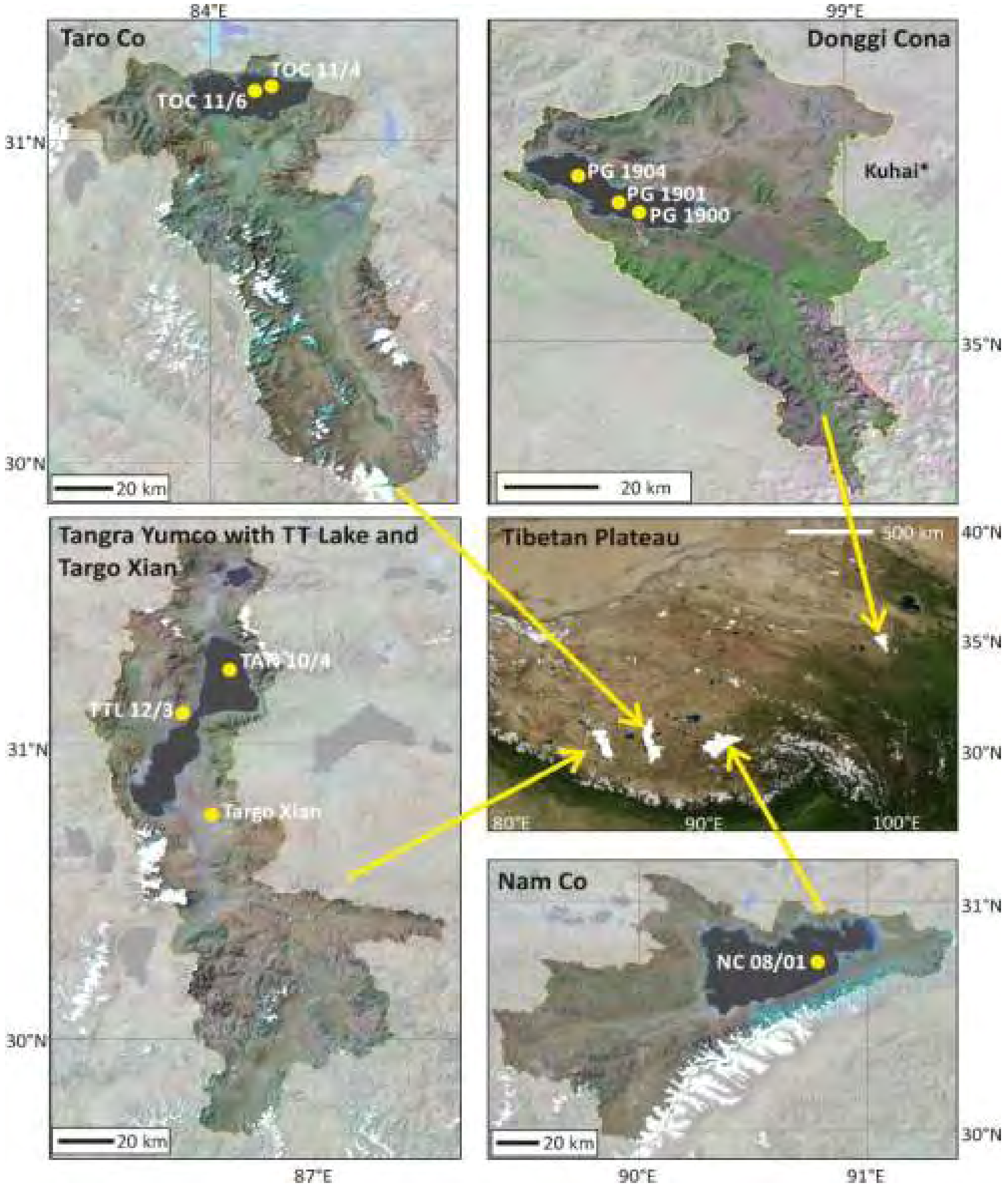


Mixing natural process end-members



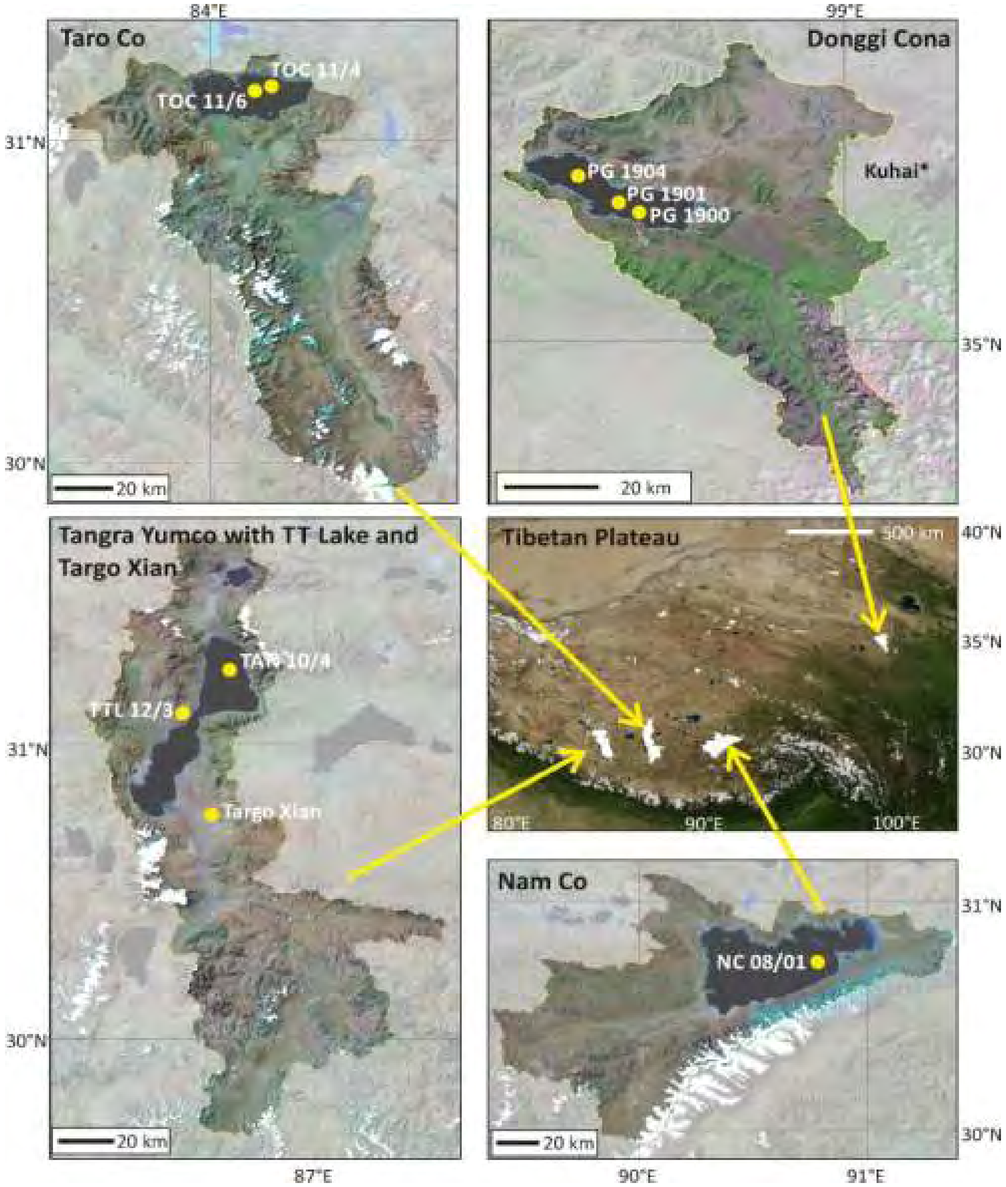


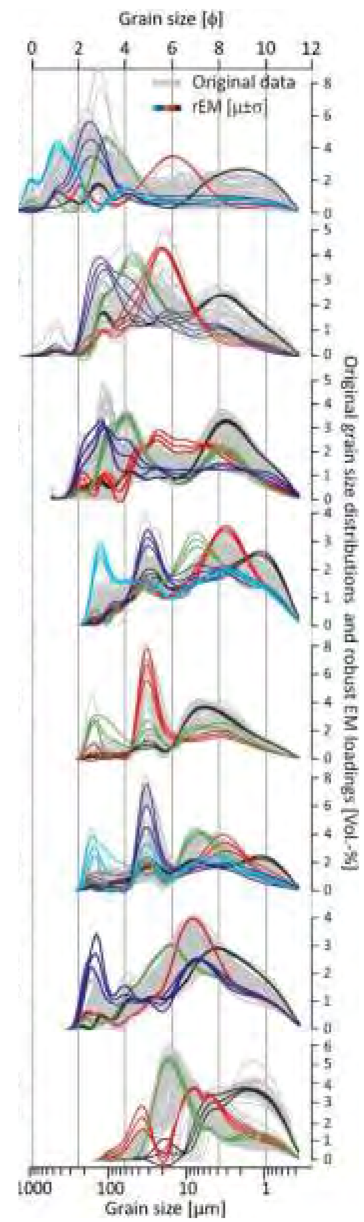
Application I - processes recorded in lakes across the Tibetan Plateau

