



# Identifying and attributing common data quality problems: temperature and precipitation observations in Bolivia and Peru

S. Hunziker<sup>1</sup>, S. Gubler<sup>2</sup>, J. Calle<sup>3</sup>, I. Moreno<sup>3</sup>, M. Andrade<sup>3</sup>, F. Velarde<sup>3</sup>, L. Ticona<sup>3</sup>, G. Carrasco<sup>4</sup>, Y. Castellón<sup>4</sup>, C. Oria Rojas<sup>5</sup>, S. Brönnimann<sup>1</sup>, M. Croci-Maspoli<sup>2</sup>, T. Konzelmann<sup>2</sup>, M. Rohrer<sup>2</sup>

<sup>1</sup> Oeschger Centre for Climate Change Research and Institute of Geography, University of Bern, Bern, Switzerland

<sup>2</sup> Federal Office of Meteorology and Climatology MeteoSwiss, Zürich, Switzerland

<sup>3</sup> Laboratorio de Física de la Atmósfera, Instituto de Investigaciones Físicas, Universidad Mayor de San Andrés, La Paz, Bolivia

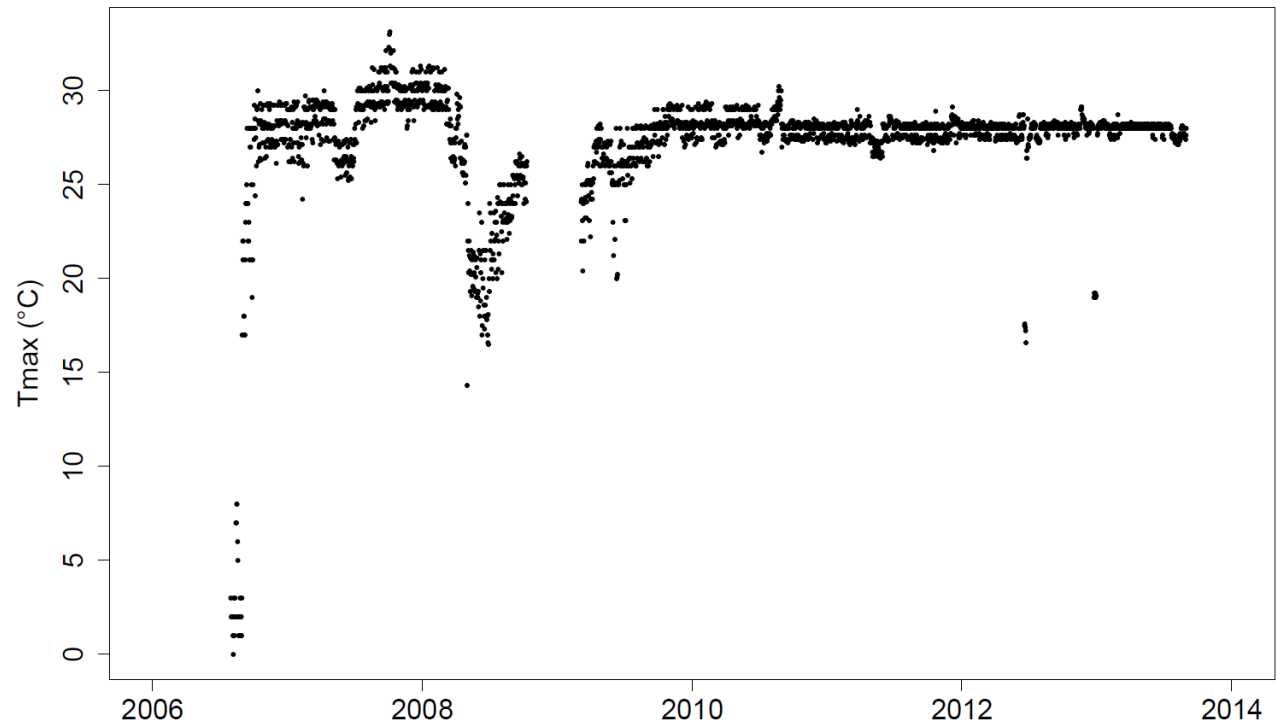
<sup>4</sup> Servicio Nacional de Meteorología e Hidrología de Bolivia, SENAMHI

<sup>5</sup> Servicio Nacional de Meteorología e Hidrología del Perú, SENAMHI



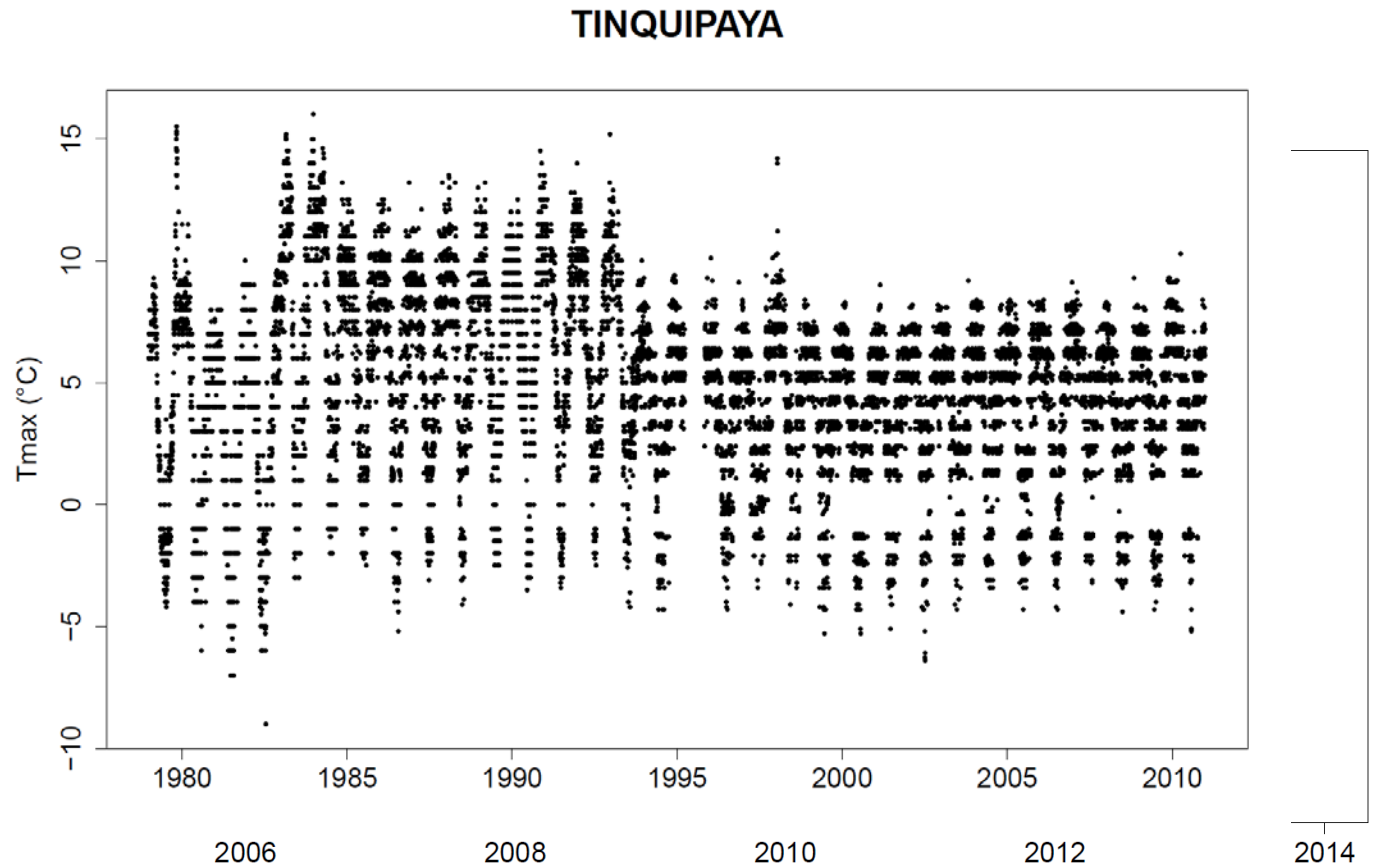
- Data availability
- Sparse station network
- Metadata is frag
- Often severe da

