Seasonal Forecasting of Ethiopian rainfall: Statistical vs ENSEMBLE system

Gulilat Tefera Diro [1], David Grimes[2], Emily Black[2] and Adrian Tompkins[1]
gtefera@ictp.it

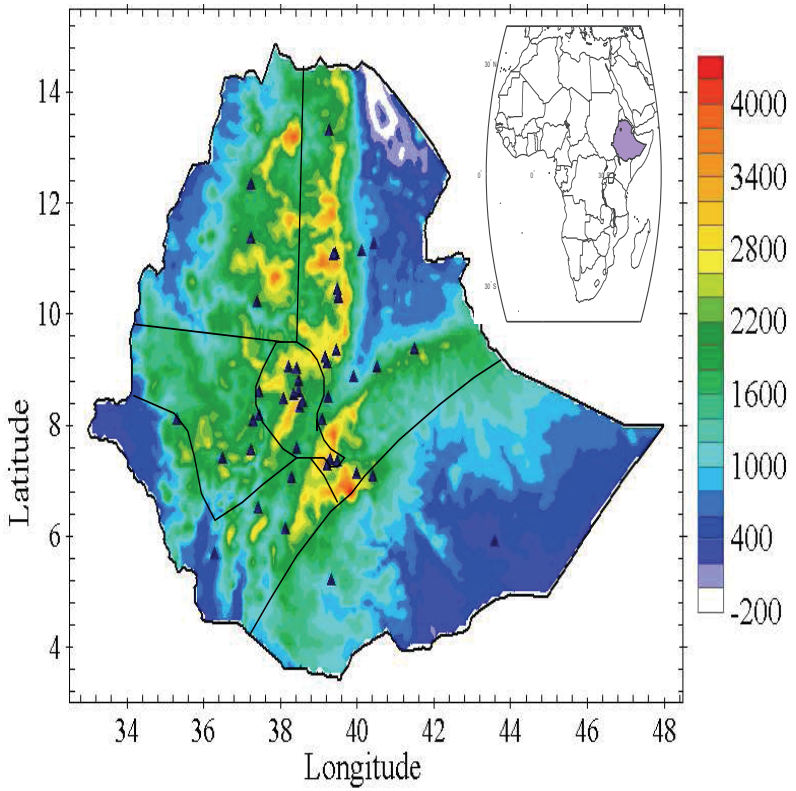
[1] Earth System Physics Section, The abdus salam ICTP
[2] Department of Meteorology, University of Reading

Outline

- Introduction
- Statistical Seasonal Forecasting Method
  - Homogeneous rainfall zones
  - Predictor Identification and Selection
  - Forecasting Models
  - Skill Assessment
- Dynamical seasonal re-forecasts
  - ECMWF ENSEMBLE hindcast
- Summary and future work

Introduction

- Rainfall is the most important climate parameter
  - 85% of the population is involved in agricultural activities
  - 98% of the Energy comes from Hydropower
- but it exhibits high spatial and temporal variation
- Need for early warning system!

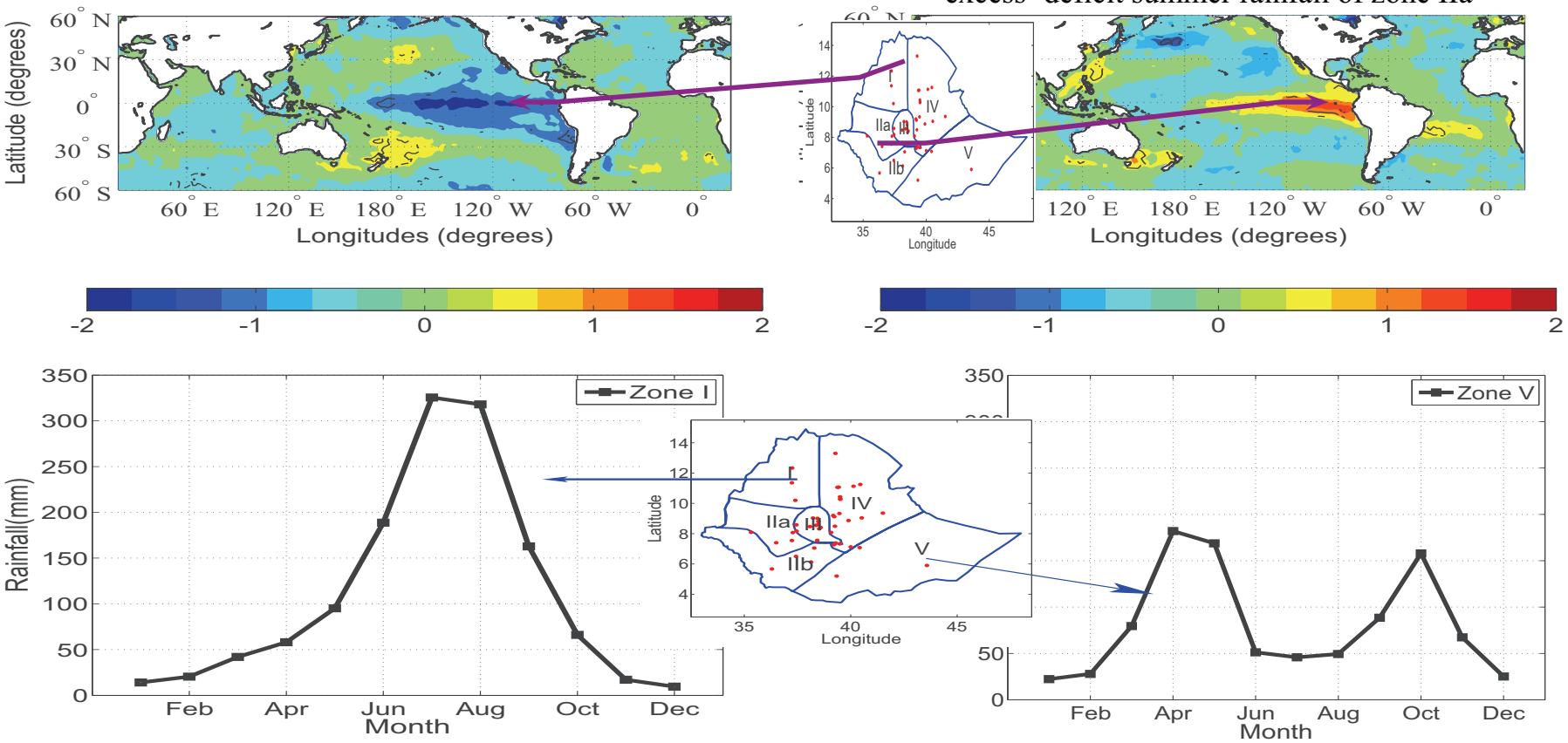


Identifying homogeneous rainfall zones

- Challenge: high spatial variability

Composite of previous winter SST based on excess−deficit summer rainfall of of zone I

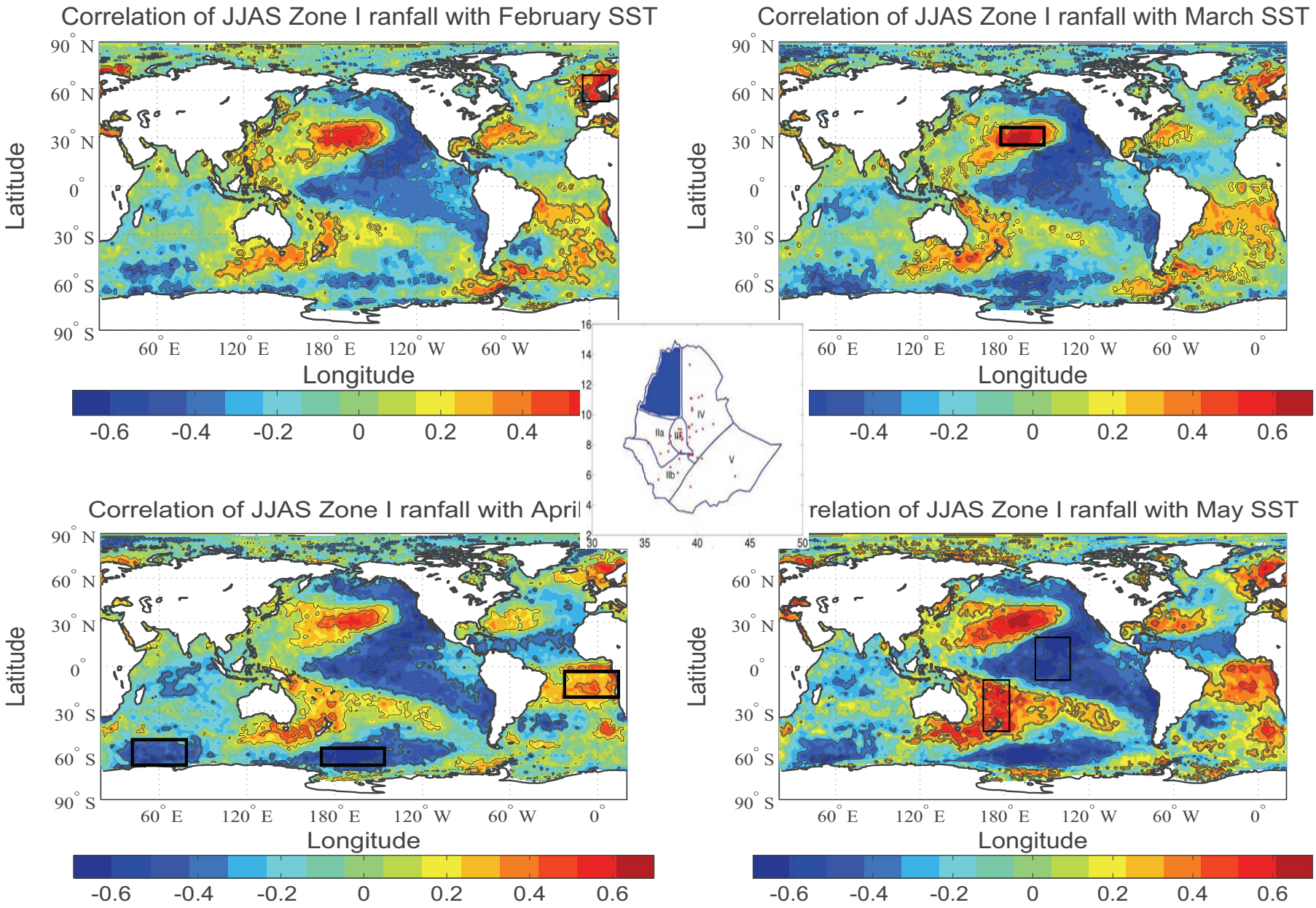
Composite of previous winter SST based on excess−deficit summer rainfall of zone IIa



- Zoning is carried out based on:
  - Seasonal cycle
  - Inter-annual variability

(Gissila et al., 2004)