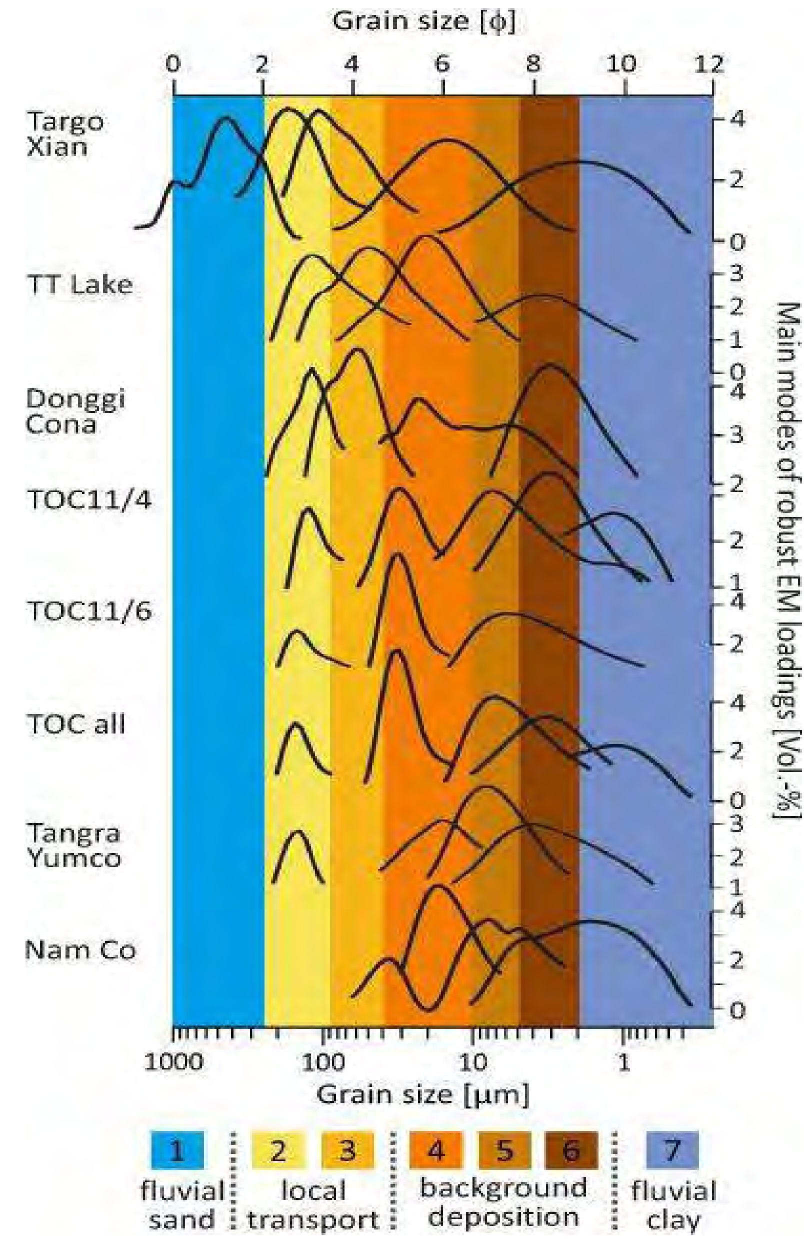


Dietze et al. (2014)

Application I - processes recorded in lakes across the Tibetan Plateau

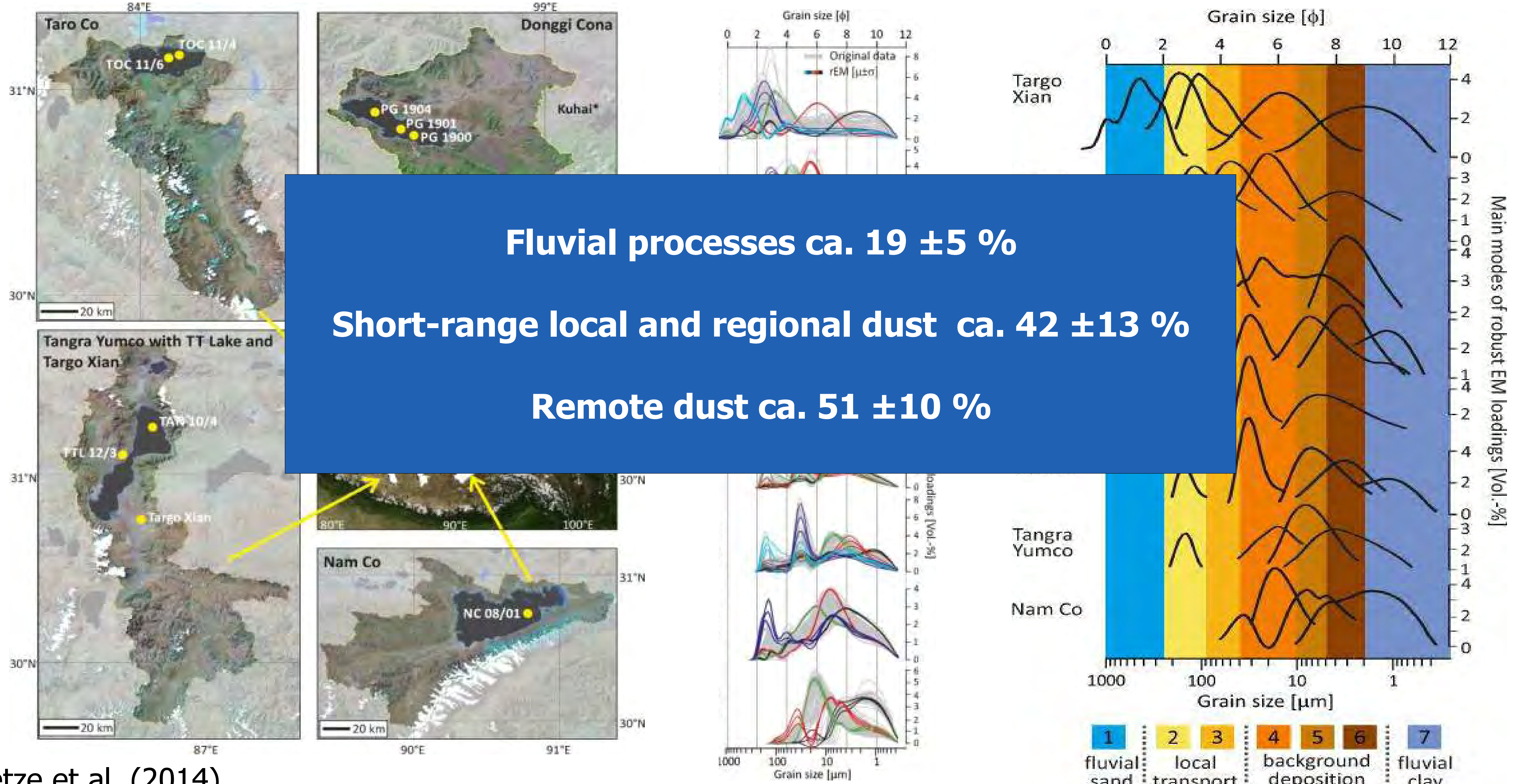






Dietze et al. (2014)