## INCAPAMPA



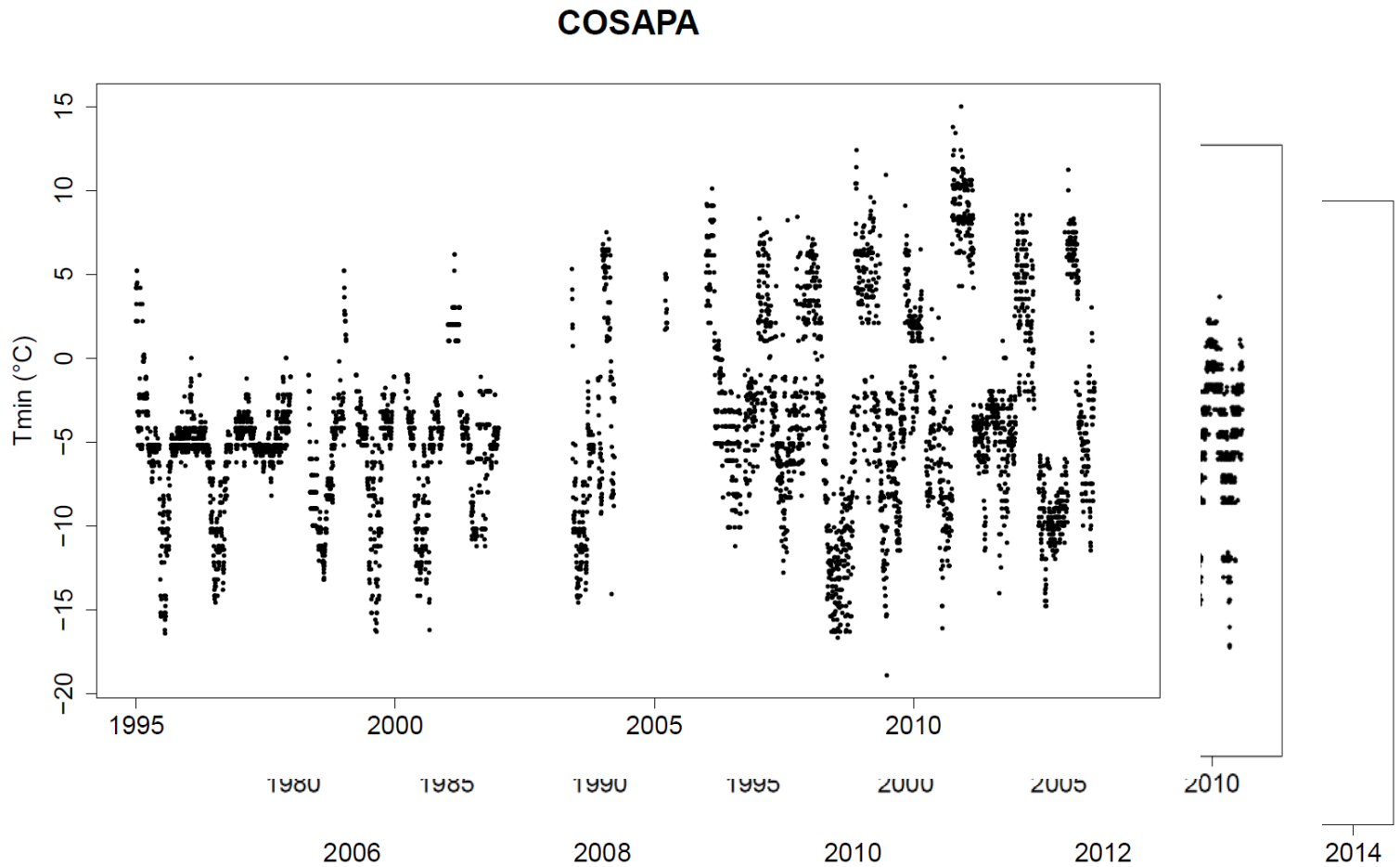
# Challenges

- Data availability
- Sparse stati
- Metadata is
- Often sever



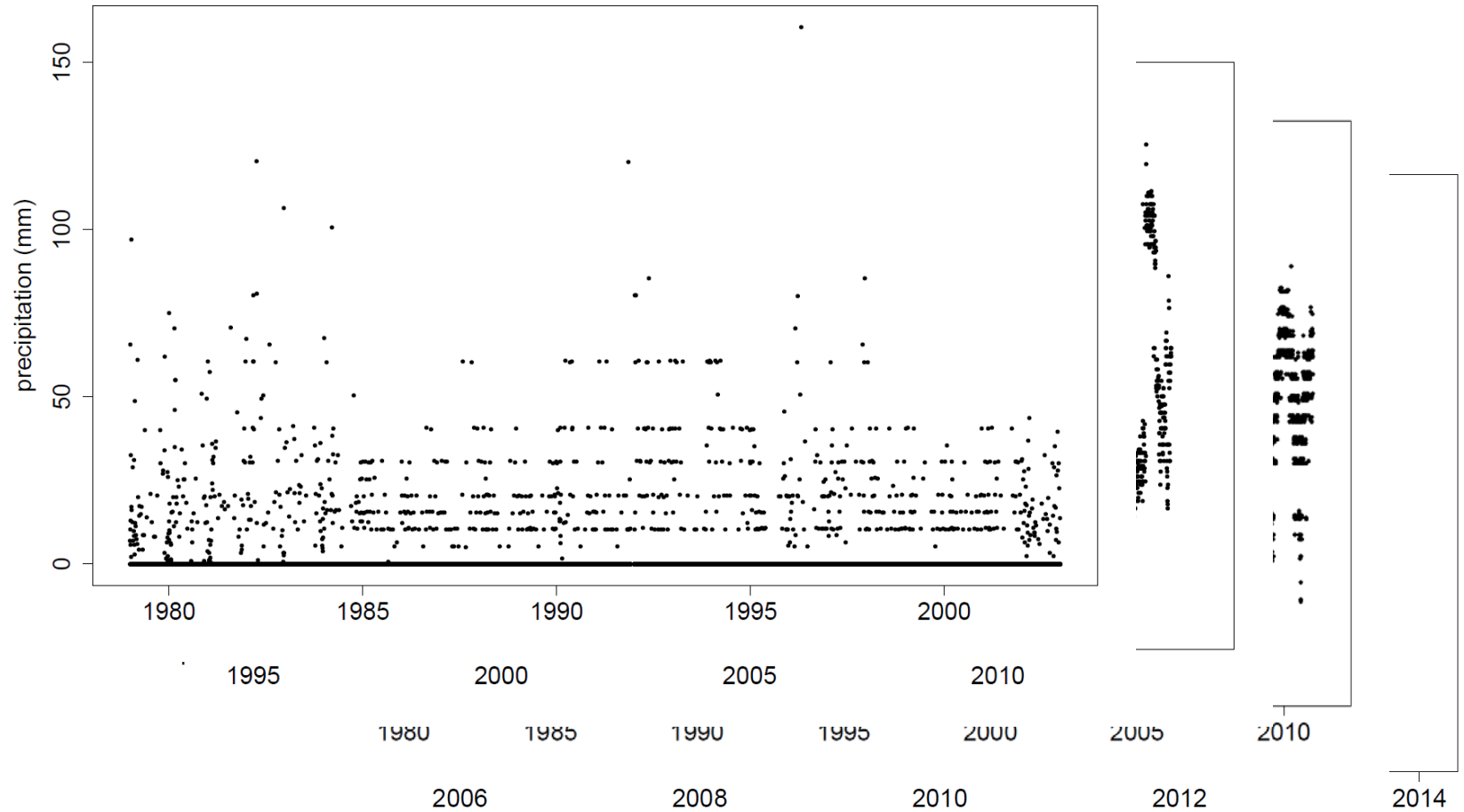
# Challenges

- Data availability
- Spars
- Metac
- Often



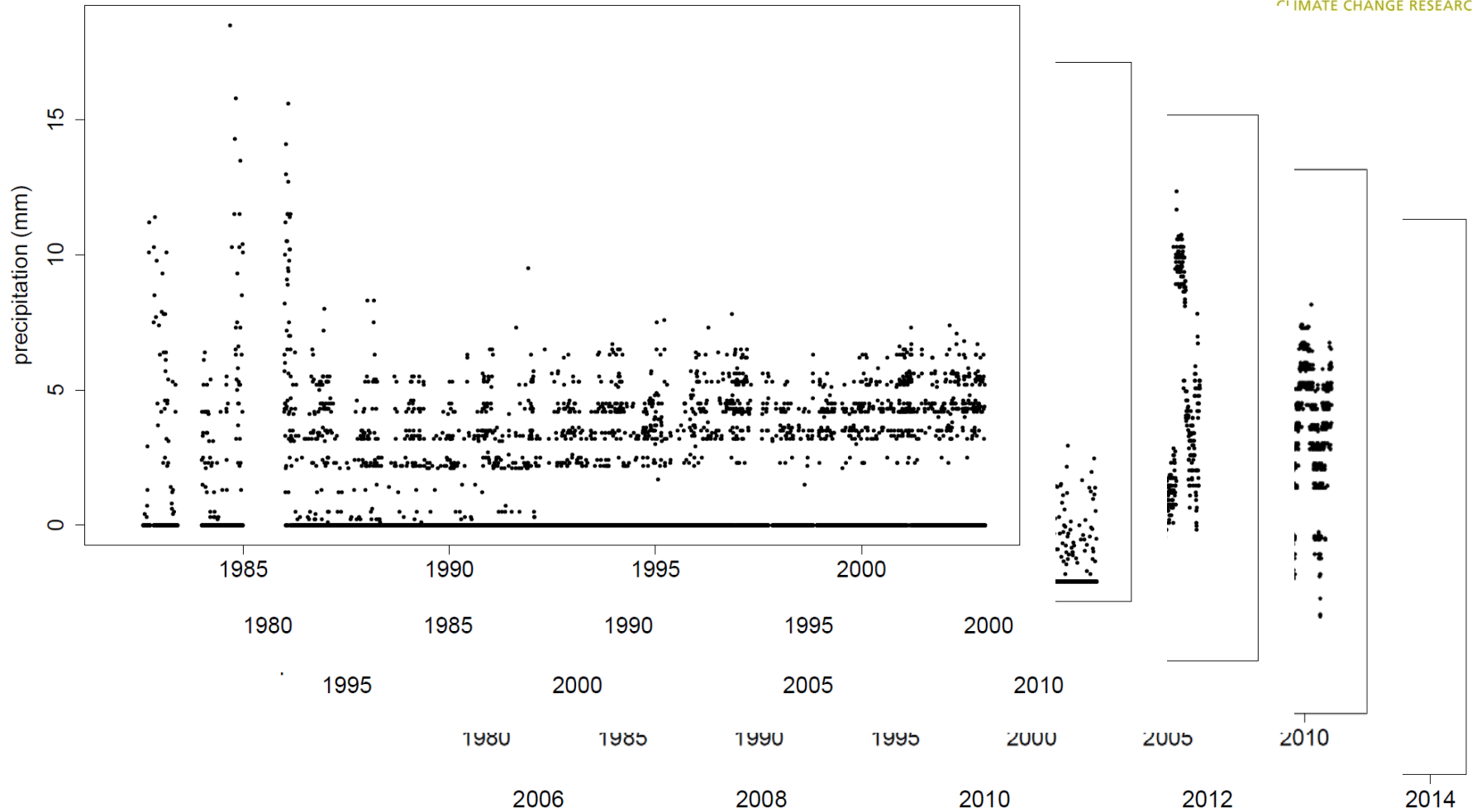
# Challenges

## QUIMOME



# Challenges

## VILOCO



- Data availability
- Sparse station network
- Metadata is fractional or missing
- Often severe data quality problems

## Frequently found errors:

- missing temperature intervals
- reduction of variability
- rounding inconsistencies
- 20mm precipitation cut
- missing low precipitation values
- untagged rainfall accumulations
- transcription errors



- Data availability
- Sparse station network
- Metadata is fractional or missing
- Often severe data quality problems

Frequently found errors:

- missing temperature intervals
- **reduction of variability**
- **rounding inconstancies**
- **20mm precipitation cut**
- **missing low precipitation values**
- **untagged rainfall accumulations**
- transcription errors

Station visits to






1. report and assess the actual state of the station
2. reconstruct the station history
3. detect sources of data errors

# Metadata

$u^b$



# Metadatos

FORMULARIO DE RECOLECCIÓN DE METADATOS ESTACIONES METEOROLÓGICAS			
Nombre de la estación	Cota Cota	Alias	
Código(s)		Tipo de estación	Automática (datalogger CR200X SN: 13809) Campbell Scientific.
Institución a cargo	Instituto de Investigaciones Físicas		
Persona de contacto	Lic. René Torrez Santalla		
Rotación del personal	Actual	Lic. René Torrez Santalla	
	Previa	No hubo nunca rotación del personal	
Cada cuanto se reporta a la oficina	Cada 2 semanas se descargan los datos		
Ubicación	Actual	Previa	Inicio de la serie de datos
Latitud	16°33'20.893"		Se inicia en tre 1998 y 1999
Longitud	68°03'59.297"		
Elevación	3445.73 m		
Fotografías desde la estación			
Vista al Sur	Vista al Norte	Foto de la estación	
			
Vista al Este	Vista al Oeste		
			
Obstáculos	Existen materiales como maderas y fierros que perjudican el tránsito		
Infraestructura	La estación está instalada en una Torre de hierro de aproximadamente 3.33 m de altura.		
Observaciones	El techo podría perjudicar la medición de la radiación solar.		