Predictors Identification

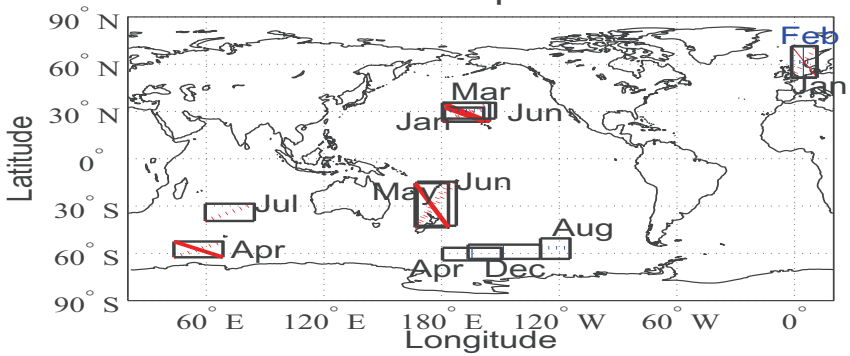


Selection procedure

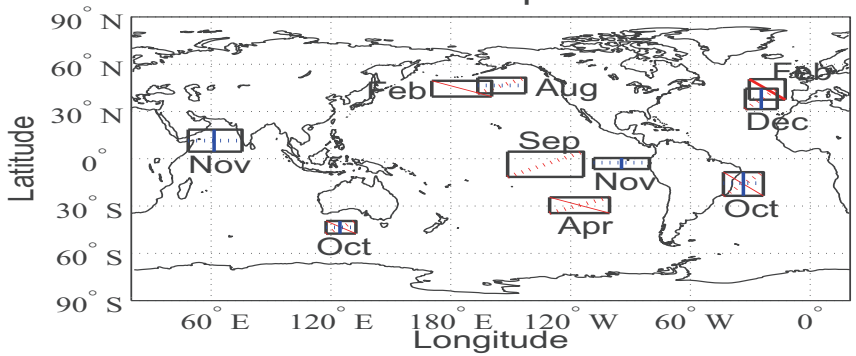
- Four sets of Predictors (A', B', A and B) are created using:
  - A': stepwise regression and including predictors from contemporaneous season
  - B': stepwise discriminant analysis and including predictors from contemporaneous season
  - A : stepwise regression and excluding predictors from contemporaneous season
  - B : stepwise discriminant analysis and excluding predictors from contemporaneous season

Selected Predictors

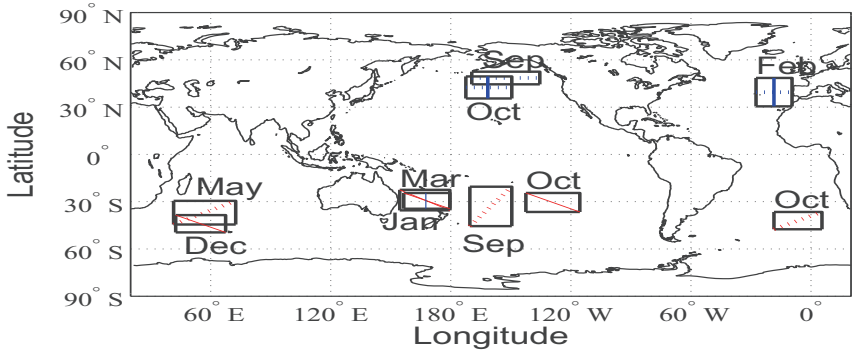
Kiremt Zone I predictors



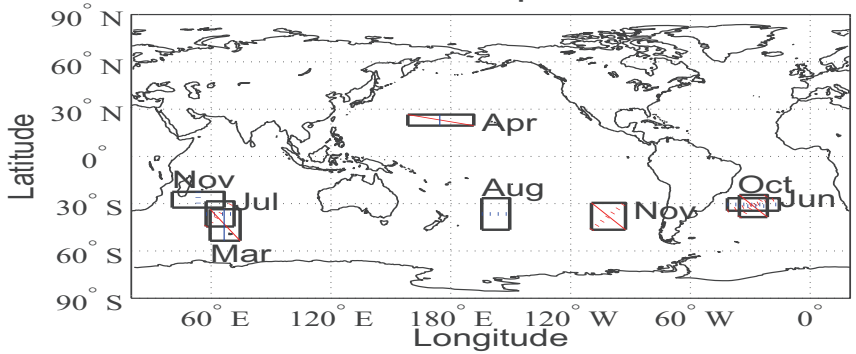
Kiremt Zone IIa predictors



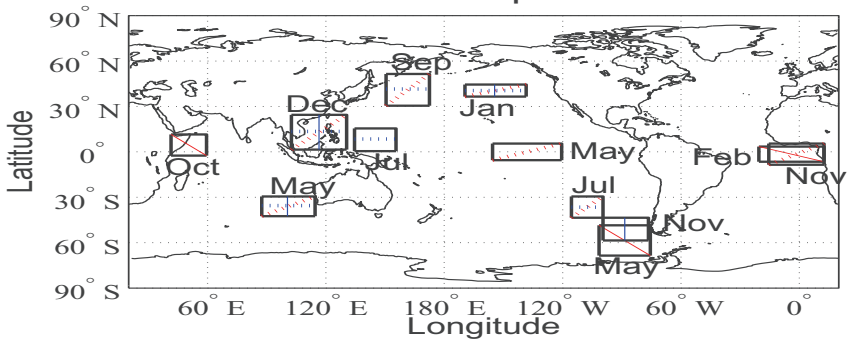
Kiremt Zone IIb predictors

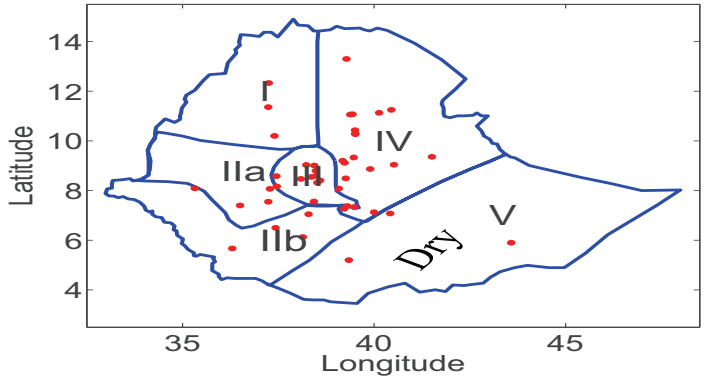


Kiremt Zone III predictors



Kiremt Zone IV predictors





Stepwise regression

Stepwise discriminant analysis

Common to both

Including JJAS predictors

Forecasting Models

Linear Discriminant Analysis (LDA)

$$Pr(W_i|X) = \frac{Pr(W_i)f(X|W_i)}{\sum_j Pr(W_j)f(X|W_j)} \quad ; ln(f(X|W_i)) = \alpha_{i1}x_1 + \alpha_{i2}x_2 + \cdots + \alpha_{in}x_n + \Gamma_i$$

$$Pr(W_i|X) = \frac{e^{d_i}}{\sum_j e^{d_j}} \quad ; where\ d_i = \alpha_{i1}x_1 + \alpha_{i2}x_2 + \cdots + \alpha_{in}x_n + \Gamma_i + ln(Pr(W_i))$$

Where $\alpha_{ij} = S_{ij}^{-1}\bar{x}_i$ and $\Gamma_i = -\frac{1}{2}\bar{x}_i^T S^{-1}\bar{x}_i$

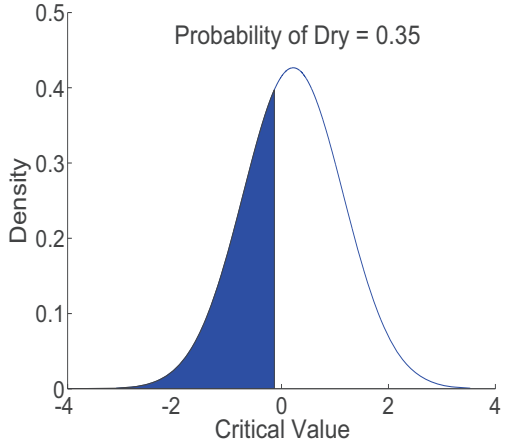
d=discriminant score, $Pr(W_i)$ = apriori probability, W_i = the i^{th} category, f = density function

Multiple Linear Regression (MLR)

$$\mathbf{Pr(Y|X) \sim N(\hat{Y}, \sigma^2)}$$

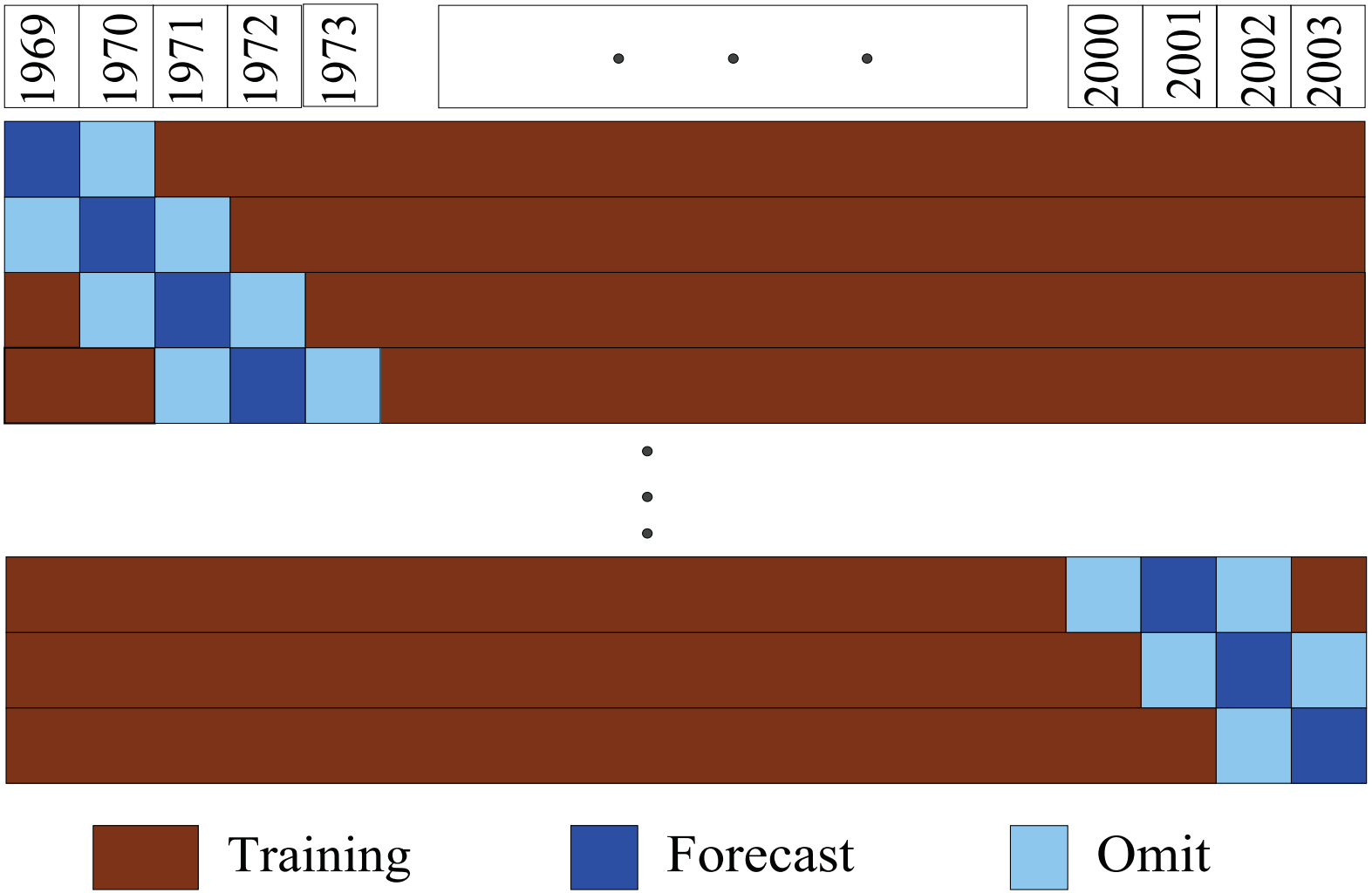
$$where; \quad \mathbf{\hat{Y} = \hat{\beta}X} \quad and \quad \mathbf{\sigma^2 = \sigma_o^2[1 + X_o'(X'X)^{-1}X_o]}$$