Application I - processes recorded in lakes across the Tibetan Plateau



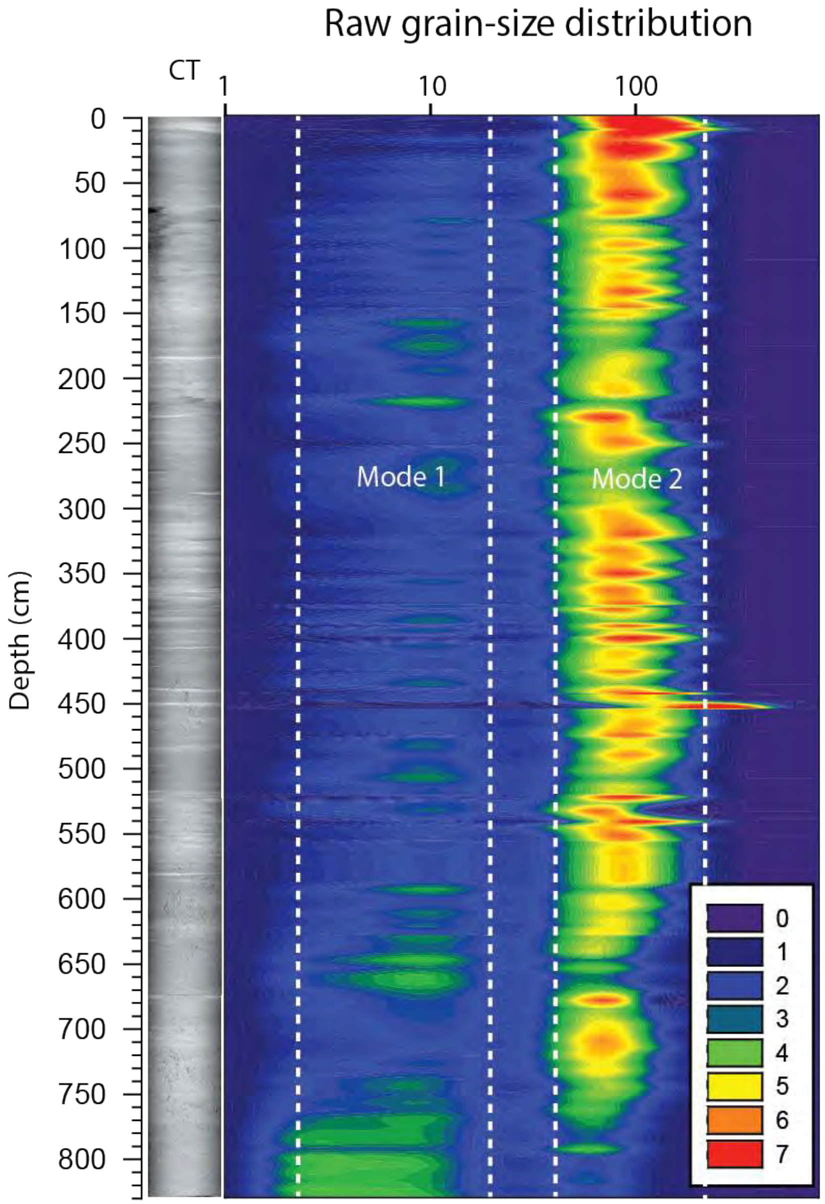
Fluvial processes ca. 19 ±5 %

Short-range local and regional dust ca. 42 ±13 %

Remote dust ca. 51 ±10 %

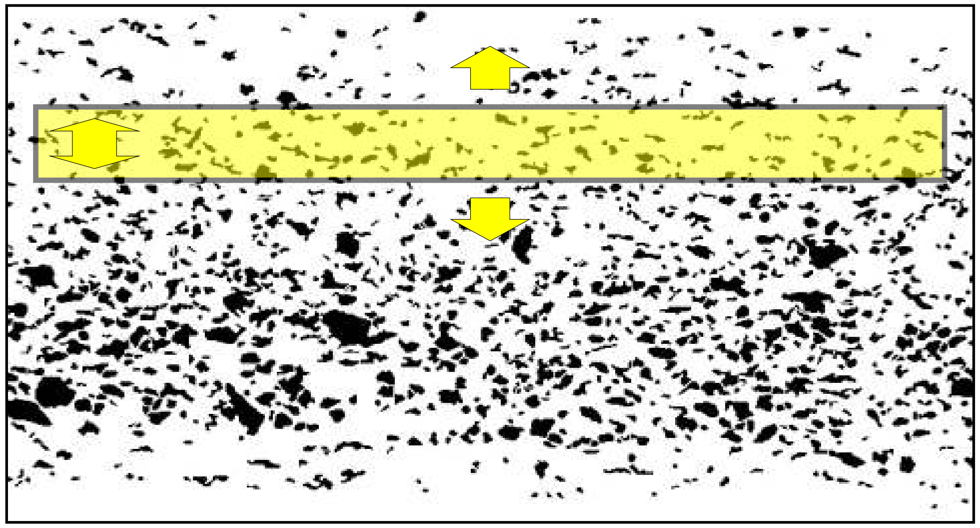
Dietze et al. (2014)