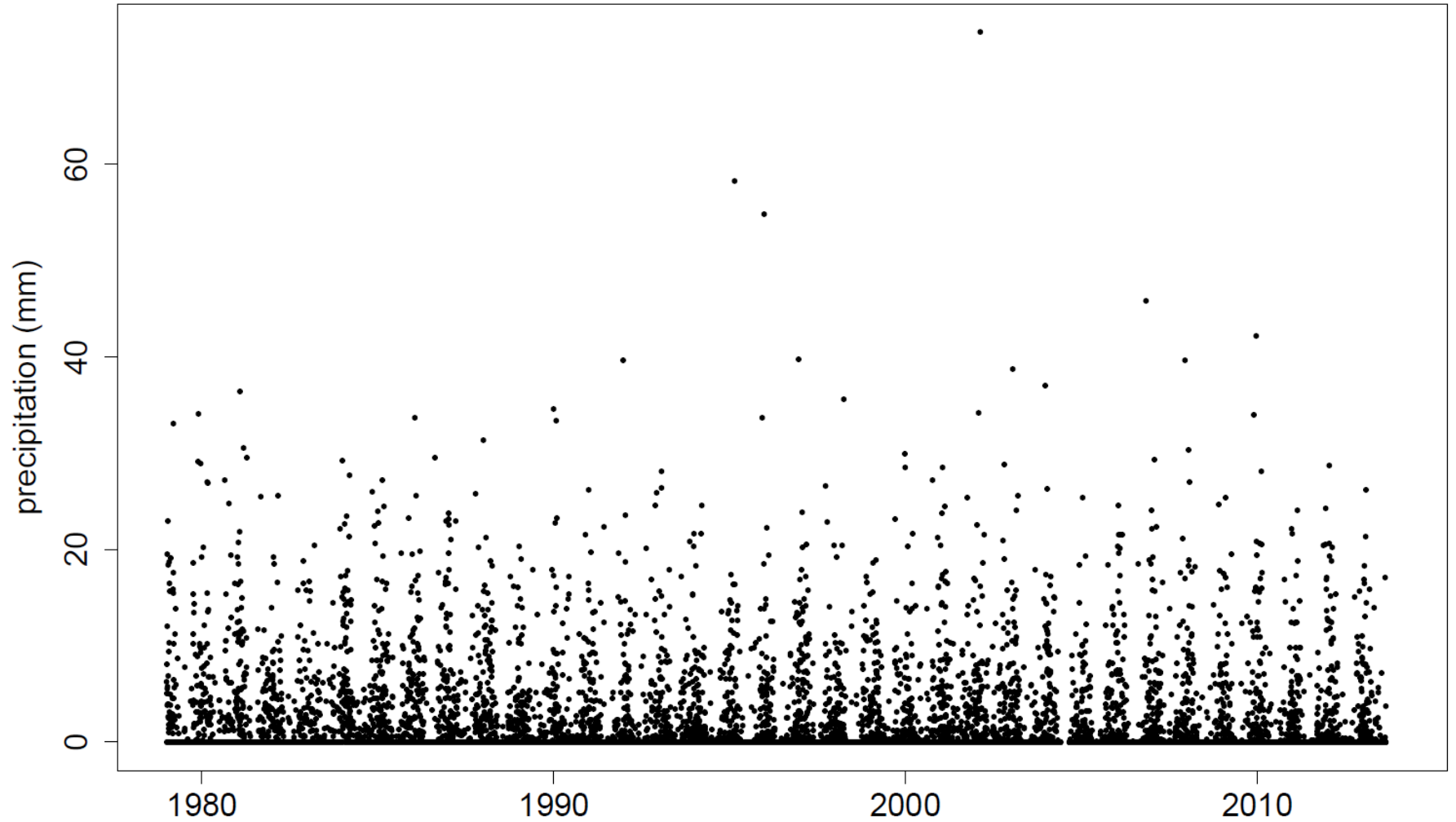
Variables que se miden	Unidades	Instrumento Marca/Modelo	Altura	Intervalo de medición	Calibración	Observaciones
Temperatura	°C	Licor 200	3.33[m]	5 [min]	No se realiza	No se realiza calibraciones, solo intercomparaciones con la estación de Patacamaya Variación de 2°C actualmente . 13/03/2014
Presión						
Humedad relativa	%	Met one instruments	3.33[m]	5 [min]	No se realiza	No se realiza calibraciones, solo intercomparaciones con la estación de Patacamaya
Velocidad del viento	m/s	Met one instruments	3.07[m]	5 [min]	No se realiza	No se realiza calibraciones, solo intercomparaciones con la estación de Patacamaya
Dirección del viento	(°)	Met one instruments	3.07[m]	5 [min]	No se realiza	No se realiza calibraciones, solo intercomparaciones con la estación de Patacamaya
Precipitación						
Radiación solar	mV		3.01[m]	5 [min]	No se realiza	No se realiza calibraciones, solo intercomparaciones con la estación de Patacamaya

**\*Notas:**

- No se tiene un periodo determinado para realizar la intercomparación
- La estación meteorológica fue armada por partes y no se tiene el número del modelo de cada parte.

BITACORA DE LA ESTACIÓN			
#	Fecha	Participante(s)	Descripción
1	12/03/2014	Decker Guzman Zabalaga	Recojo de bitacora No existe bitacora de la estación.
2	No existe fecha exacta	Lic. Rene Torrez Santalla	Retiro de radiómetro Aproximadamente hace un año y medio se hizo el retiro de radiómetro .
3	20/02/2014	Lic. Rene Torrez Santalla	Descarga de datos Se descargaron los datos de la estación por medio de una computadora conectada al Datalogger

## VINO\_TINTO



# Metadata

*u<sup>b</sup>*

---

<sup>b</sup>  
**UNIVERSITÄT  
BERN**

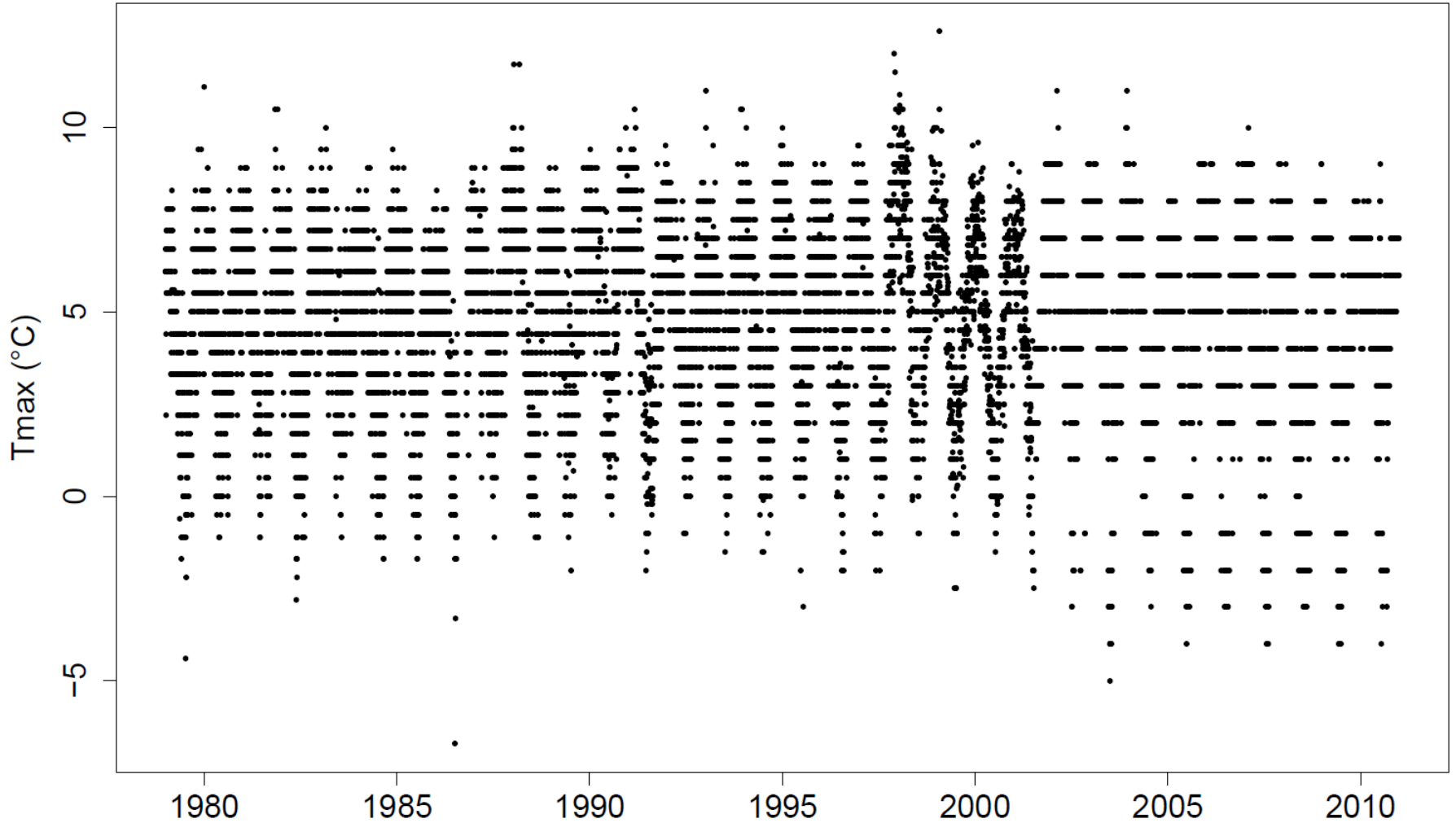
**OESCHGER CENTRE**  
CLIMATE CHANGE RESEARCH