Y = Rainfall ; X= predictors (SSTA); β = regression parameter



Hindcast

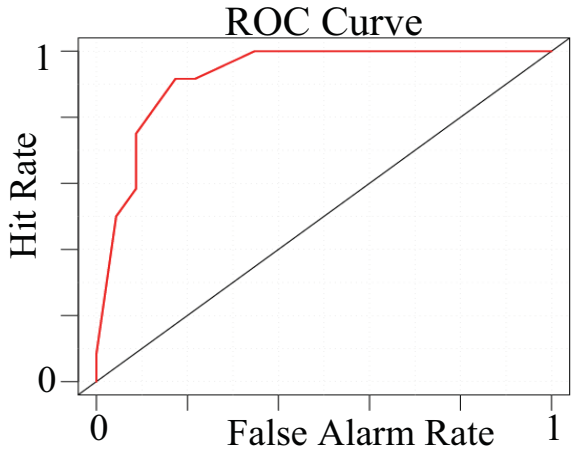
- Cross-validation with leaving 3 years out



Skill Assessment

Relative Operational Characterstics Score (ROCS)

- Area under ROC curve
- Compares against a random forecast
- For a skillful forecast $\rightarrow$ ROCS > 0.5



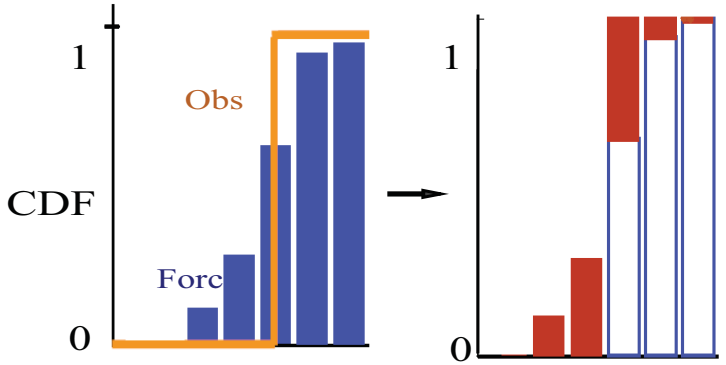
	Dry		Normal		Wet	
Predictors	MLR	LDA	MLR	LDA	MLR	LDA
Set A'	**0.90**	**0.87**	*0.74*	0.64	**0.92**	**0.84**
Set B'	**0.88**	**0.84**	0.70	0.63	**0.92**	**0.88**
Set A	**0.83**	**0.81**	*0.70*	0.59	**0.85**	**0.83**
Set B	**0.85**	**0.81**	0.61	0.59	**0.83**	**0.82**

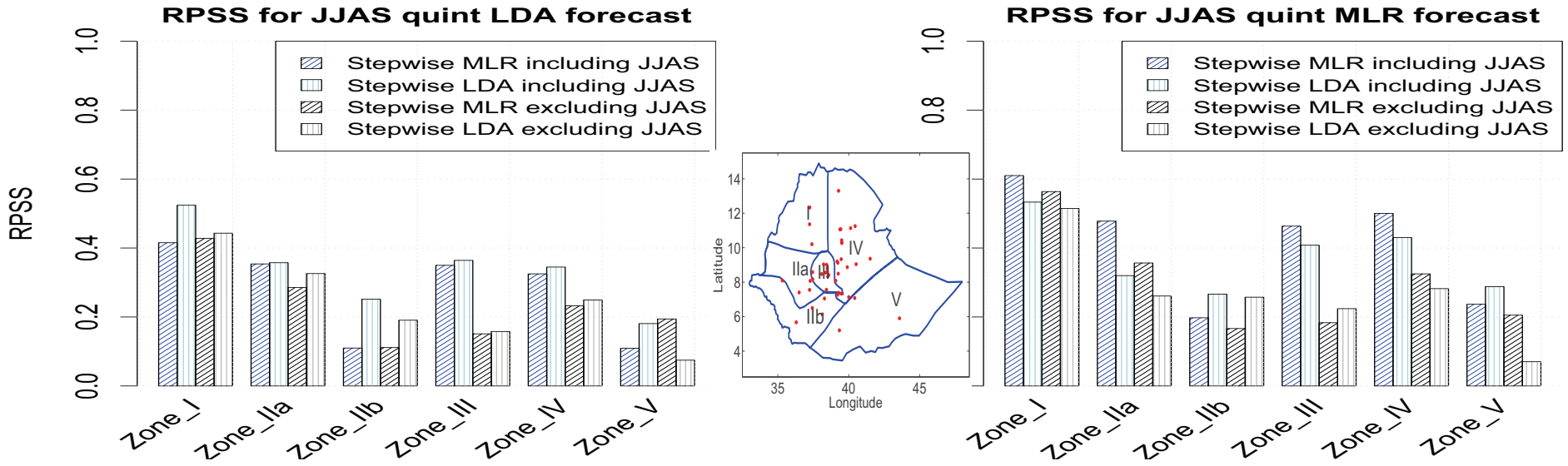
- Better than a random forecast
- Lower skill in the near Normal category

Skill Assessment II

Ranked Probability Skill Score (RPSS)

- $RPSS = 1 - \frac{RPS_{forc}}{RPS_{clim}}$
- $RPS = \frac{1}{K-1}[\sum_{i=1}^{K}(CDF_{forc,i} - CDF_{obs,i})^2]$
- For a skillful forecast $\rightarrow$ RPSS > 0





- Better than climatology
- Highest skill for Zone I

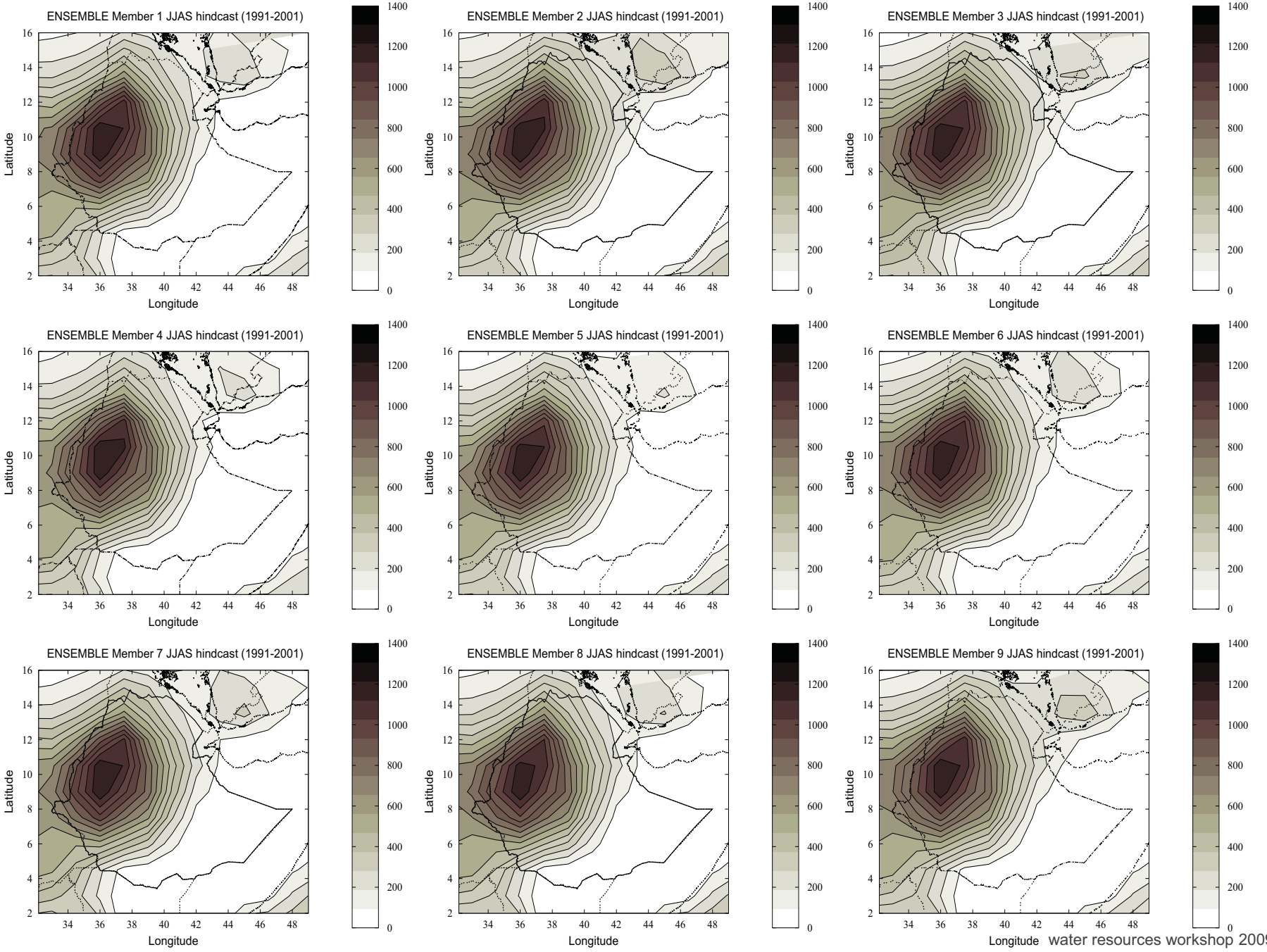
Summary on statistical forecasting

- Due to high spatial variation forecasting should be done for each homogeneous rainfall zones separately
- Linear approach yields a better skill compared to climatology or random forecast
- Outer categories have more skill than the inner ones
- The skill of the forecast is highest for northwest
- MLR tend to slightly outperform LDA
- Including contemporaneous SST has a huge impact for central part of Ethiopia