Application II - quasi-continuous EMMA on laminated marine sediments

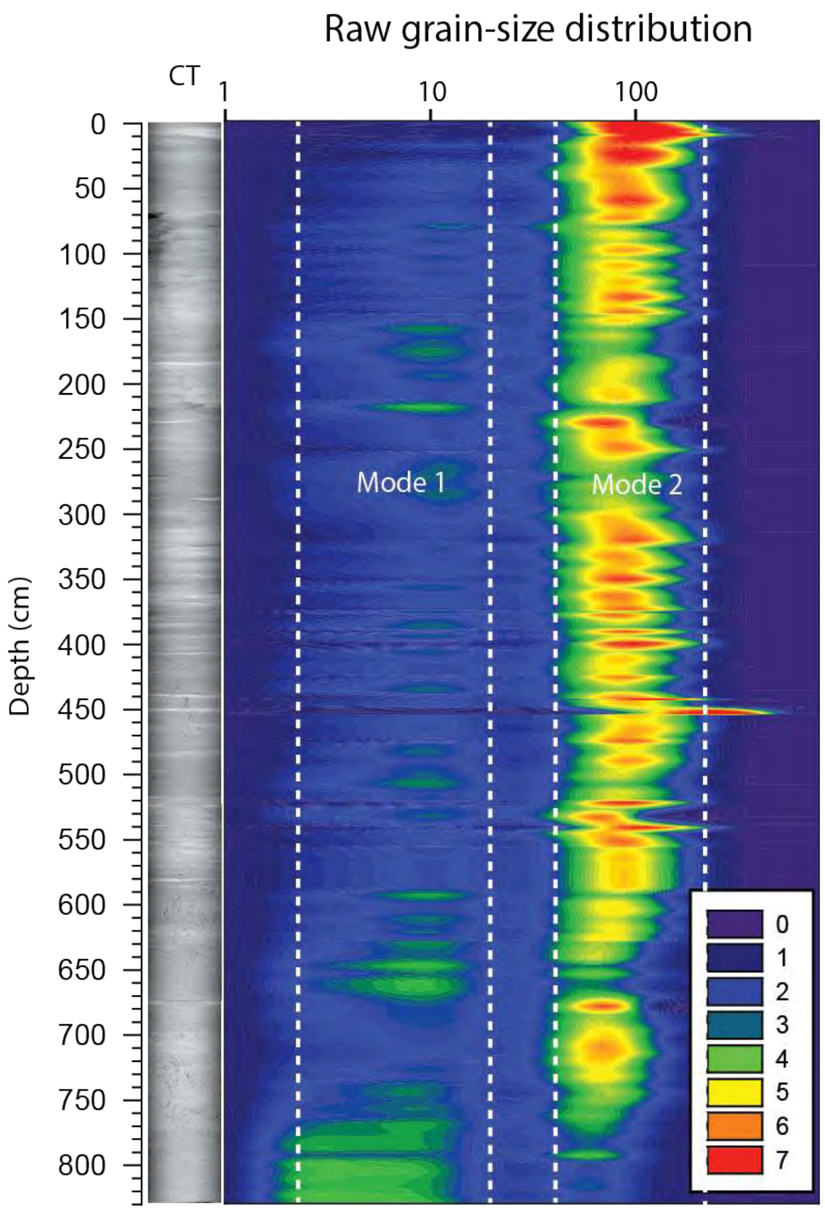


SEM-image classification yields spatially continuous grain-size information. A moving window filter calculates continuous grain-size distributions that are used for EMMA.

based on Lewis et al. (2010)

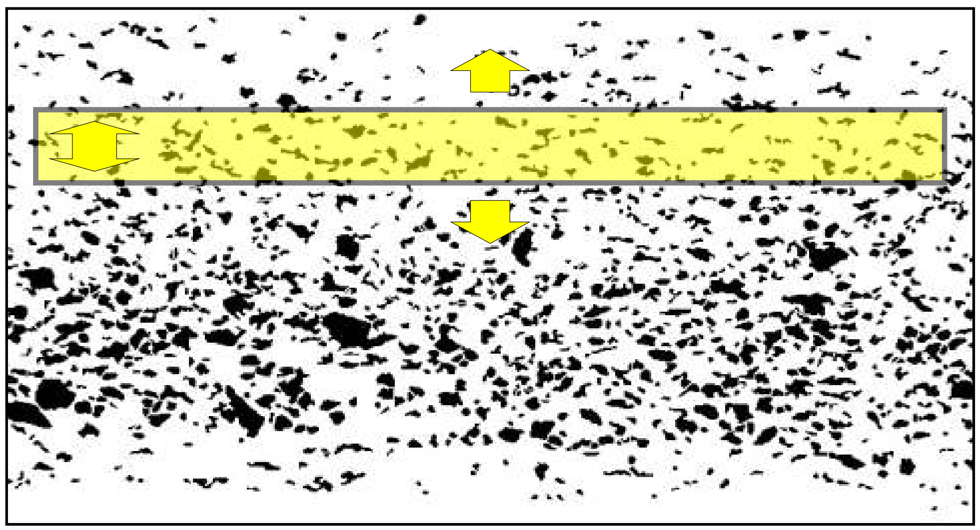


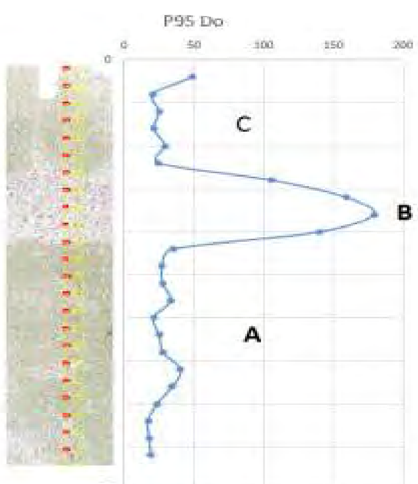
Application II - quasi-continuous EMMA on laminated marine sediments