# Rounding inconsistencies

$u^b$

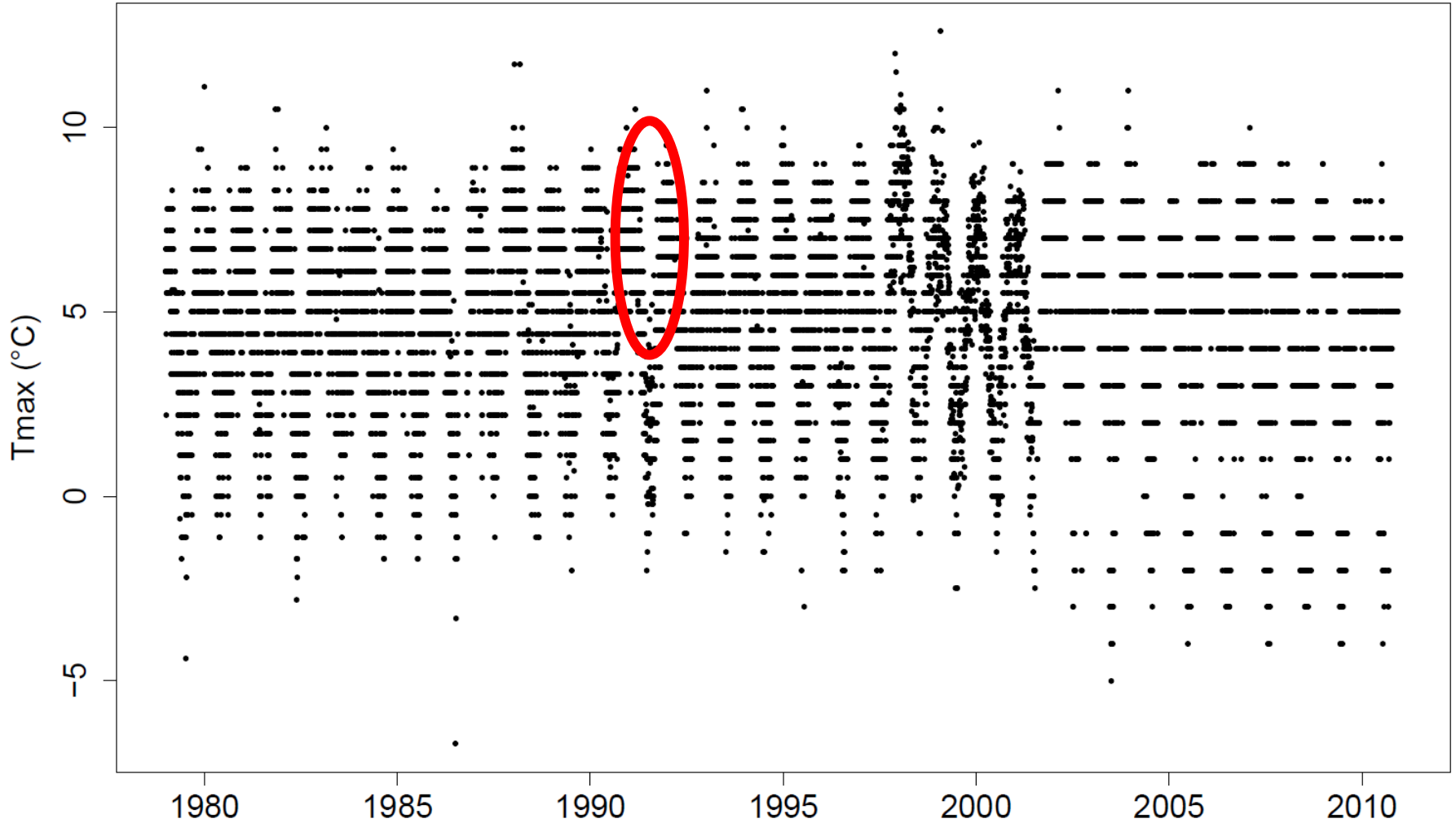
TIRAQUE



# Rounding inconsistencies

$u^b$

TIRAQUE





# Reduced variability in all variables

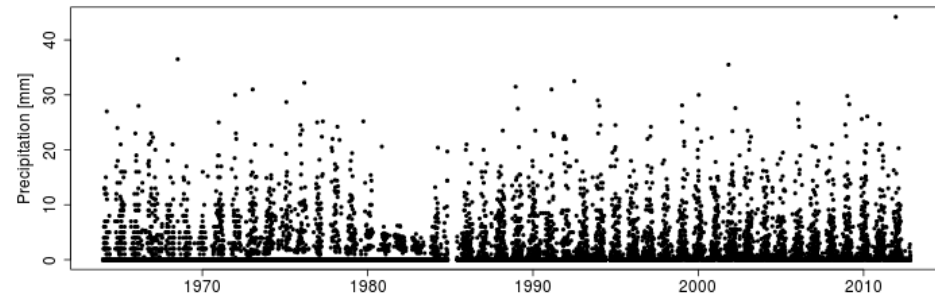
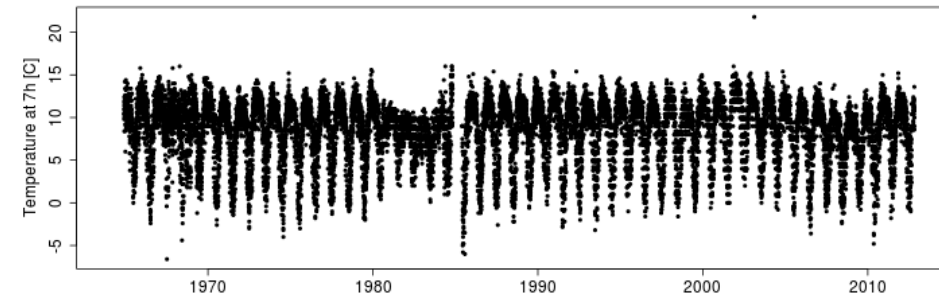
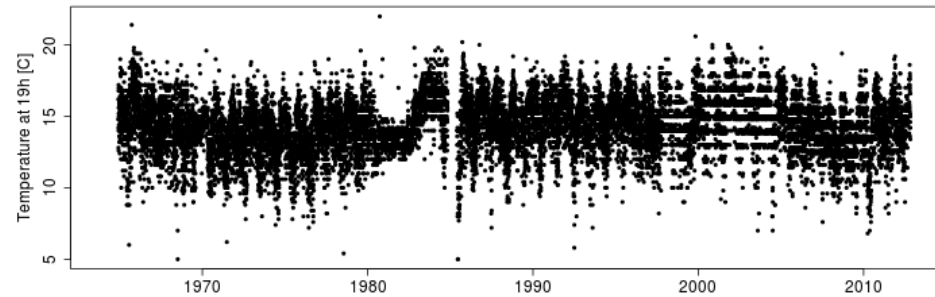
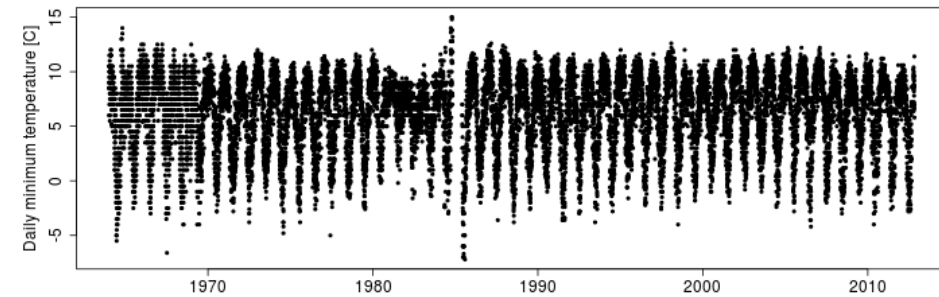
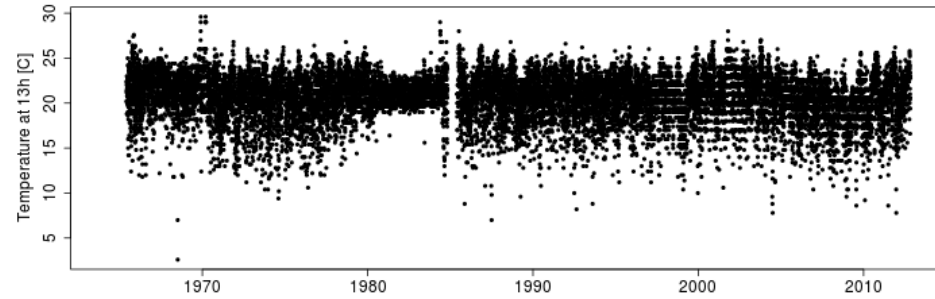
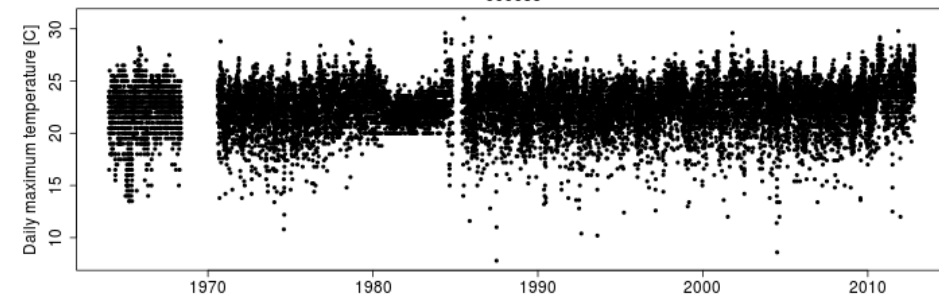
$u^b$