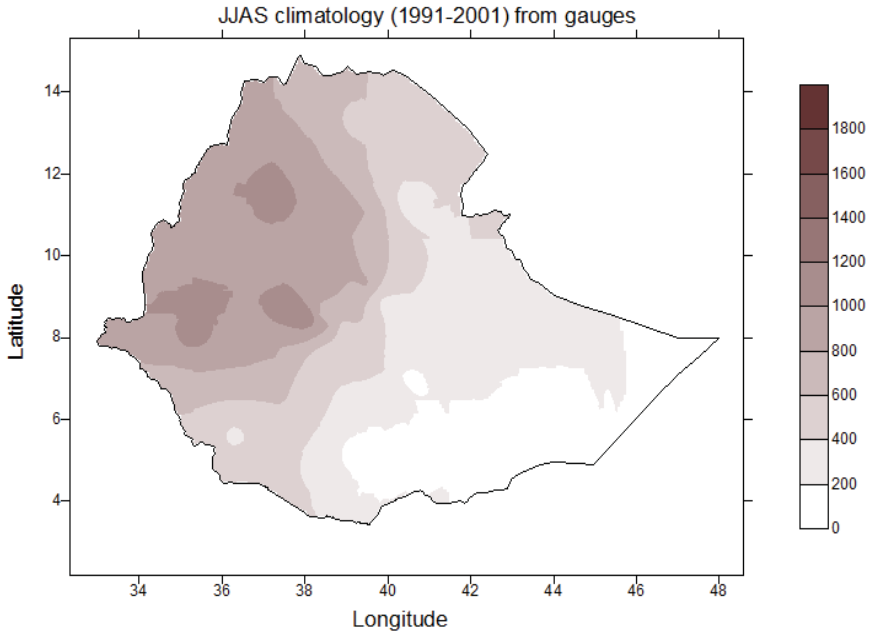
ECMWF ENSEMBLE Hindcast

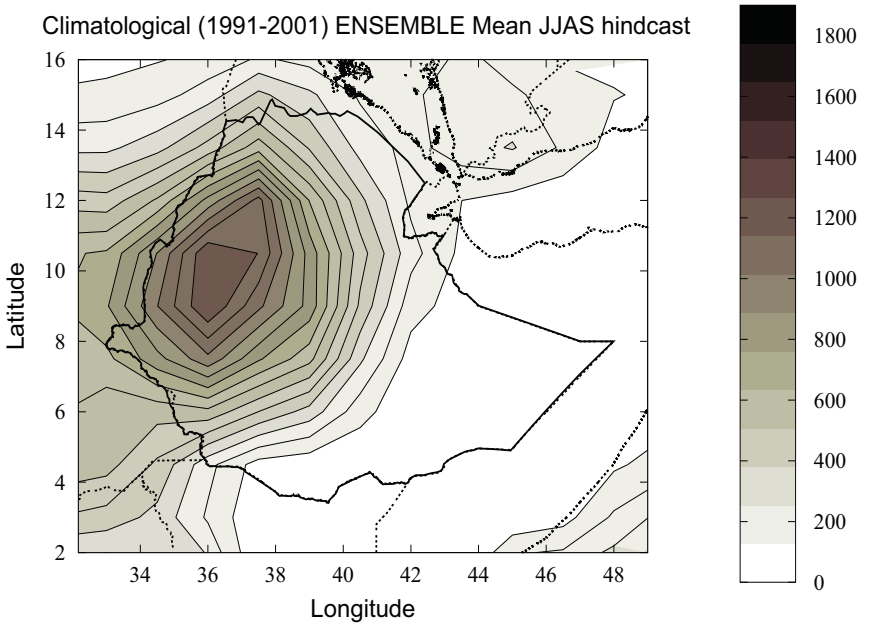
- Resolution : $1.5^0 \times 1.5^0$ in horizontal and 40L in vertical
- Hindcast period: 1991-2001
- 9 member ensembles
- addressing forecast uncertainity
  - uncertainity in initial condition: Perturbed initial conditions
  - model error: Perturbed physics
- Two start dates (May and November): we use the May 1st start
- 6 month hindcast (May 1st - November 1st)

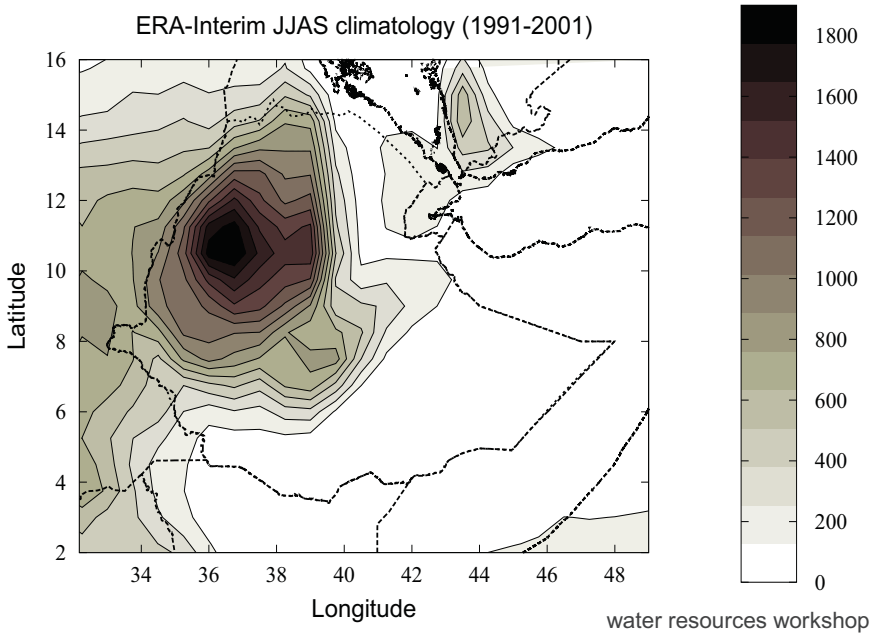
Spatial variability: Climatology



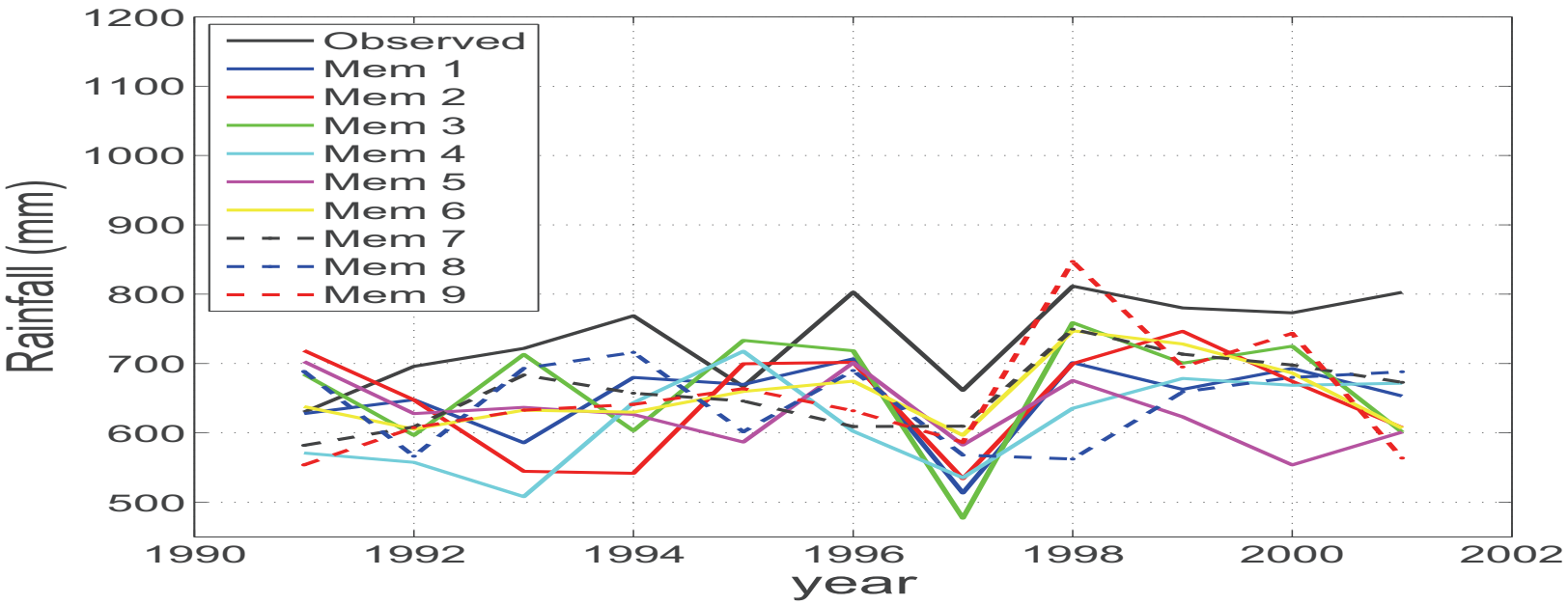
Ensemble mean climatology vs ERA-Interim