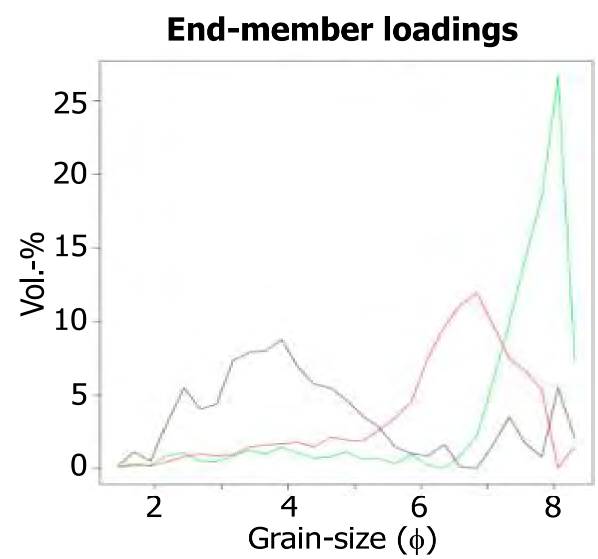


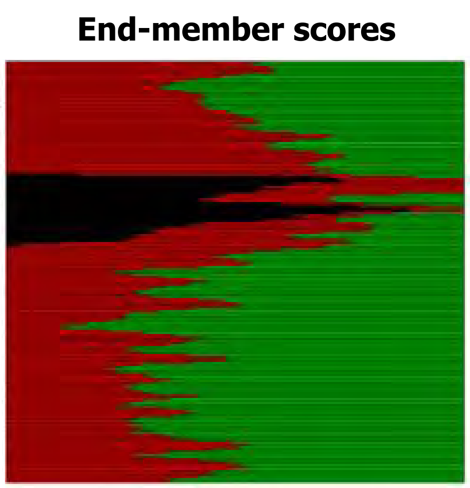
based on Lewis et al. (2010)

SEM-image classification yields spatially continuous grain-size information. A moving window filter calculates continuous grain-size distributions that are used for EMMA.