<sup>b</sup>  
UNIVERSITÄT  
BERN

OESCHGER CENTRE  
CLIMATE CHANGE RESEARCH

## Urubamba, Peru

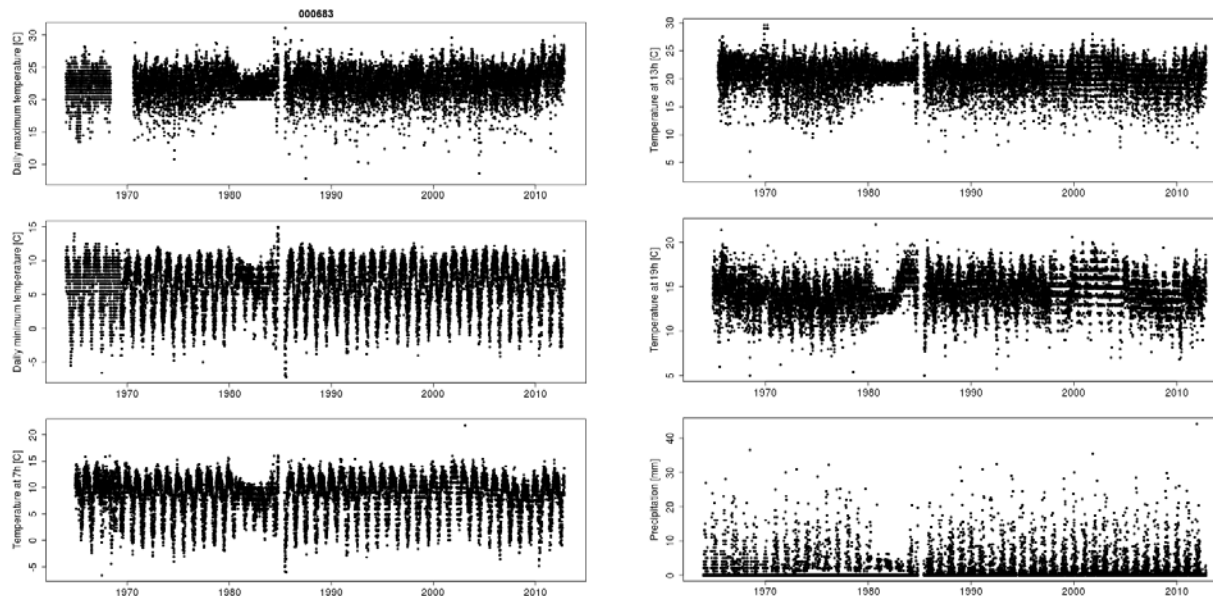
000683



# Reduced variability in all variables

## Urubamba, Peru

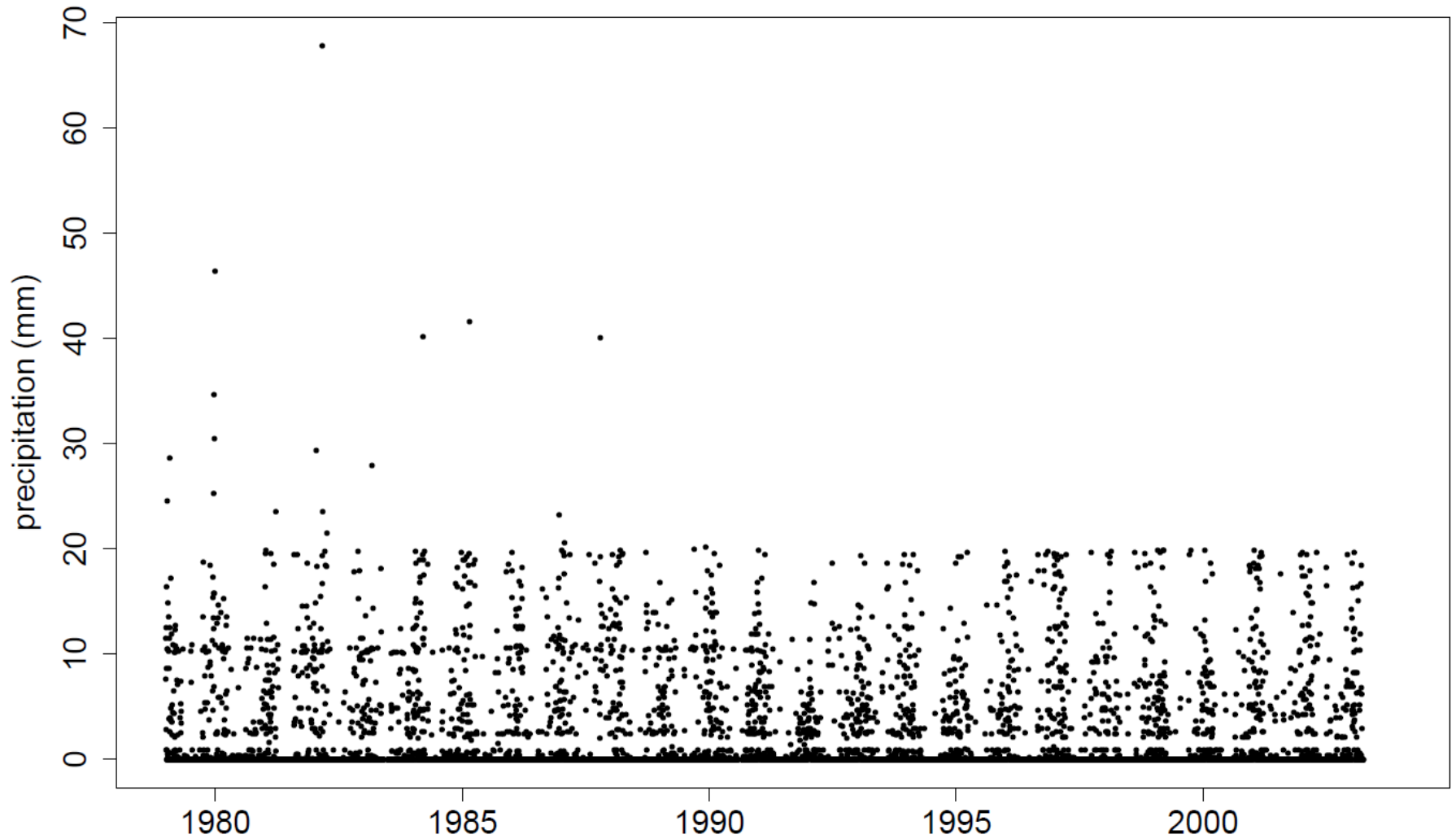
- Rebellion of the Sendero Luminoso against the Peruvian state (escalating in the early 80s)
- Data gaps and errors are found in many stations in that time
- All parameters affected → observer error
- Exact source of error source is unknown



