9. DRAFT

Introduction to the Data Library (DL):

Trends and De-trending

Training Module
November 29, 2016
Version 1.0

International Research Institute for Climate and Society
EARTH INSTITUTE | COLUMBIA UNIVERSITY

International Research Institute for Climate and Society (IRI), (2016). Introduction to the Data Library (DL)- Trends and De-trending. November 29, Version 1.0. Palisades: IRI.

CONTENTS

CHAPTER

ONE

INTRODUCTION TO THE DATA LIBRARY (DL) - TRENDS AND DE-TRENDING

1.1 Introduction

The IRI Climate Data Library is a library of datasets. By library we mean a collection of datasets, collected from various sources, designed to make them more accessible for the library's users (Bluementhal, 2004). For this module we will be expanding on trends and de-trending. Traditional GIS platforms are now widely used by planners and decision makers in society. However, they are highly-focused on geospatial capabilities and have limited functionality for temporal analysis. Without information on the latter, meaningful inference about the causation of disease outbreaks is impossible (Jacquez 2000). Furthermore, many tools are unable to readily process the vast quantities of space-time data associated with, for example, the outputs of a global climate model. The IRI Climate Data Library overcomes the limitations imposed by GIS platforms by being based on a much more general multi-dimensional data model that includes both space and time dimensions. All datasets, including GIS features (such as points, lines, and polygons) are geo-located and temporally referenced in a uniform framework.

1.2 Overview

What will be working on?

For this manual we will be working on trend calculations [and subtraction]:

- Linear trend and De-trending
- "Differences" in time
- Subtraction of an average

1.3 Access

The IRI Data Library can be accessed with the following links:

- Worldwide: http://iridl.ldeo.columbia.edu/
- Chile: http://www.climatedatalibrary.cl/
- Venezuela: http://datoteca.ole2.org/
- Uruguay: http://dlibrary.snia.gub.uy/
- Rwanda: http://maproom.meteorwanda.gov.rw/
- Ethiopia: http://www.ethiometmaprooms.gov.et:8082/

- Tanzania: http://maproom.meteo.go.tz/
- Mali: http://197.155.140.164/
- Ghana: http://maps.meteo.gov.gh:89/
- Zambia: http://41.72.104.142/
- Madagascar: http://map.meteomadagascar.mg/
- Peru: http://ons.snirh.gob.pe/
- Niger: http://cradata.agrhymet.ne/
- Kenya (KMD): http://kmddl.meteo.go.ke:8081/
- Kenya (ICPAC): http://digilib.icpac.net/

1.4 Linear Trend and De-trending

The calculation of the linear trend depends on the period chosen E.g.: Case of the overall mean temperature (global mean surface temperature)