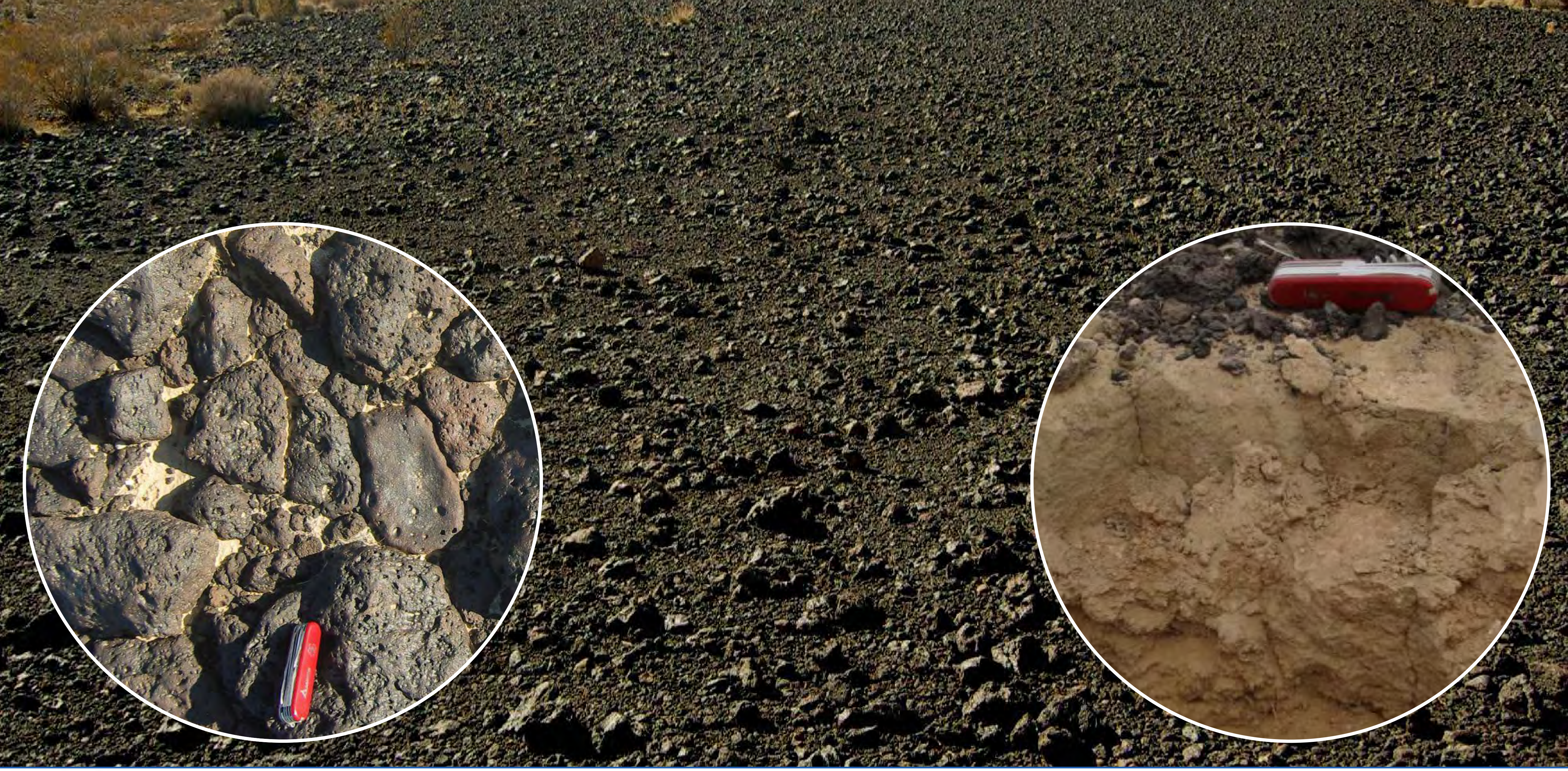




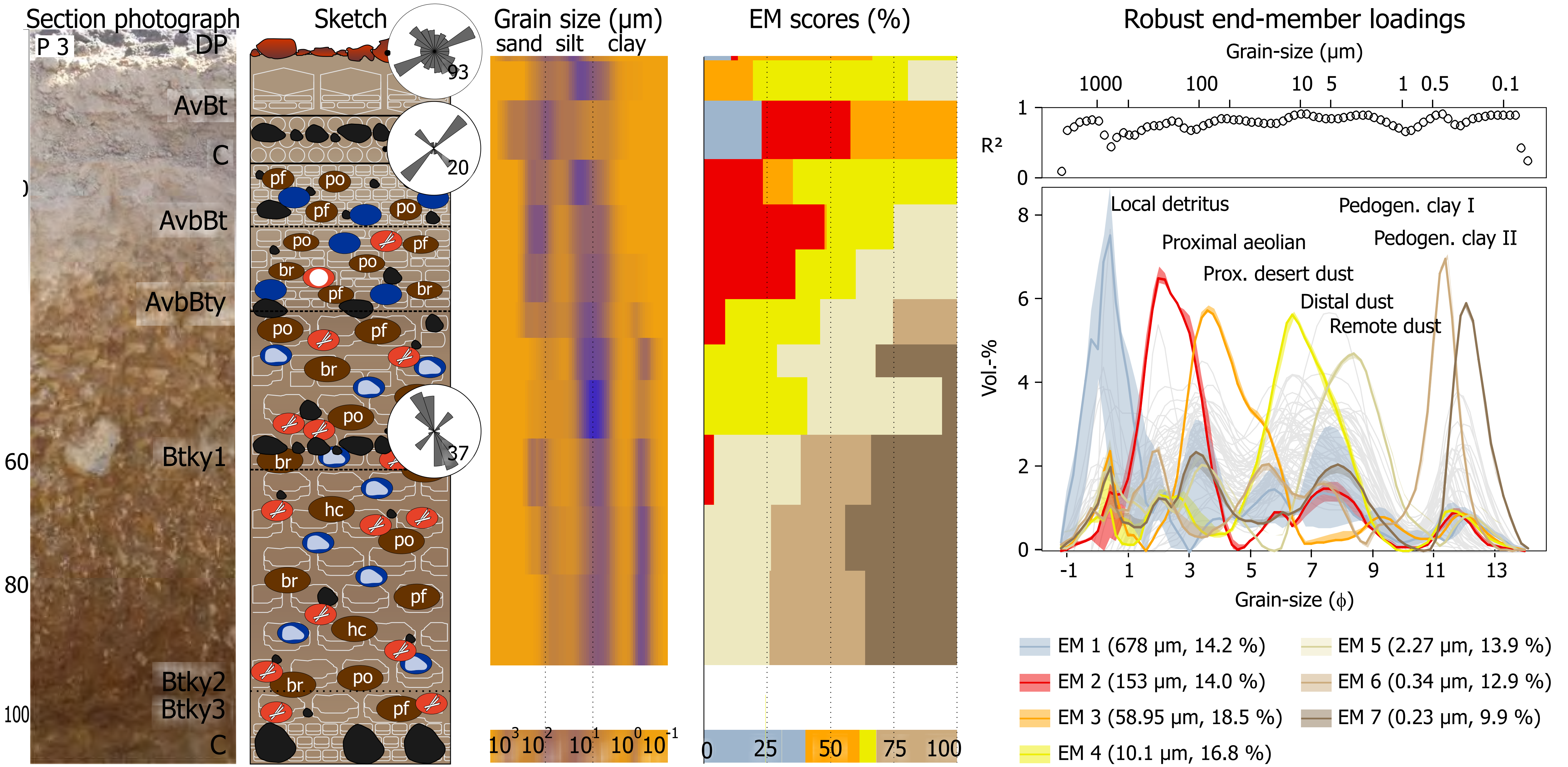




Application III - Stone-covered landforms in deserts

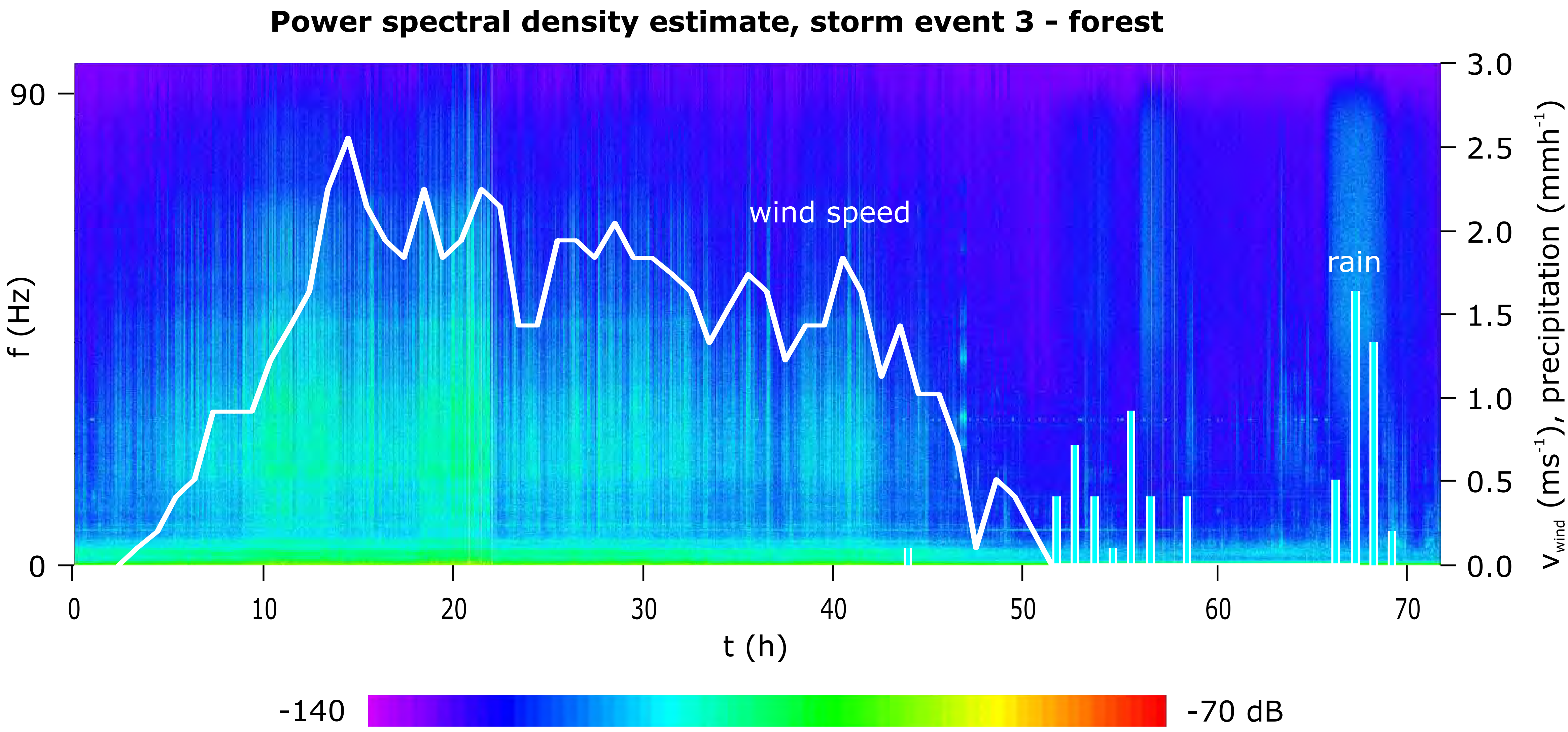


Application III - Stone-covered landforms in deserts

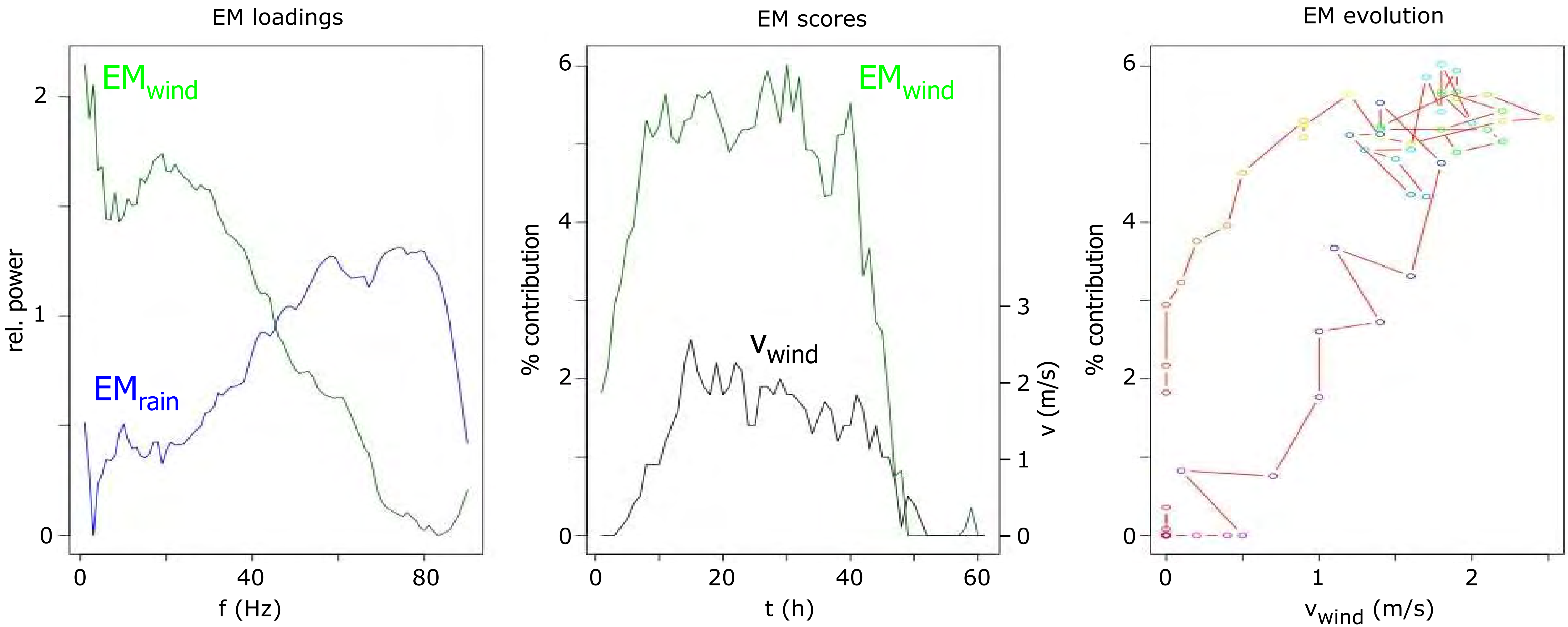


Dietze et al. (2012, 2013)

Application IV - unmixing environmental signals in seismic records



Application IV - unmixing environmental signals in seismic records



Preliminary Summary

Grain-size data is suspect to either redundancy or uncertainty

due to discretisation of a continuous distribution.

Direct linkage of grain-size properties to Earth surface dynamics may be limited

due to multplicity, non-linearity, path-dependency, cascade systems, ...

Preliminary Summary

Grain-size data is suspect to either redundancy or uncertainty
due to discretisation of a continuous distribution.

Direct linkage of grain-size properties to Earth surface dynamics may be limited
due to multplicity, non-linearity, path-dependency, cascade systems, ...

EMMA allows identification and quantification of generic sediment transport regimes
along with estimation of uncertainty (loadings & scores) and inspection of sources of uncertainty.

EMMA can be applied to a variety of depositional systems and data beyond grain-size
given data transformation is possible to fulfil EMMA constraints.

Thank you!