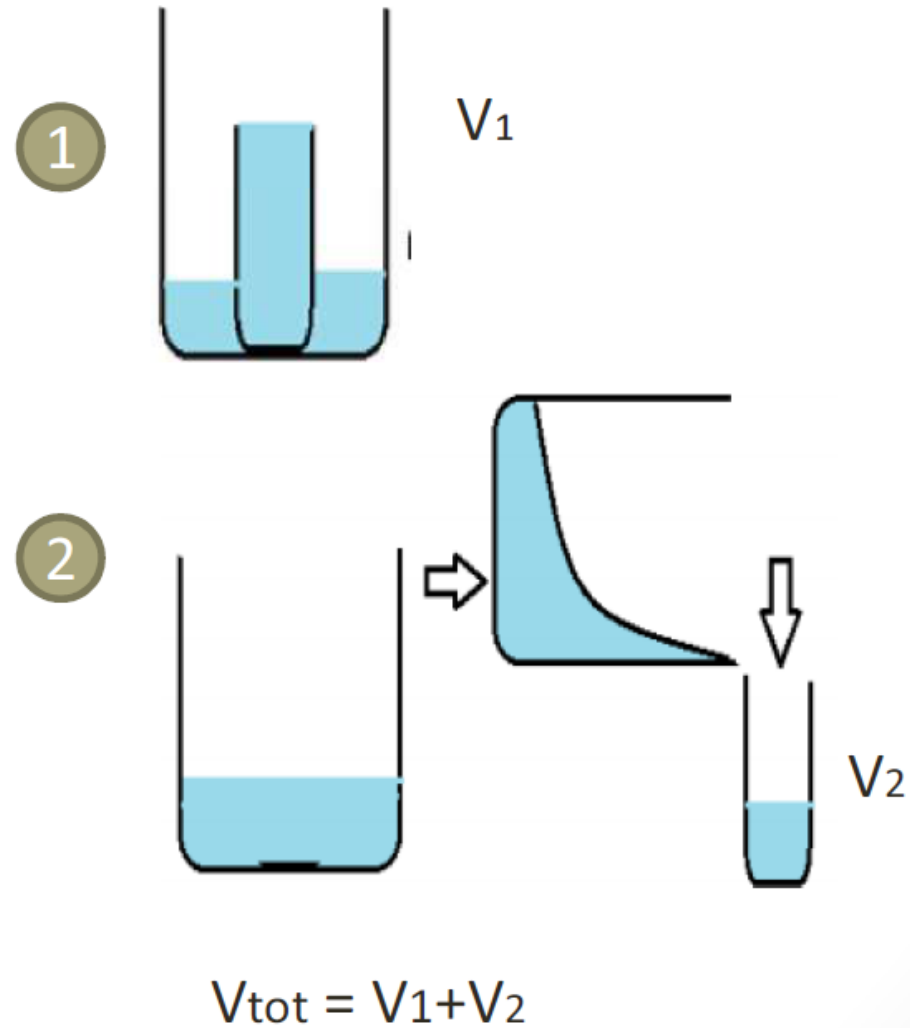
# 20mm precipitation cut

$u^b$

## AGUIRRE



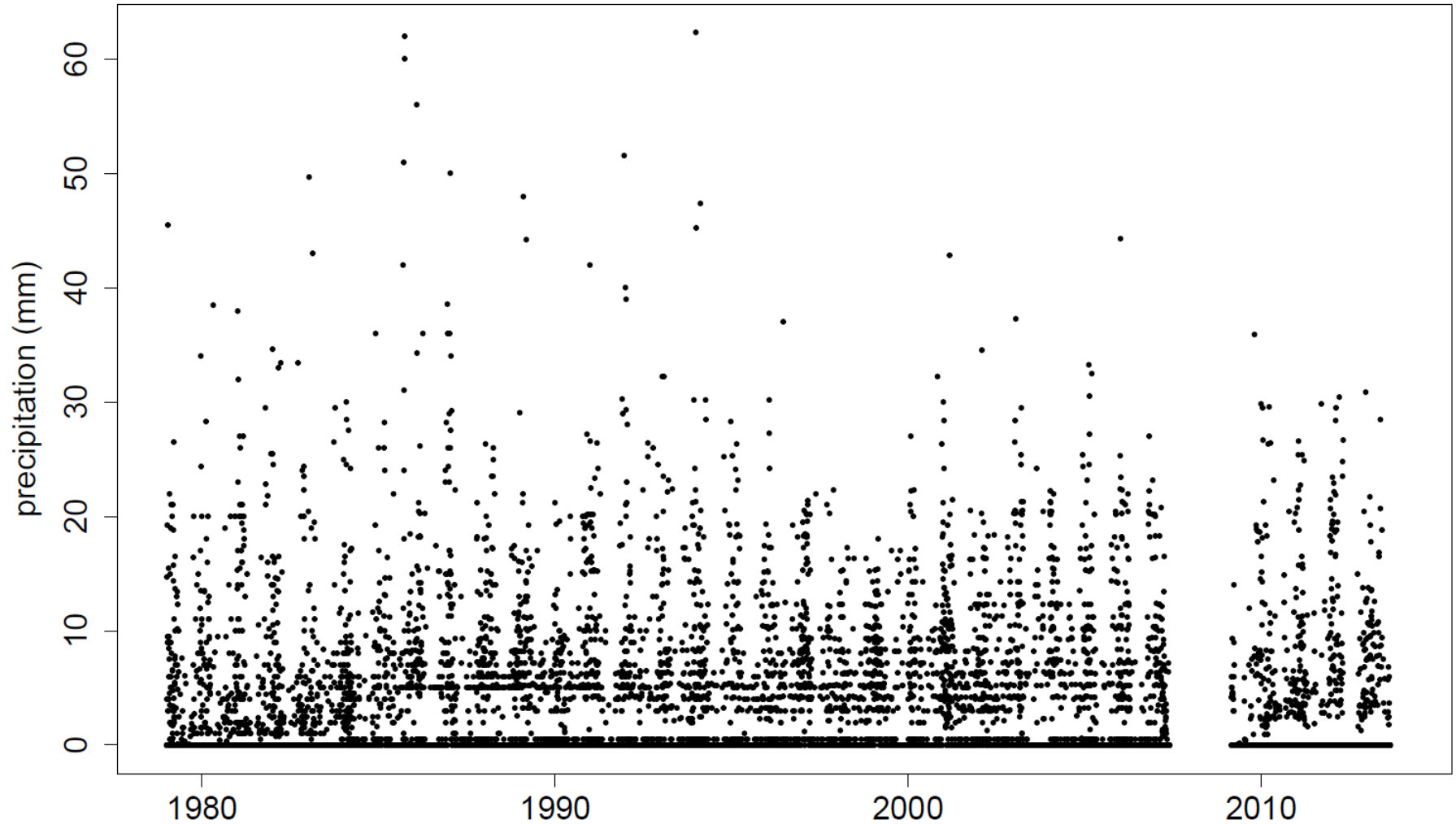
# 20mm precipitation cut



# Low precipitation gap

$u^b$

## QUIABAYA



# Low precipitation gap

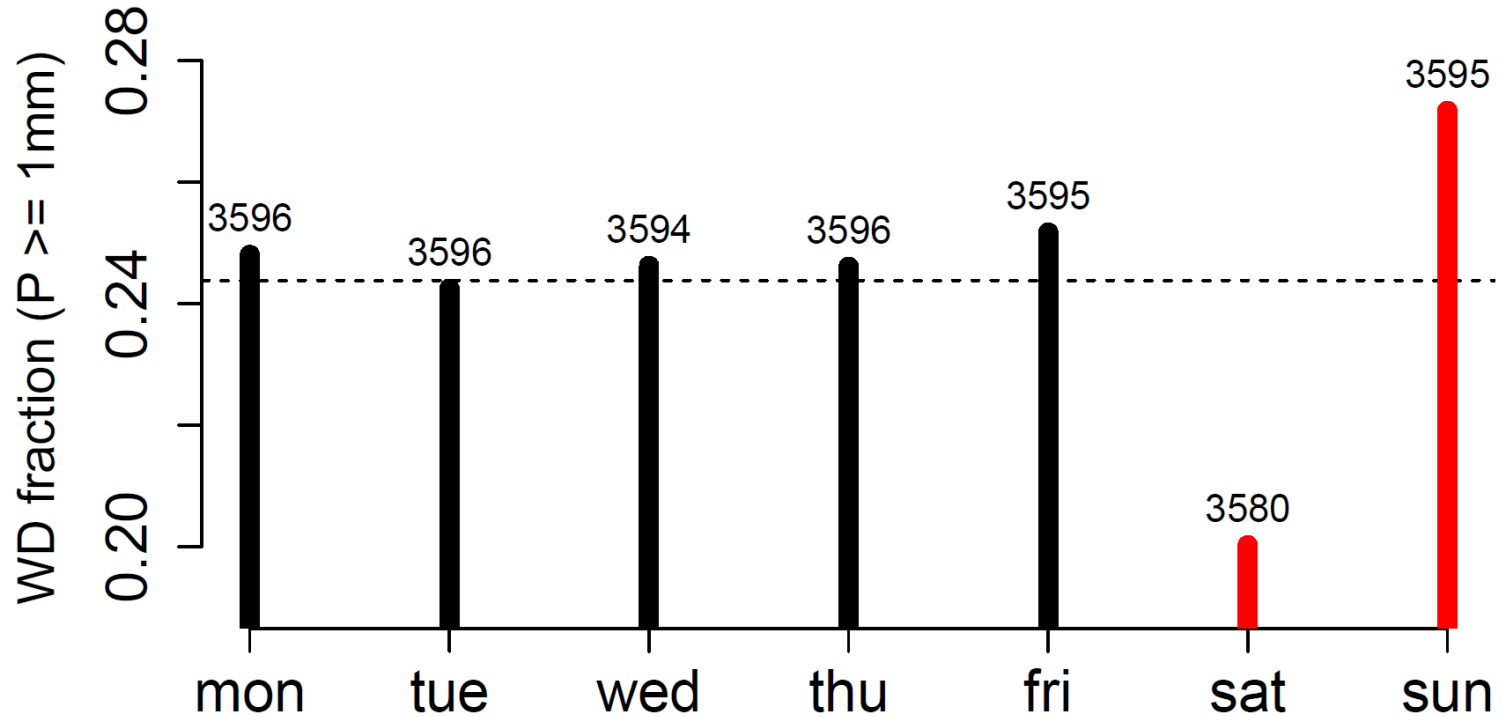
*u<sup>b</sup>*

<sup>b</sup>  
UNIVERSITÄT  
BERN

OESCHGER CENTRE  
CLIMATE CHANGE RESEARCH



## SAN\_CALIXTO



----- expected WD fraction

3596 total number of days

■ WD fraction inside the 95% CI

■ WD fraction outside the 95% CI

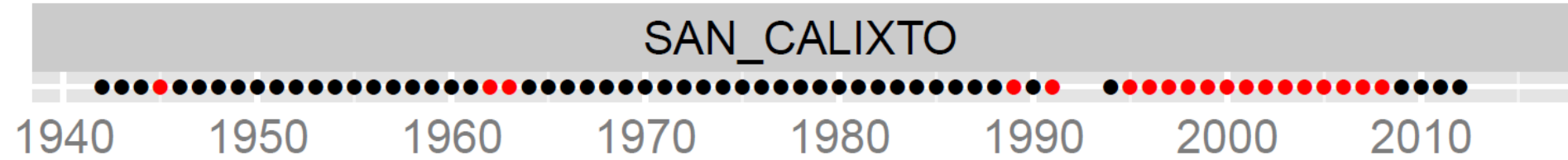
# Untagged rainfall accumulations

*u<sup>b</sup>*

b  
UNIVERSITÄT  
BERN

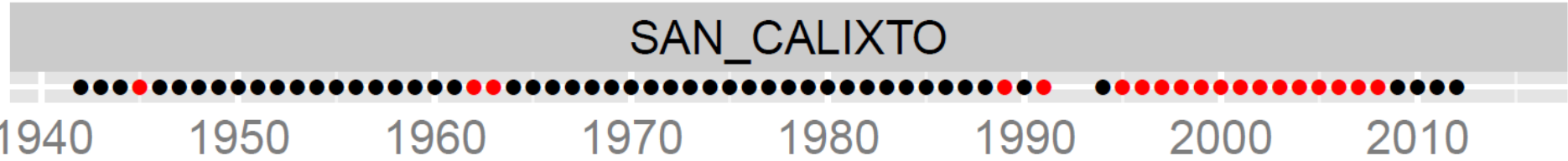
OESCHGER CENTRE  
CLIMATE CHANGE RESEARCH