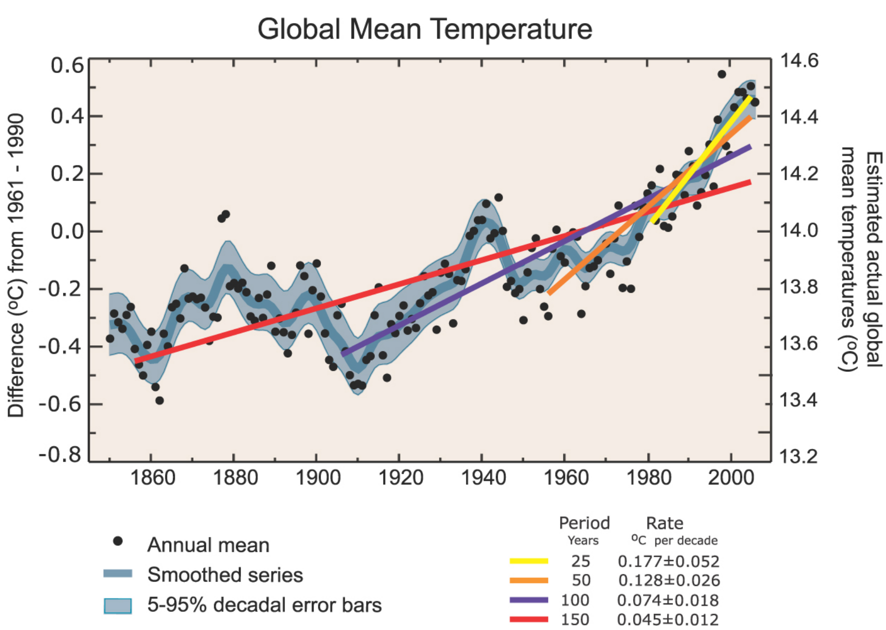


Fig. 1.1: Global Mean Temperature

Least-squares best-fit linear trend. It is a "special" case of linear regression, where x=time (look at Fig 1.2).

In this least-squares best-fit context, we can assess the statistical significance of the trend. This can be done by comparing Fig 1.3 (a), the hypothesis to test is whether a2 is significantly different from zero, for example Fig 1.3 (b).

Linear trend in DL, the function to be used is: detrend-bfl (Fig 1.4).

Fig 1.5 is the NASA/GISS global land average temperature with its linearly de-trended version.

$$y = a_1 + a_2 x$$

$$a_2 = \frac{\sigma_{yx}}{\sigma_x^2}$$

$$\sigma_{a_2} = \frac{1}{\sqrt{n-1}} \frac{\sigma_y}{\sigma_x}$$

$$\sigma_{yx} = \frac{1}{n-1} \sum_i (x_i - \bar{x})(y_i - \bar{y})$$

Fig. 1.2: Regression

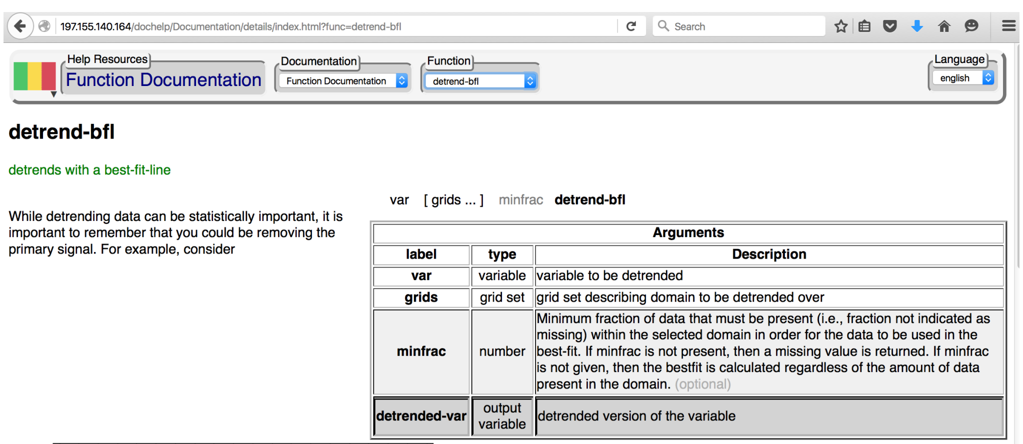
a)

$$a_2, \sigma_{a_2}$$

b)

$$a_2 \pm 2\sigma_{a_2} = 0$$

Fig. 1.3: Regression Comparison



detrend-bfl

detrends with a best-fit-line

While detrending data can be statistically important, it is important to remember that you could be removing the primary signal. For example, consider

var [grids ...] minfrac **detrend-bfl**

Arguments		
label	**type**	**Description**
var	variable	variable to be detrended
grids	grid set	grid set describing domain to be detrended over
minfrac	number	Minimum fraction of data that must be present (i.e., fraction not indicated as missing) within the selected domain in order for the data to be used in the best-fit. If minfrac is not present, then a missing value is returned. If minfrac is not given, then the bestfit is calculated regardless of the amount of data present in the domain. (optional)
detrended-var	output variable	detrended version of the variable

bfl = best fit line [quadratic fit]