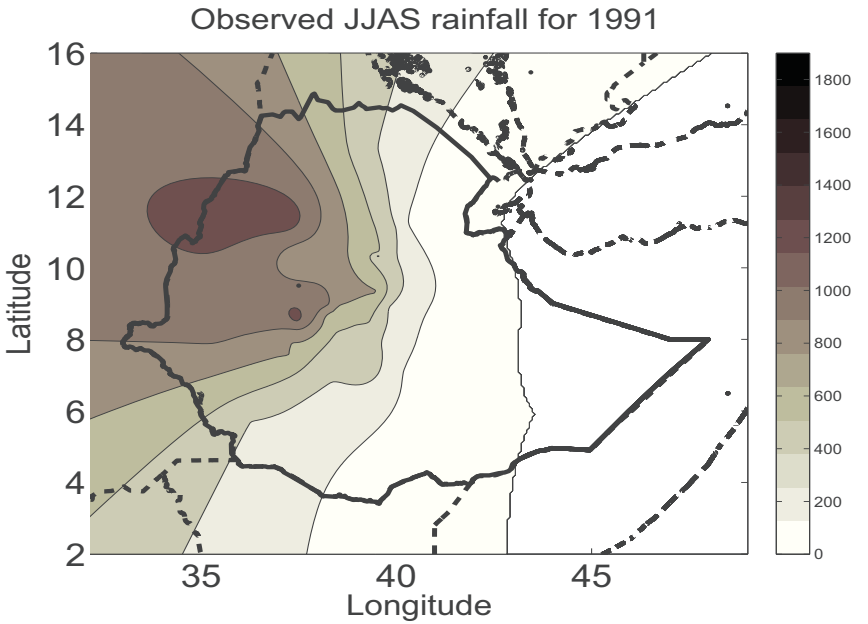


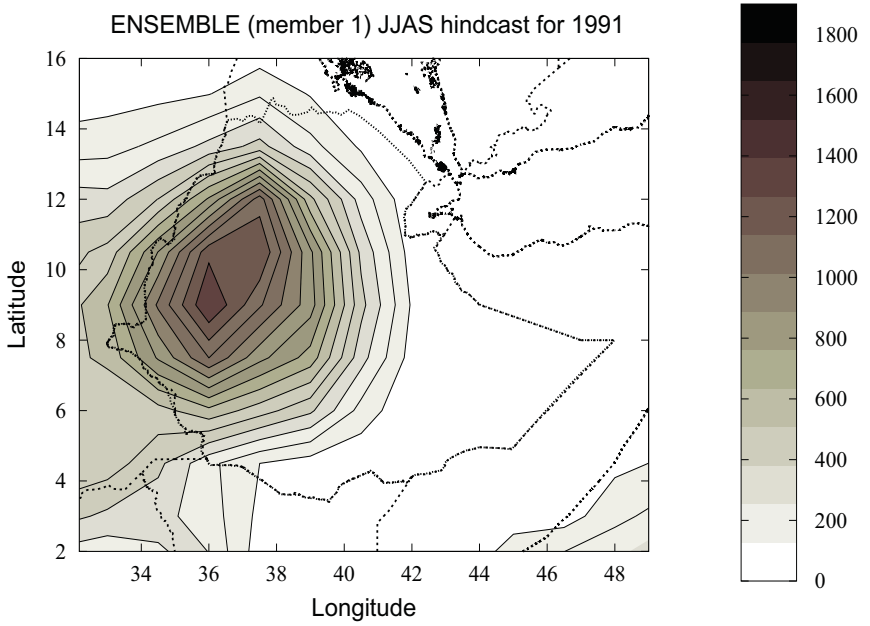
Interannual variability

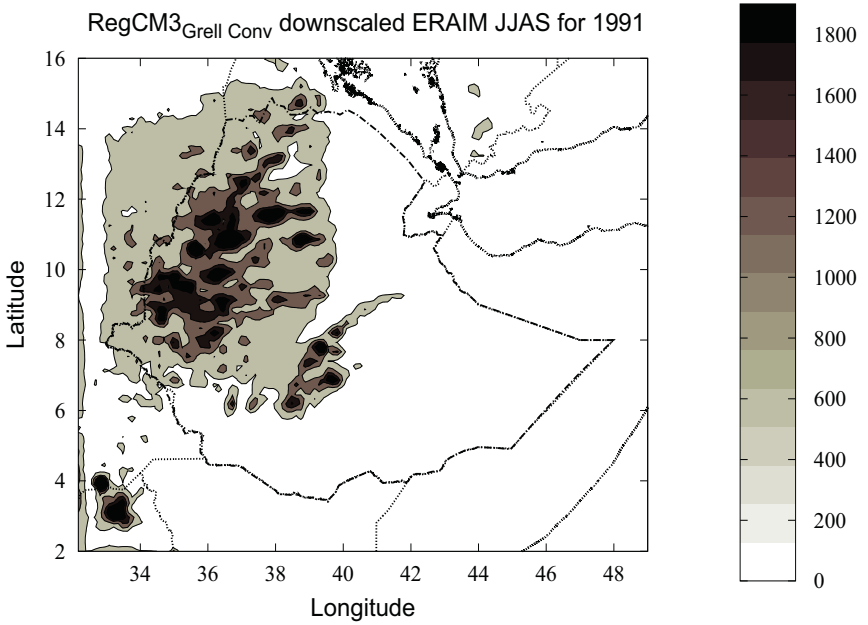


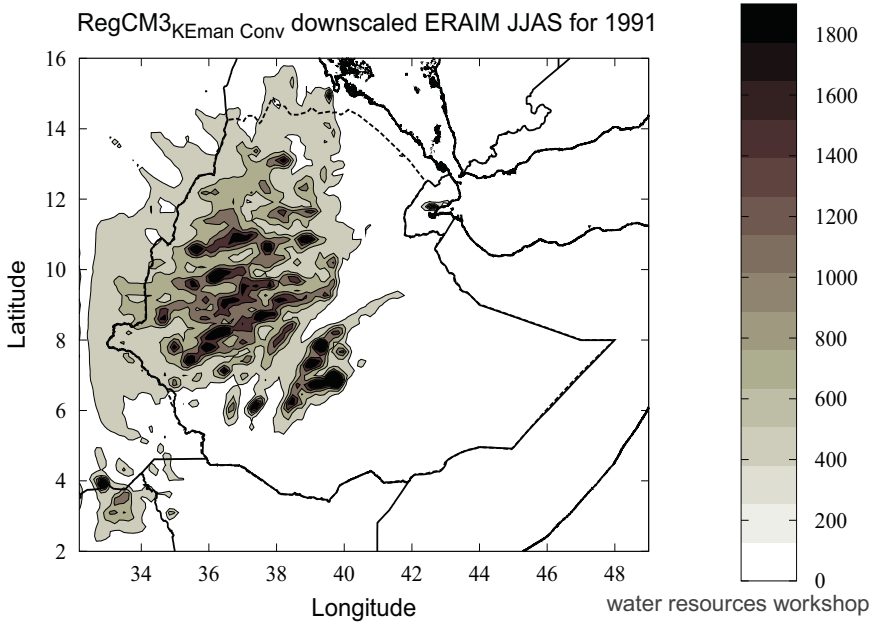
- Summary and future work
  - Spatial variability is reasonably reproduced
  - It seems that the model captures ENSO related rainfall variability
  - It needs to be downscaled if we have to use it as an input for impact models
  - The skill between statistical and dynamically downscaled rainfall should be compared