SAN\_CALIXTO

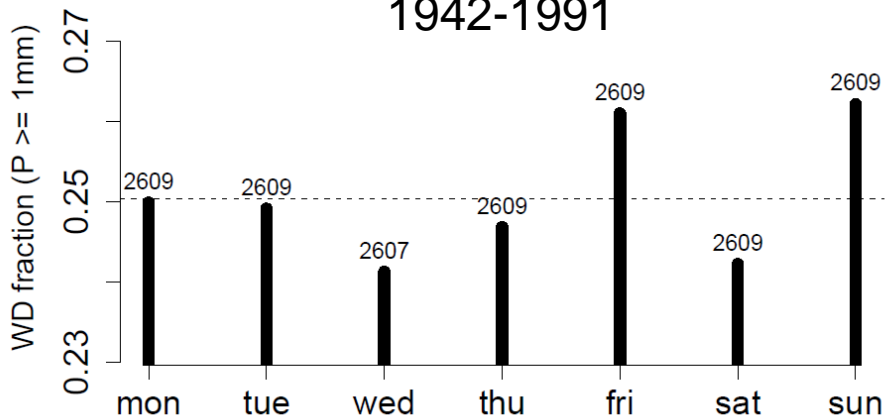




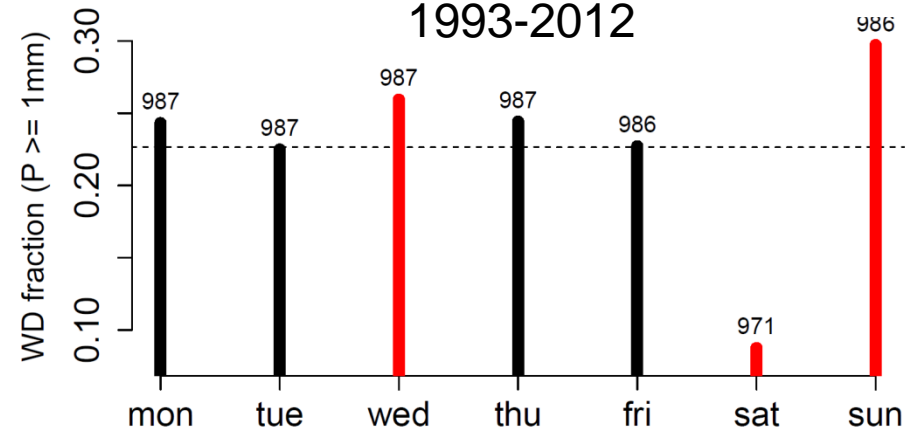
# Untagged rainfall accumulations



**SAN\_CALIXTO**  
1942-1991



**SAN\_CALIXTO**  
1993-2012



----- expected WD fraction

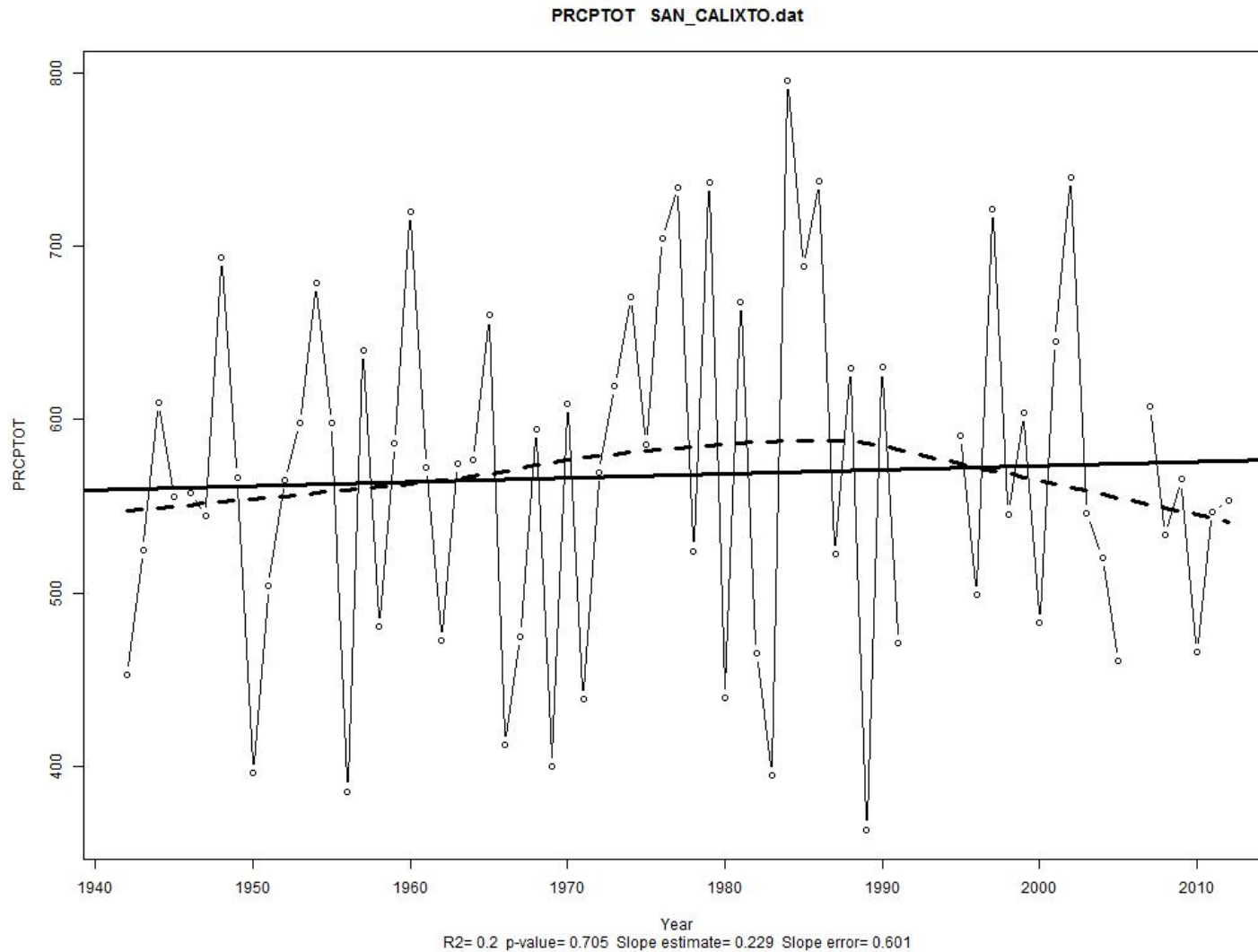
3596 total number of days

█ WD fraction inside the 95% CI

█ WD fraction outside the 95% CI

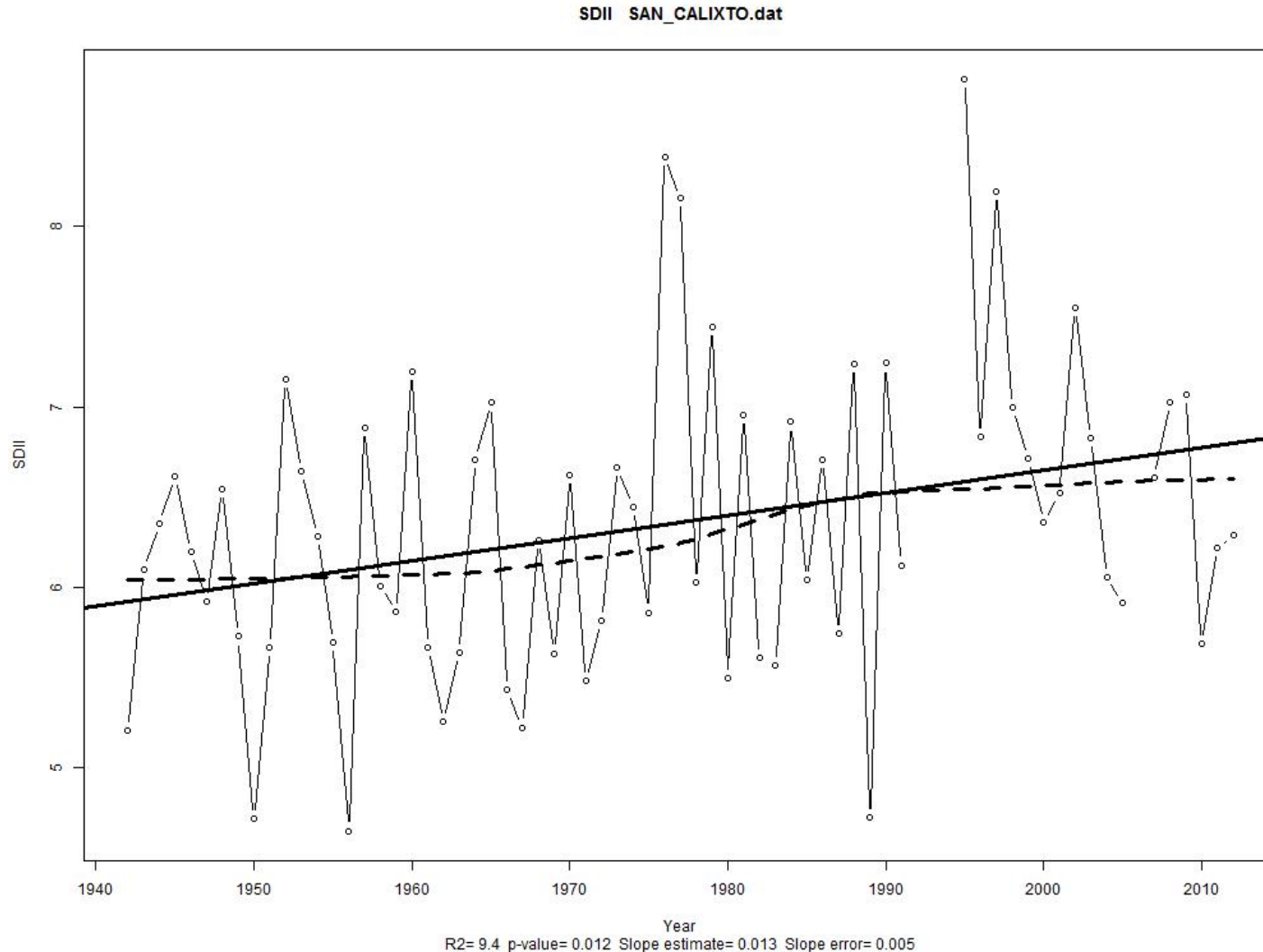
# Untagged rainfall accumulations

PRCPTOT: annual total wet-day precipitation



# Untagged rainfall accumulations

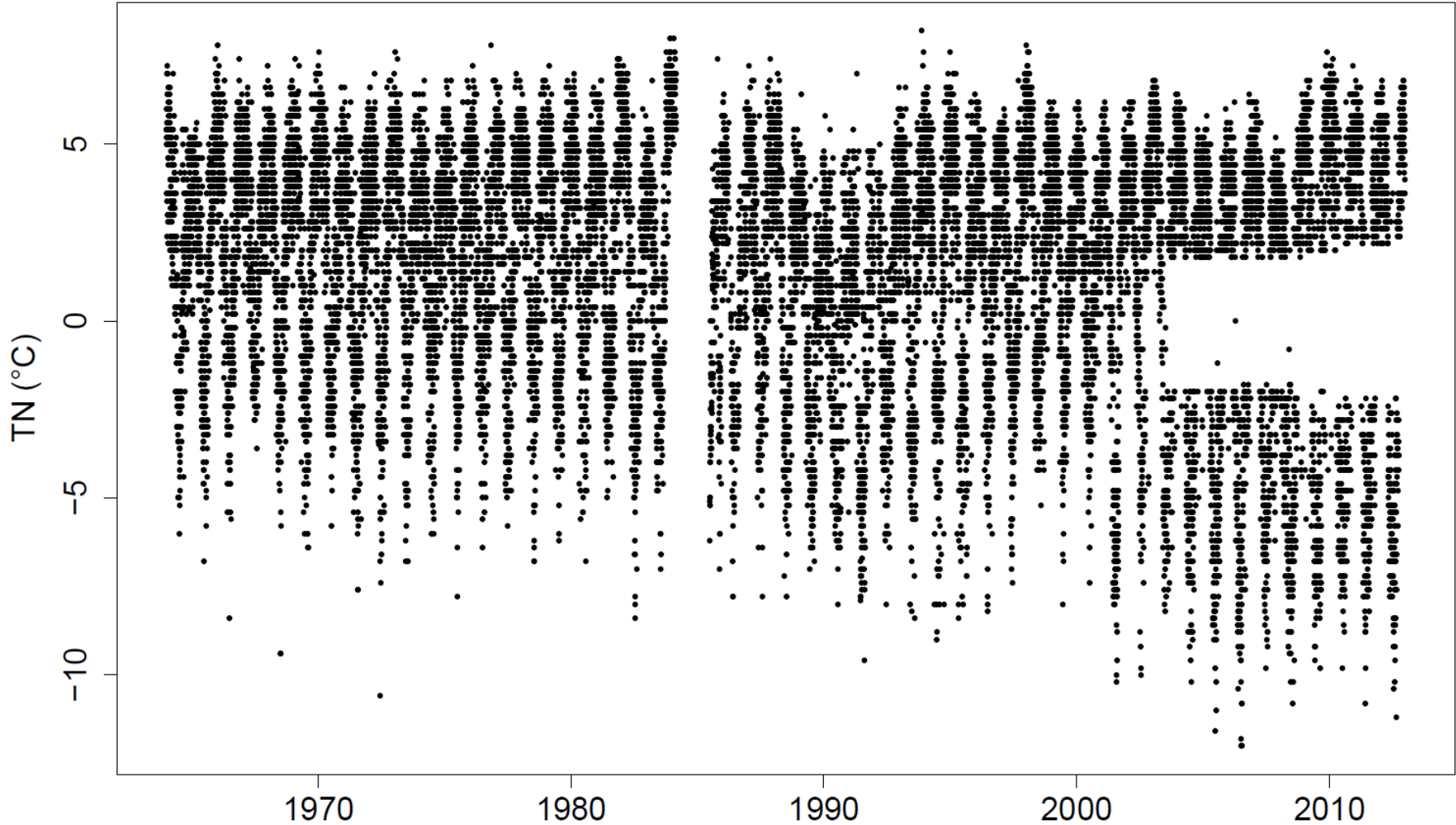
SDII (Simple Daily Intensity Index): annual total precipitation divided by the number of wet days in the year