Fig. 1.4: [detrend-bfl] Function

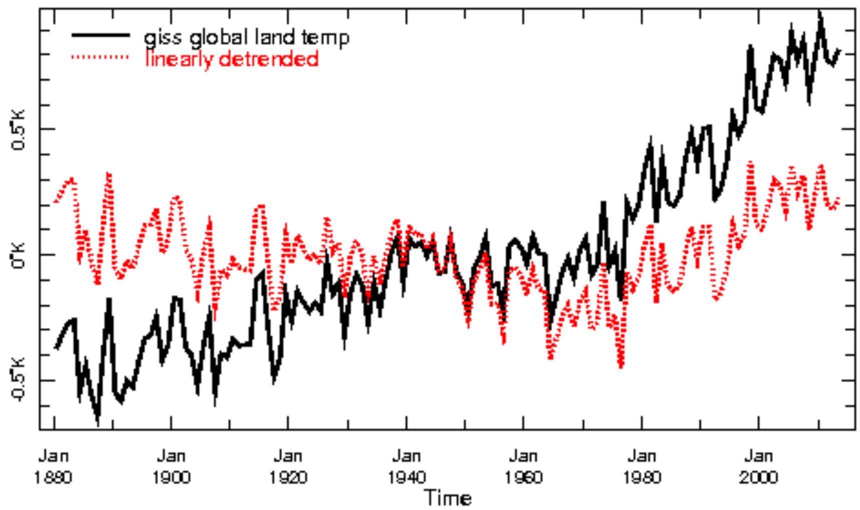


Fig. 1.5: NASA/GISS Global Land Average Temperature (black line) with its linearly de-trended version (red line)

The expert mode of Fig 1.5:

```
SOURCES .NASA .GISS .GISSTEMP .Global .Ts
T 12 boxAverage
a:
   /fullname (giss global land temp) def
:a:
   [T]detrend-bfl
   /fullname (linearly detrended) def
:a
T fig: medium black line red line :fig
```

or:

```
SOURCES .NASA .GISS .GISSTEMP .Global .Ts
T 12 boxAverage
/fullname (giss global land temp) def
dup
[T]detrend-bfl
/fullname (linearly detrended) def
```

[detrend-bfl] subtracts the linear trend. But the question on how to exploit it to calculate the trend itself is what follows. In other words - where is the trend in the figure 1.4. The trend is the difference between the two series !!!

1.4.1 Expert Mode

:: SOURCES .NASA .GISS .GISSTEMP .Global .Ts

T 12 boxAverage

dup

[T]detrend-bfl

sub

T fig: line :fig

Resulting in Figure 1.5.

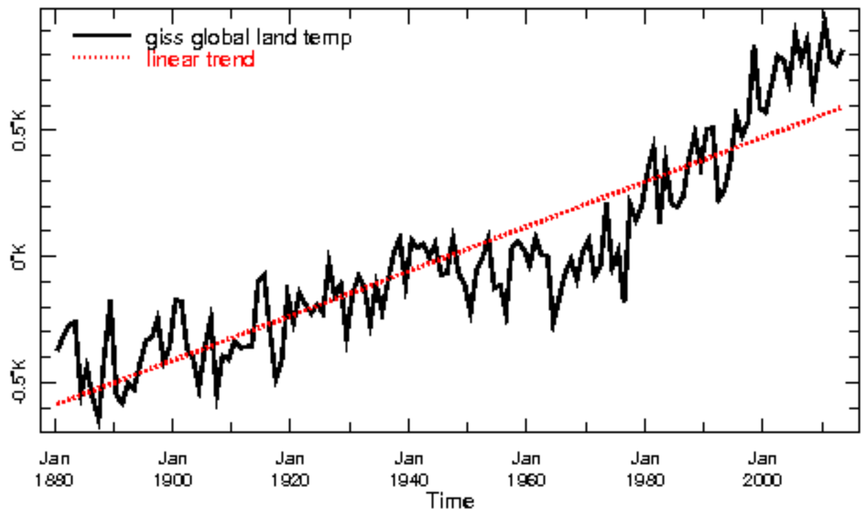


Fig. 1.6: NASA/GISS global land average temperature with its linear trend

:: SOURCES .NASA .GISS .GISSTEMP .Global .Ts T 12 boxAverage a: /fullname (giss global land temp) def :a: :a: :a [T]detrend-bfl sub /fullname (linear trend) def T fig: medium black line red line :fig

1.5 How to interpret a trend map

First we observe the trends in annual precipitation, refer to Fig. 1.6.