Future work: dynamical downscaling



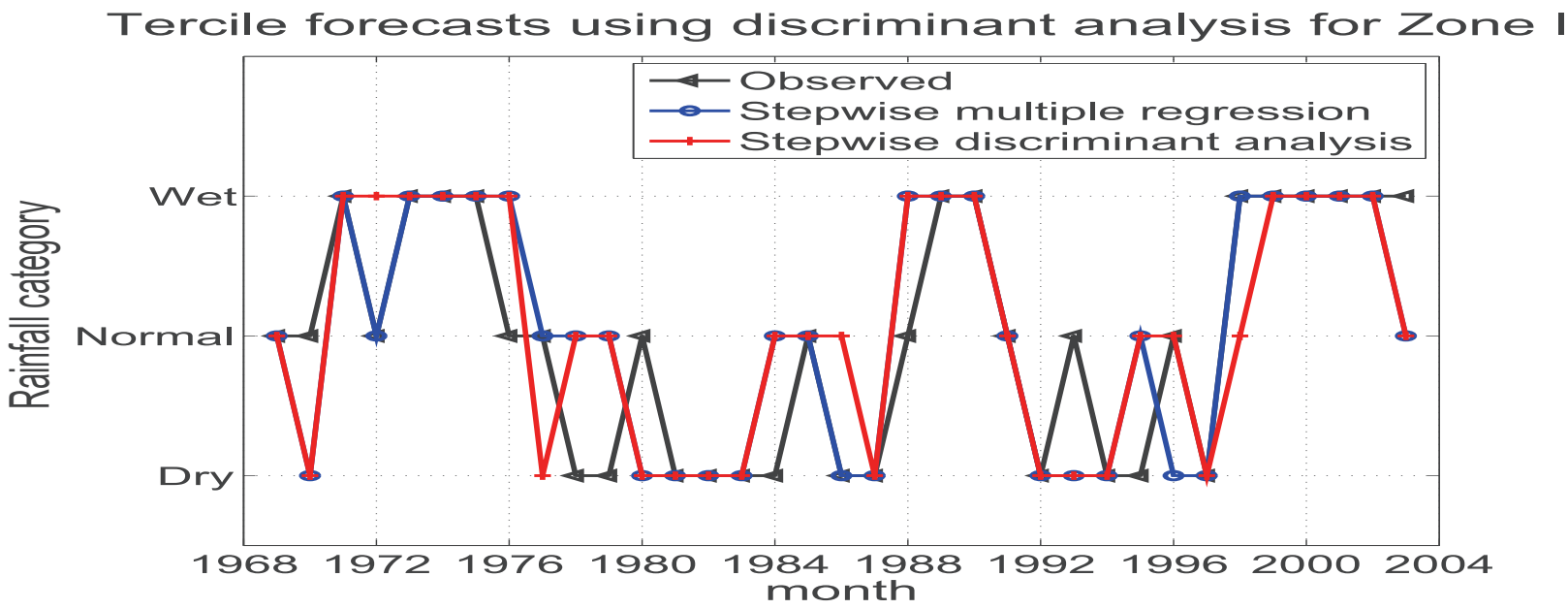






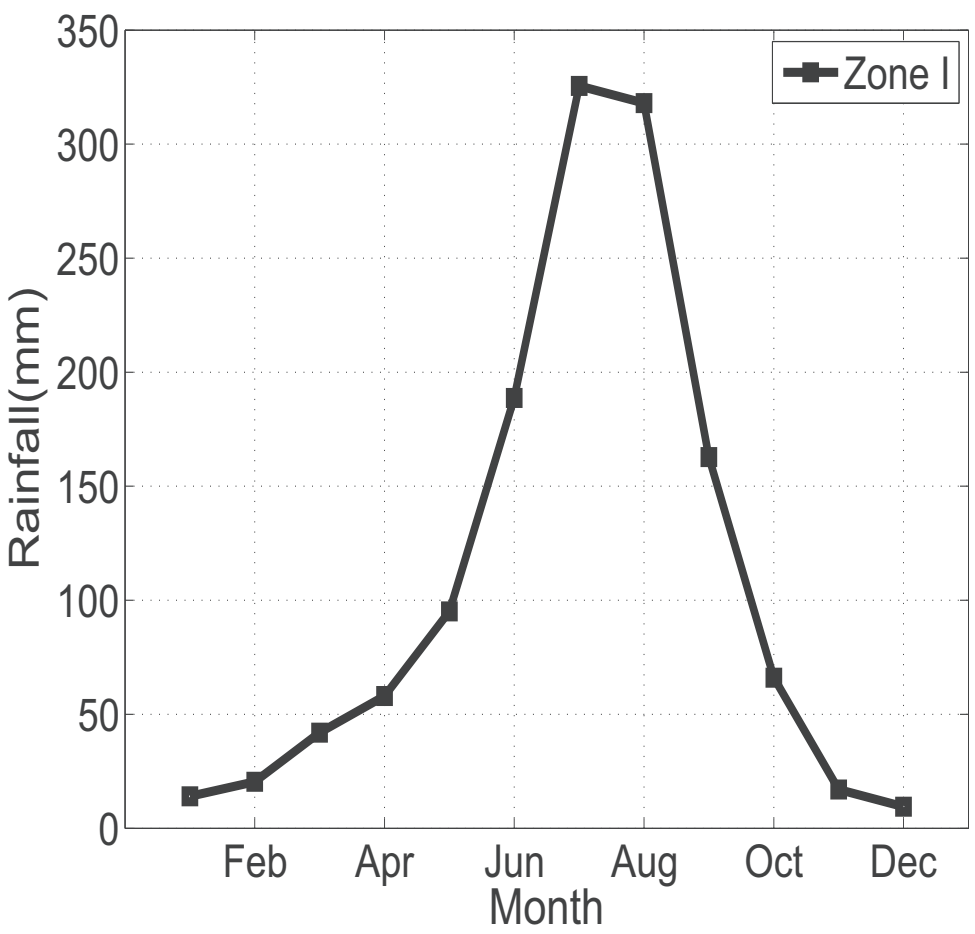
Thanks

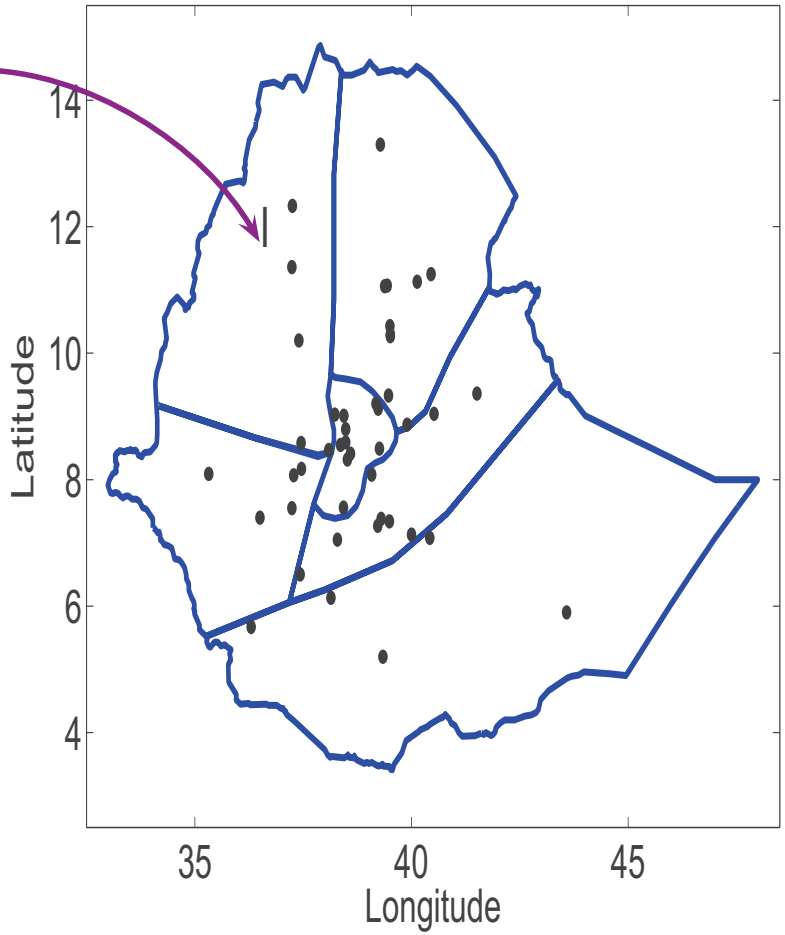
hindcast


Tercile forecasts using discriminant analysis for Zone I
Observed
Stepwise multiple regression
Stepwise discriminant analysis
Rainfall category
Wet
Normal
Dry
1968
1972
1976
1980
1984
1988
1992
1996
2000
2004
month