# Error correction

$u^b$

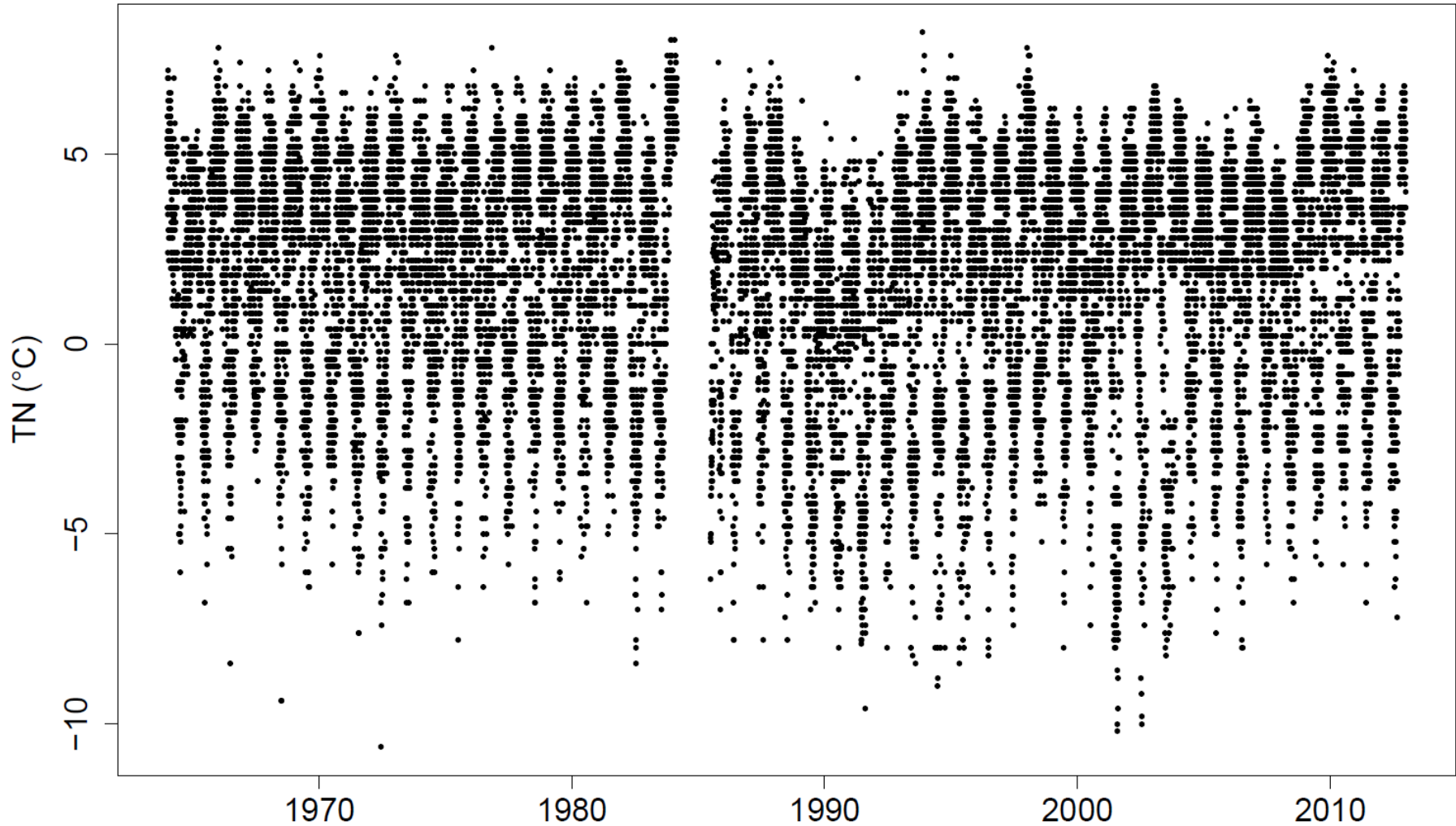
Progreso



# Error correction

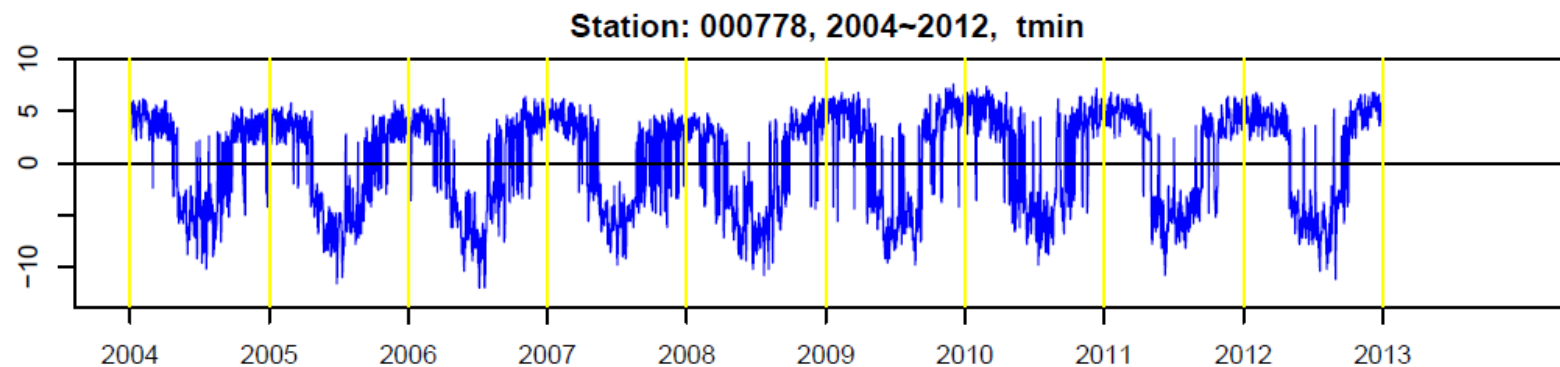
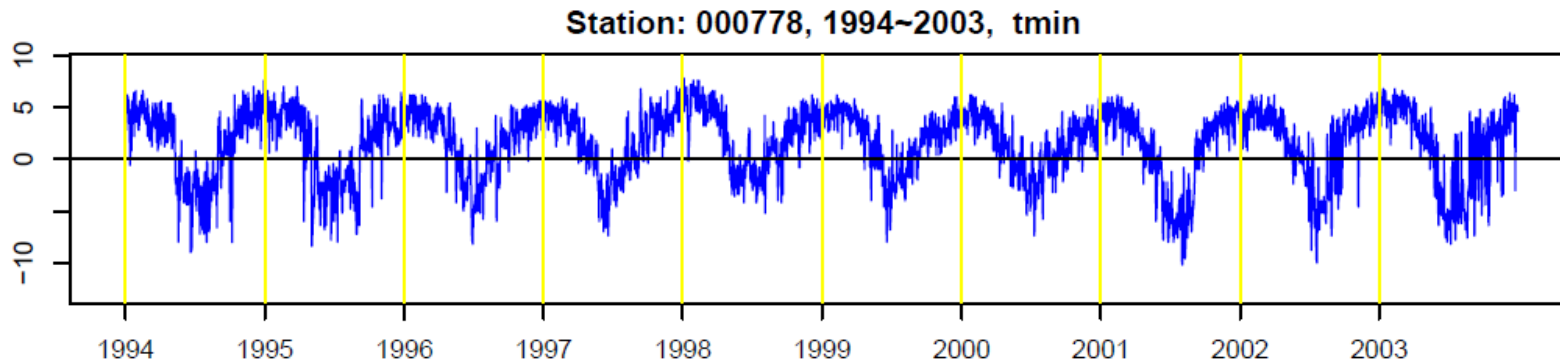
$u^b$

Progreso



# Error detection

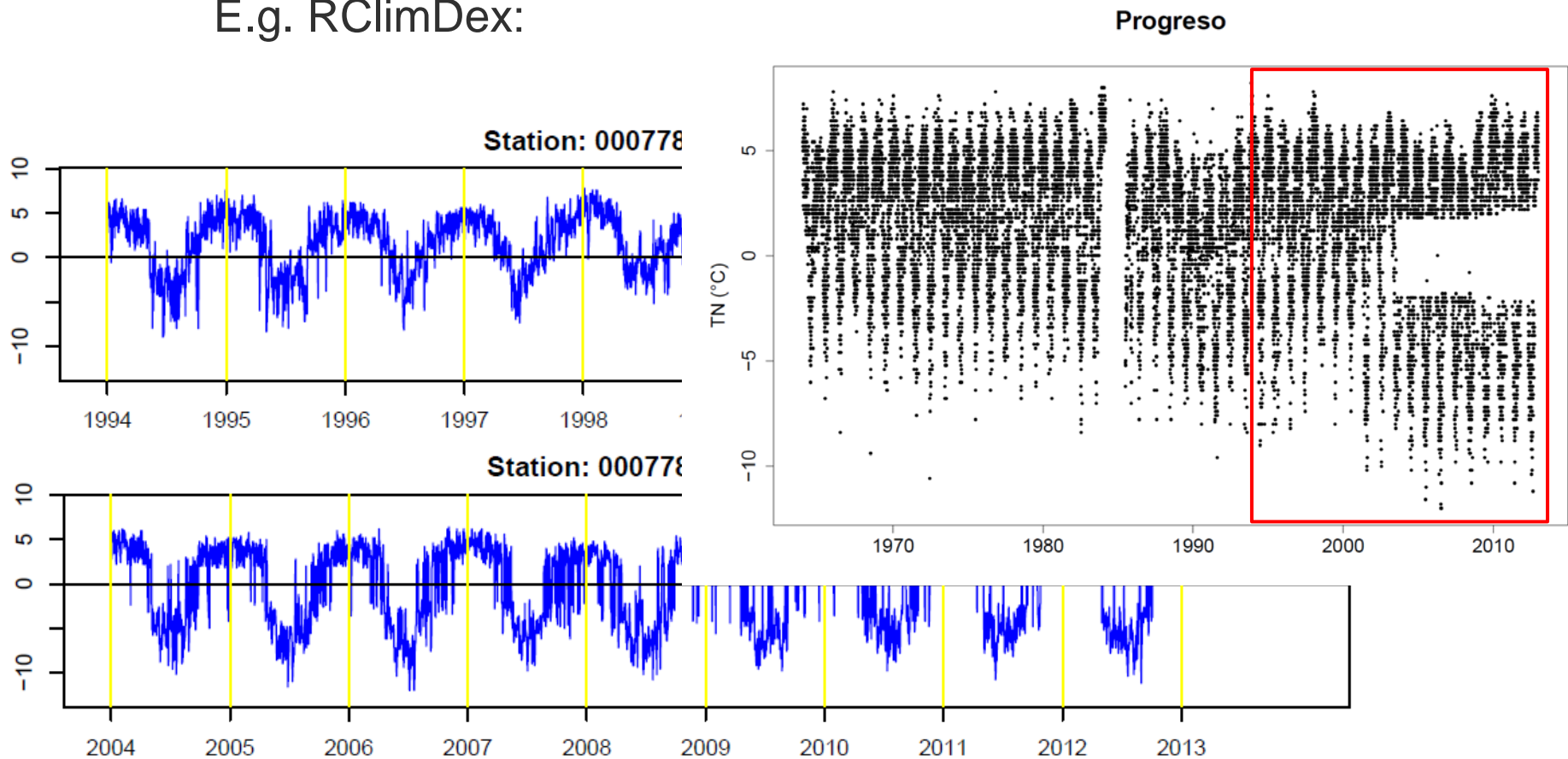
- QC is included in many programs
- Programs often create plots for visual quality control, e.g. RClimDex:



# Error detection

- Quality control is included in many programs
- Programs often create plots for visual quality control:

E.g. RClimDex:



- High priority to QC before analyzing data from Bolivia and Peru
- Ideally, metadata should be checked before using the data
- Visual QC is very helpful to detect patterns
  - use point instead of lines plots
- Reporting errors and observations in data
  - create additional metadata
- Knowing the source of the error allows to
  - possibly correct the error
  - decide if the error affects the data application of interest





$u^b$

<sup>b</sup>  
UNIVERSITÄT  
BERN

OESCHGER CENTRE  
CLIMATE CHANGE RESEARCH

# Thank you!