The we look at the observed trend in annual precipitation 1983-2014 through following expert mode and resulting in Fig 1.7.:

SOURCES .MaliMeteo .ENACTS .rainfall .MON .dekadly .rfe_merged T (Jan 1983) (Dec 2014) RANGE

monthlyAverage T 12 boxAverage

dup [T]detrend-bfl sub

dup T last VALUE exch T first VALUE sub

prcp_anomaly

DATA -100 100 RANGE

X Y fig: colors :fig

If we look back at 1.5 the best fit line doesn't quite capture the nature of the global mean temperature increase.

1.6 De-trend Using Time Differences

There is another simple technique to "de-trend", a technique that does not impose a shape to the fit, rather adapts to the data. It takes differences in time, between one time step and the previous. In the DL: [T] differences

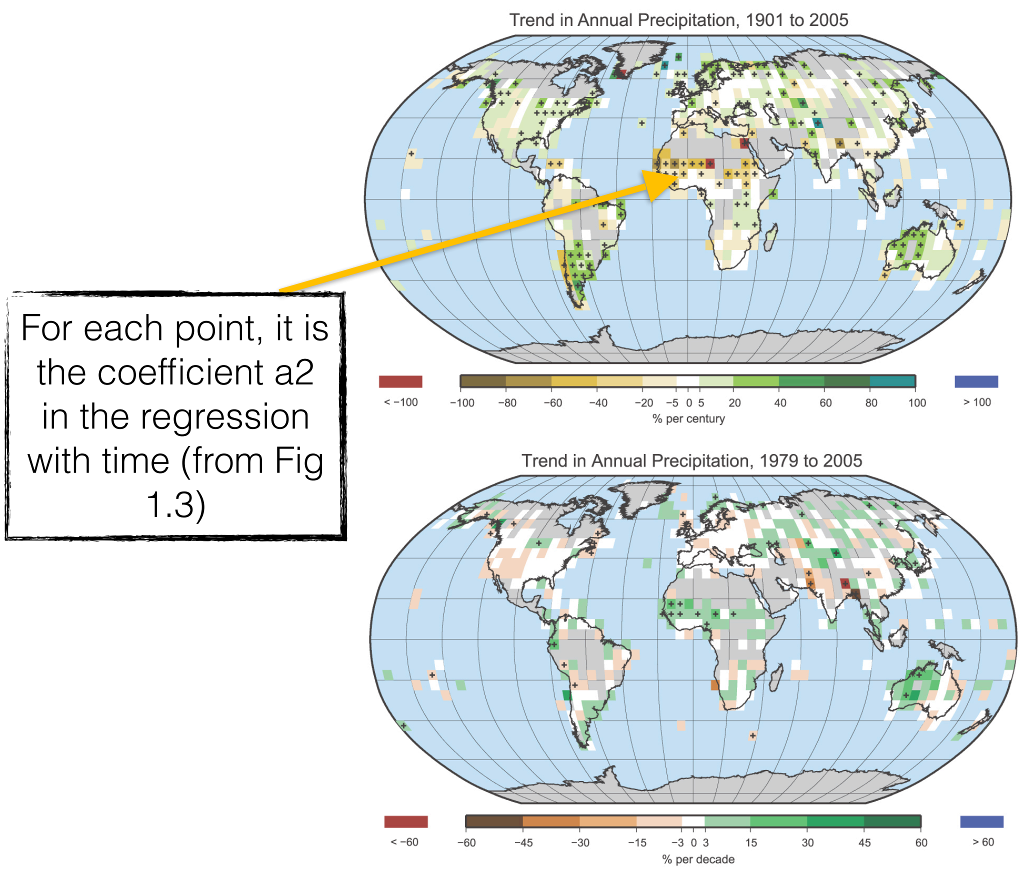


Fig. 1.7: Trends in Annual Precipitation in 1983 to 2005 in Mali

1.6.1 Example

Fig 1.8 shows the agricultural production and climate in Mali 1961-2009. The sum of production in: millet, sorghum, maize and rice.