Identifying homogeneous rainfall zones

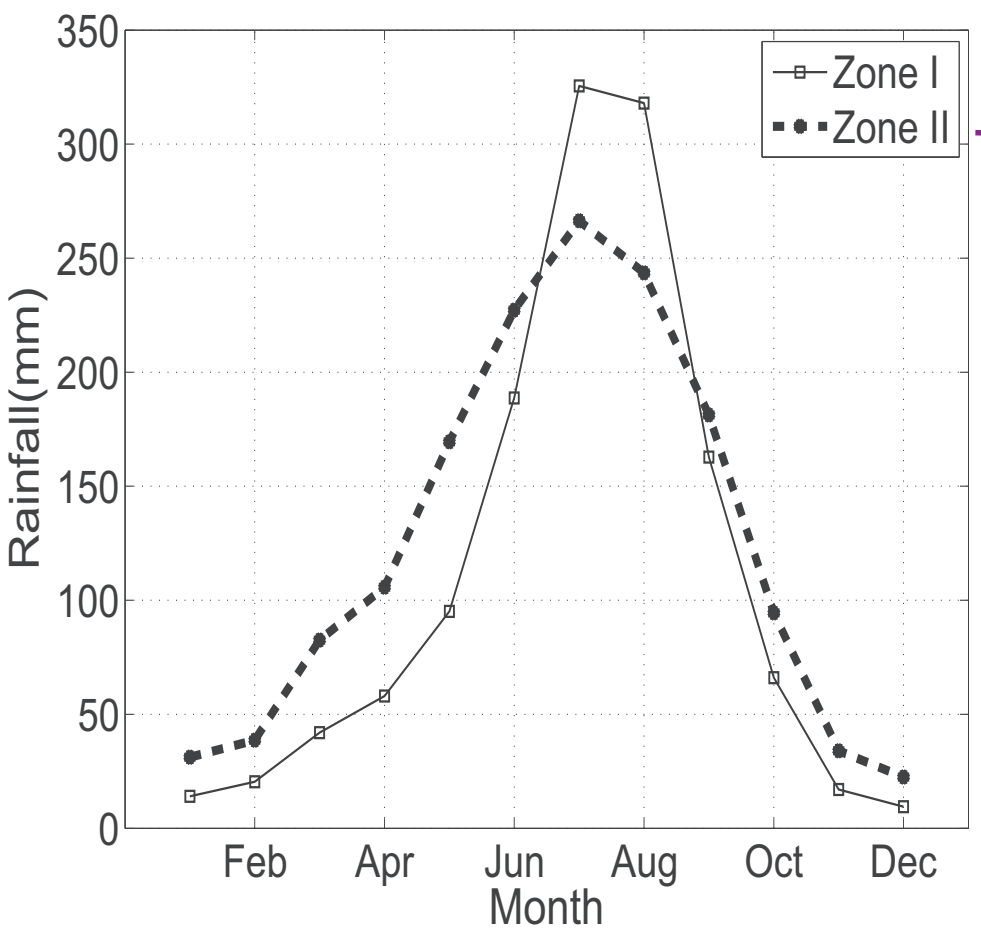
Classification based on seasonal cycle

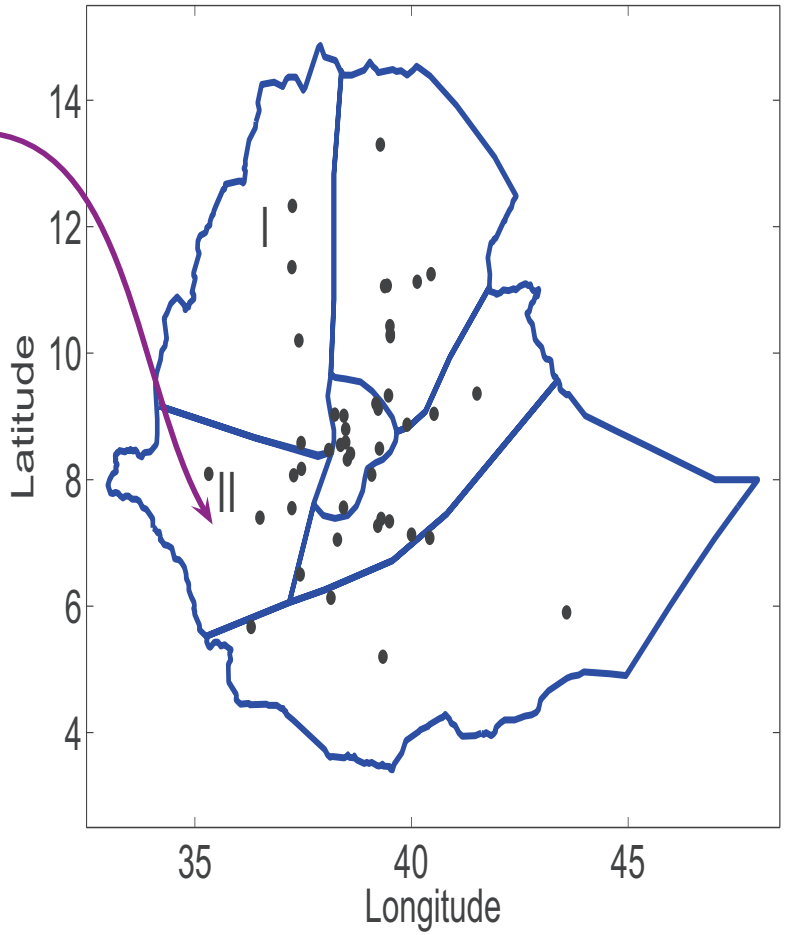




Identifying homogeneous rainfall zones

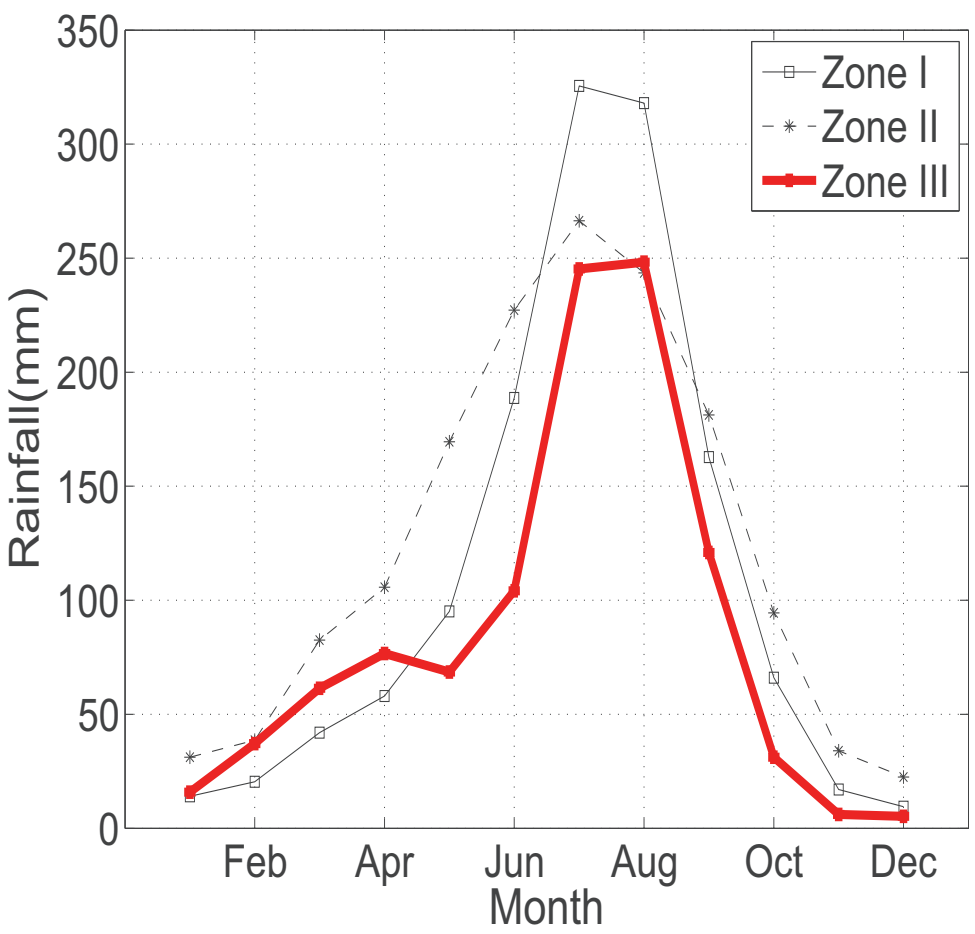
Classification based on seasonal cycle

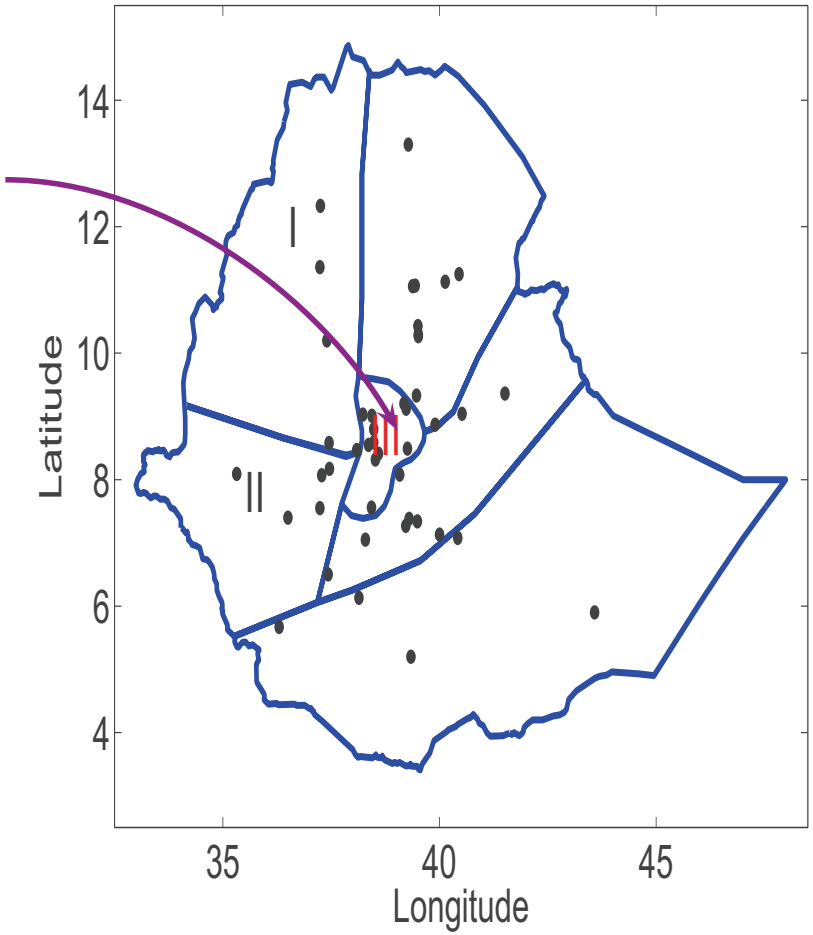




Identifying homogeneous rainfall zones

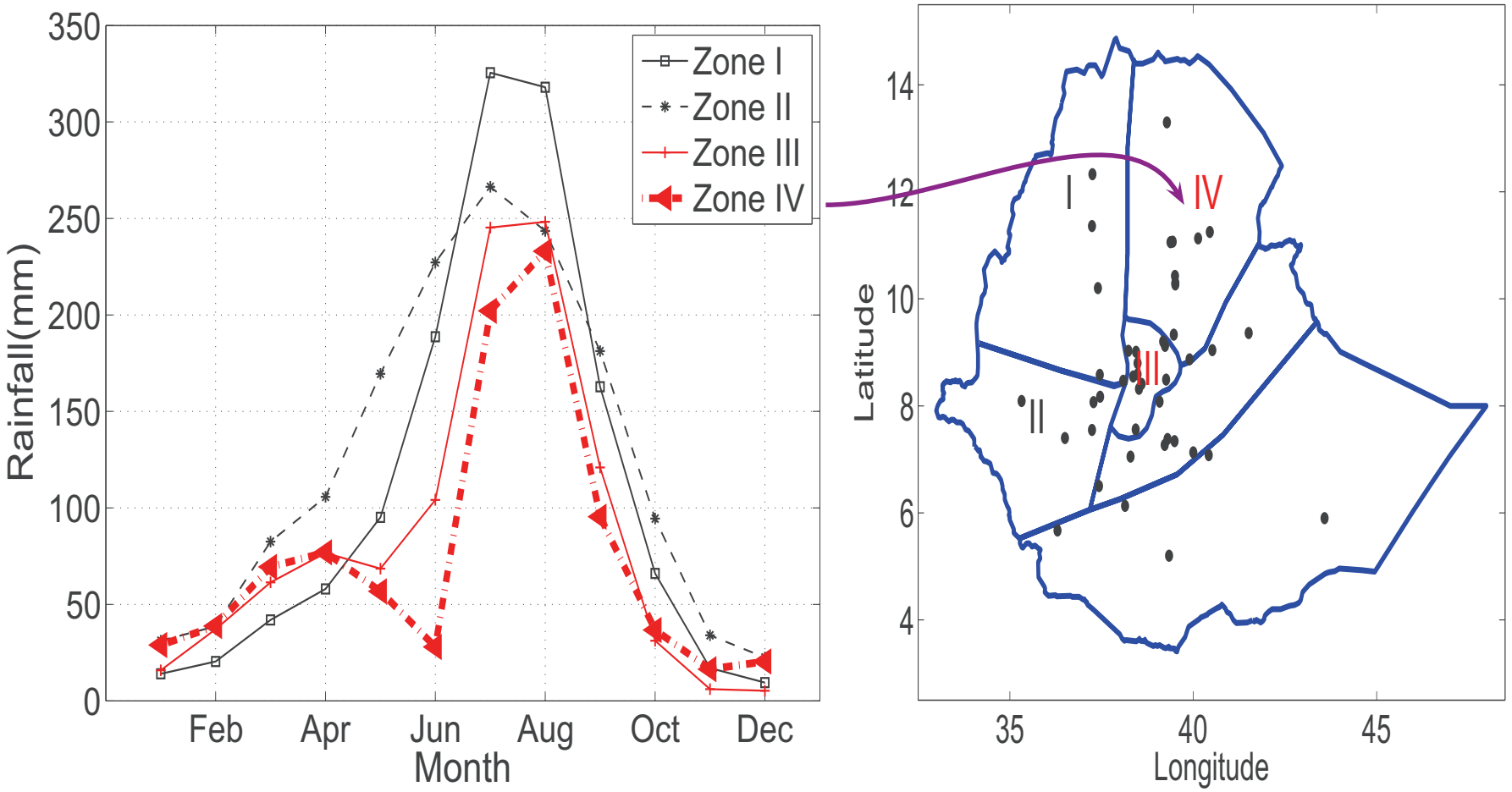
Classification based on seasonal cycle





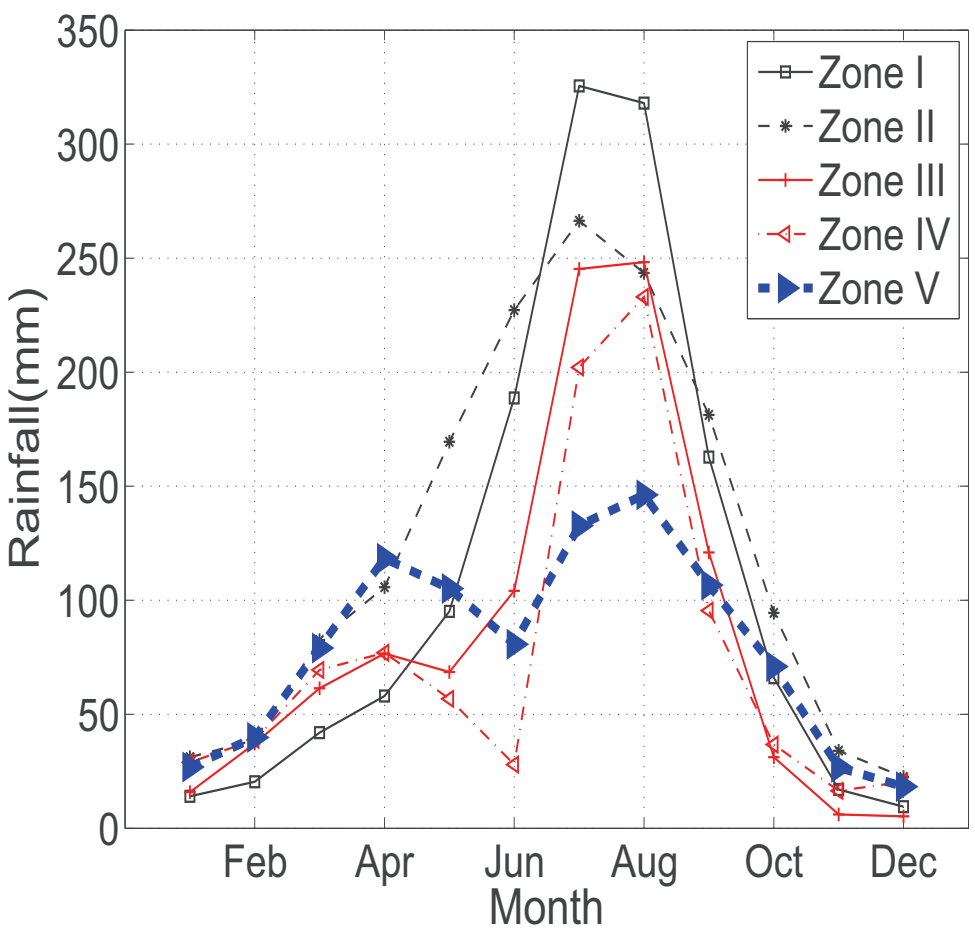
Identifying homogeneous rainfall zones

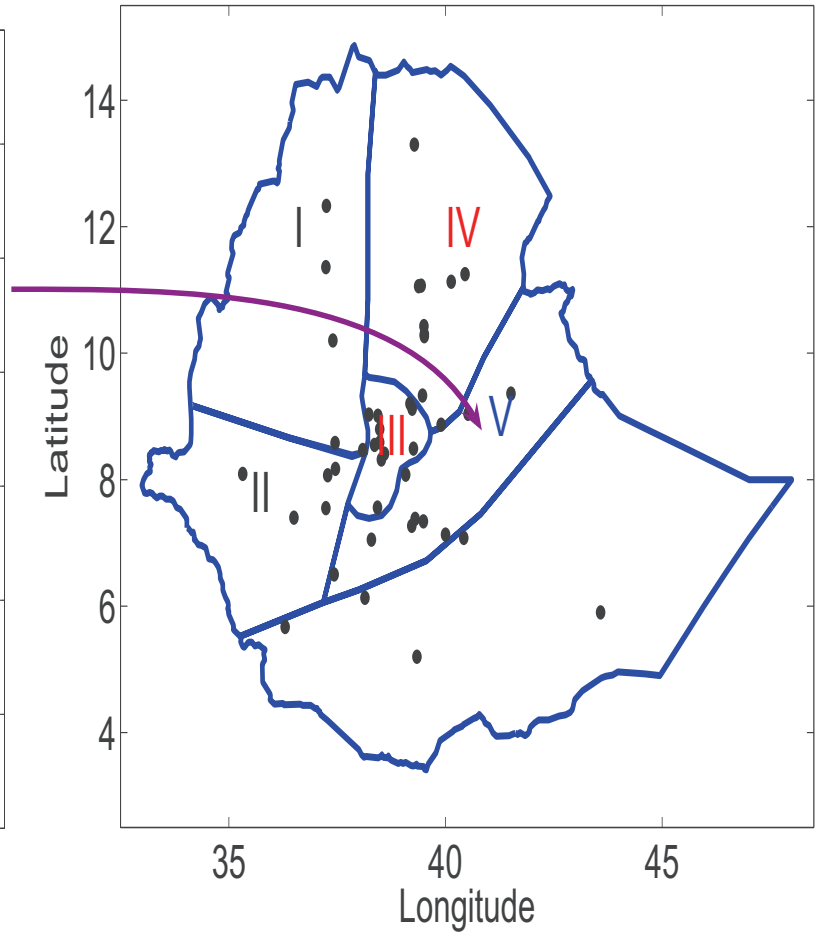
Classification based on seasonal cycle



Identifying homogeneous rainfall zones

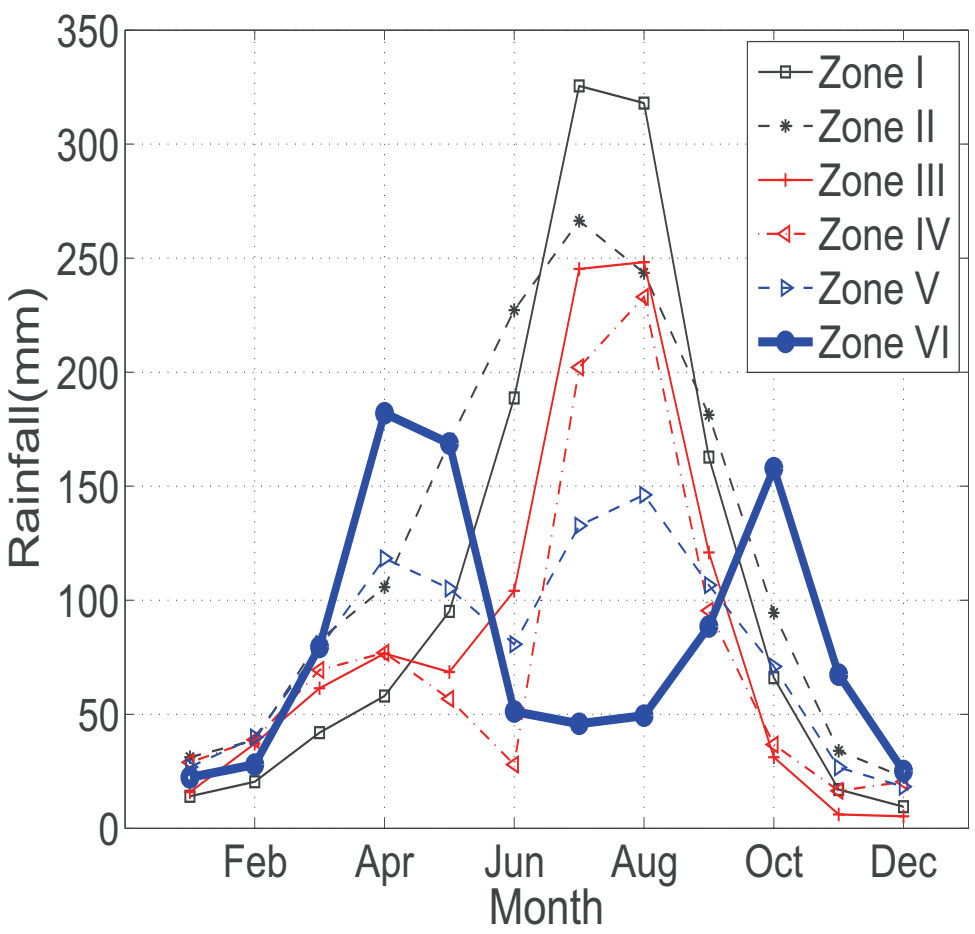
Classification based on seasonal cycle

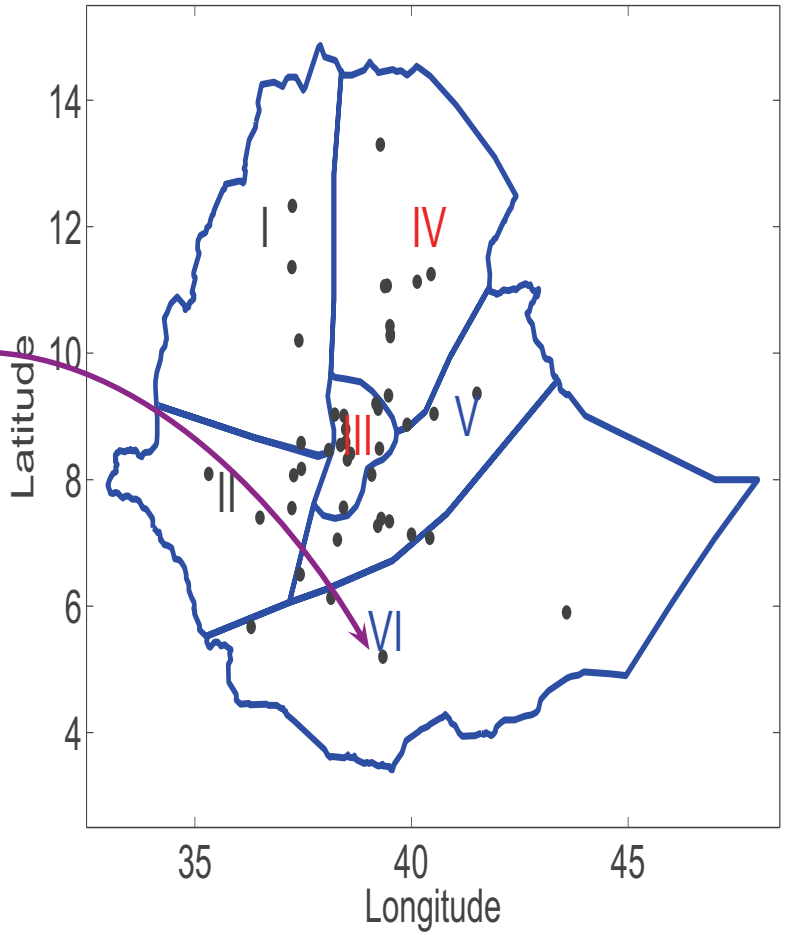




Identifying homogeneous rainfall zones

Classification based on seasonal cycle

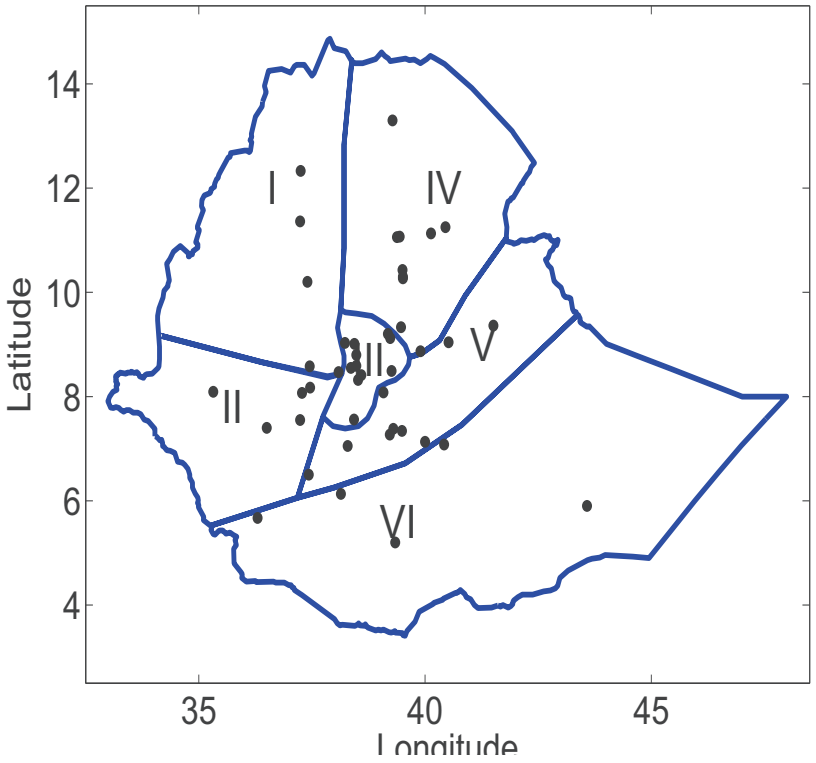