Expert Mode:

SOURCES .UN .FAO .FAOSTAT .Mali .Feb2012 .Production .Crops .production

item_codes (79) (83) (56) (27) VALUES [item_codes]sum

T fig: red medium line :fig

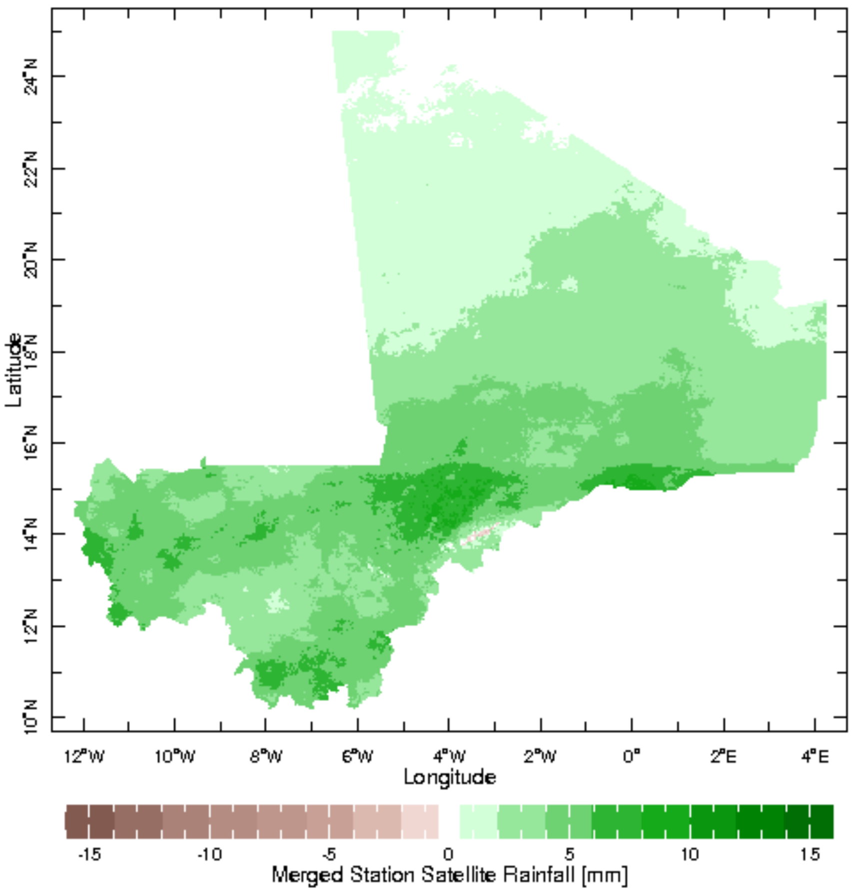


Fig. 1.8: Agricultural production in Mali 1961-2009 (http://iridl.ldeo.columbia.edu/SOURCES/.UN/.FAO/.FAOSTAT/.Mali/.Feb2012/.Production/.Crops/.production/item_codes/%2879%29%2883%29%2856%29%2827%29VALUES%5Bitem_codes%5Dsum%5BT%5Ddifferences/SOURCES/.WCRP/.GCOS/.GPCC/.FDP/.version5/.2p5/.prcp/Y/12/16/RANGE/X/-15/5/RANGE%5BX/Y%5Daverage/T/%28May-Oct%29seasonalAverage/T/%28May-Oct%201961%29%28May-Oct%202009%29RANGE/T/-1.5/shiftGRID%5BT%5Ddifferences/T/fig:/twolines/:fig/)

We add the precipitation into the figure with expert mode:

SOURCES .UN .FAO .FAOSTAT .Mali .Feb2012 .Production .Crops .production item_codes (79) (83) (56) (27) VALUES [item_codes]sum

SOURCES .WCRP .GCOS .GPCC .FDP .version5 .2p5 .prcp

Y 12 16 RANGE X -15 5 RANGE [X Y]average

T (May-Oct) seasonalAverage

T (May-Oct 1961) (May-Oct 2009) RANGE T -1 shiftGRID

T fig: twolines :fig

The result is Fig 1.9, showing the crop production and mean precipitation in Mali between 1961-2009.