Identifying homogeneous rainfall zones

Classification by Interannual variability

JJAS	Z-I	Z-II	Z-III	Z-IV	Z-V	Z-VI
Z-I	0.03	0.08	0.17	0.12	0.08	-0.09
Z-II	0.08	0.13	0.14	0.10	0.18	0.07
Z-III	0.17	0.14	0.27	0.23	0.21	0.02
Z-IV	0.12	0.10	0.23	0.51	0.24	0.00
Z-V	0.08	0.18	0.21	0.24	0.20	0.12
Z-VI	-0.09	0.07	0.02	0.00	0.12	0.15



Identifying homogeneous rainfall zones

Classification by Interannual variability

JJAS	Z-I	Z-II	Z-III	Z-IV	Z-V	Z-VI
Z-I	0.03	0.08	0.17	0.12	0.08	-0.09
Z-II	0.08	0.13	0.14	0.10	0.18	0.07
Z-III	0.17	0.14	0.27	0.23	0.21	0.02
Z-IV	0.12	0.10	0.23	0.51	0.24	0.00
Z-V	0.08	0.18	0.21	0.24	0.20	0.12
Z-VI	-0.09	0.07	0.02	0.00	0.12	0.15

JJAS	Z-I	Z-IIa	Z-IIb	Z-III	Z-IV	Z-V
Z-I	0.26	0.16	0.02	0.16	0.16	-0.06
Z-IIa	0.16	0.24	0.15	0.15	0.14	0.05
Z-IIb	0.02	0.15	0.33	0.16	0.13	0.08
Z-III	0.16	0.15	0.16	0.25	0.17	-0.08
Z-IV	0.16	0.14	0.13	0.17	0.31	-0.06
Z-V	-0.06	0.05	0.08	-0.08	-0.06	0.14