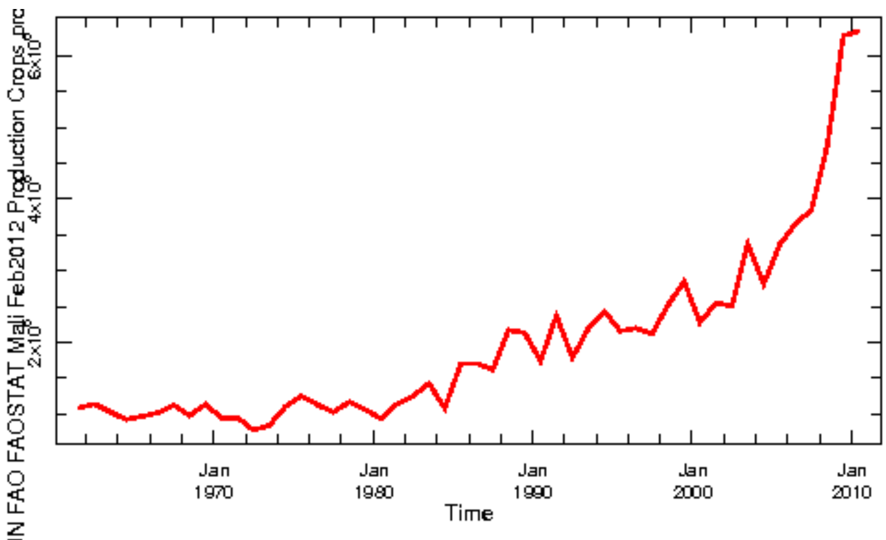


Fig. 1.9: Agricultural production and Climate in Mali 1961-2009 (http://iridl.ldeo.columbia.edu/SOURCES/.UN/.FAO/.FAOSTAT/.Mali/.Feb2012/.Production/.Crops/.production/item_codes/%2879%29%2883%29%2856%29%2827%29VALUES%5Bitem_codes%5Dsum/SOURCES/.WCRP/.GCOS/.GPCC/.FDP/.version5/.2p5/.prcp/Y/12/16/RANGE/X/-15/5/RANGE%5BX/Y%5Daverage/T/%28May-Oct%29seasonalAverage/T/%28May-Oct%201961%29%28May-Oct%202009%29RANGE/T/-1/shiftGRID/T/fig:/twolines/:fig/)

There are significant “trends” in both production and precipitation, but they are not linear.

Year-to-year differences in staple production and rainfall bring out their relationship clearly. So we de-trend by using the expert mode and resulting in Fig 1.10:

SOURCES .UN .FAO .FAOSTAT .Mali .Feb2012 .Production .Crops .production

item_codes (79) (83) (56) (27) VALUES [item_codes]sum

[T]differences

SOURCES .WCRP .GCOS .GPCC .FDP .version5 .2p5 .prcp

Y 12 16 RANGE X -15 5 RANGE [X Y] average

T (May-Oct) seasonalAverage

T (May-Oct 1961) (May-Oct 2009) RANGE T -1 shiftGRID

[T]differences

T fig: twolines :fig

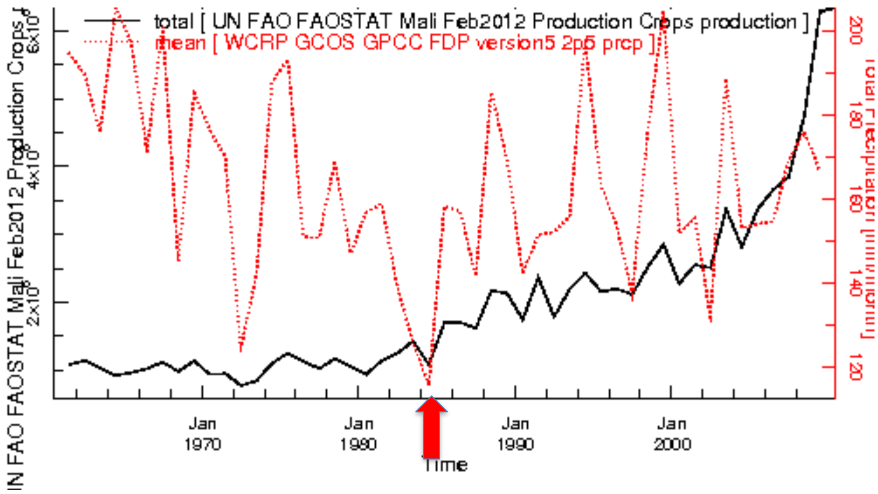


Fig. 1.10: Agricultural production and Climate in Mali 1961-2009 De-trended (http://iridl.ldeo.columbia.edu/SOURCES/.UN/.FAO/.FAOSTAT/.Mali/.Feb2012/.Production/.Crops/.production/item_codes/%2879%29%2883%29%2856%29%2827%29VALUES%5Bitem_codes%5Dsum%5BT%5Ddifferences/SOURCES/.WCRP/.GCOS/.GPCC/.FDP/.version5/.2p5/.prcp/Y/12/16/RANGE/X/-15/5/RANGE%5BX/Y%5Daverage/T/%28May-Oct%29seasonalAverage/T/%28May-Oct%201961%29%28May-Oct%202009%29RANGE/T/-1.5/shiftGRID%5BT%5Ddifferences/T/fig:/twolines/:fig/)

1.7 De-trend Using [runningAverage]

For this section we will be subtracting a dynamic average and de-trending.

The expert mode:

SOURCES .UEA .CRU .TS2p1 .monthly .prcp

X -20 40 RANGE Y 12 18 RANGE [X Y]average

T (Jan 1901) (Dec 2000) RANGE yearly-anomalies T (Jul 1901) (Sep 2000) RANGE T 3 boxAverage T 12 STEP

dup T 252 runningAverage

sub

T fig: red line :fig

1.7.1 The Results (Fig 1.11)

1.8 Summary

For this module we hope you are able to calculate Trends, De-trending, Differences in Time and Subtracting dynamic averages.

1.9 Quiz

Please answer the following questions using the IRI Data Library

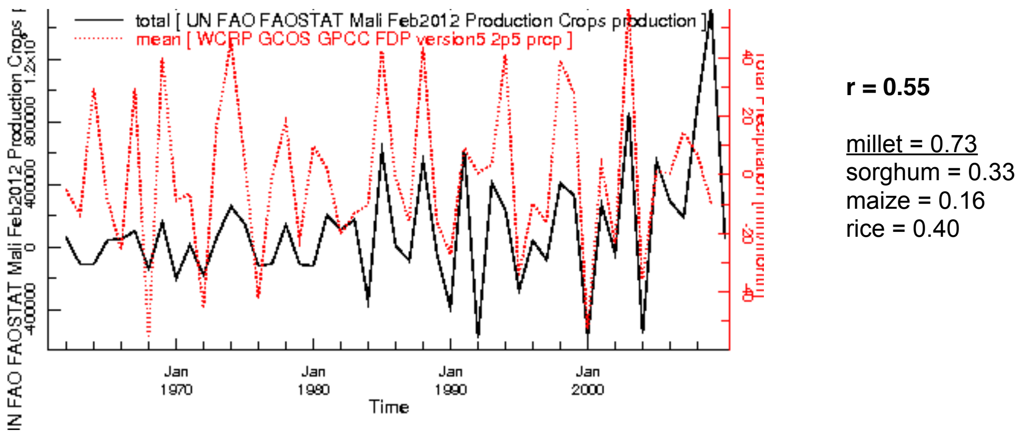


Fig. 1.11: Comparing original series vs the series after dynamic average is subtracted

Q1. What does the [detrend-bfl] function do?

Q2. What is the use of [T] differences?

1.9.1 Quiz - Answers

A1.[detrend-bfl] function, detrends with a best-fit -line.

A2. [T] differences takes the differences in time between one time step and the previous one.

1.10 Reference(s)

-