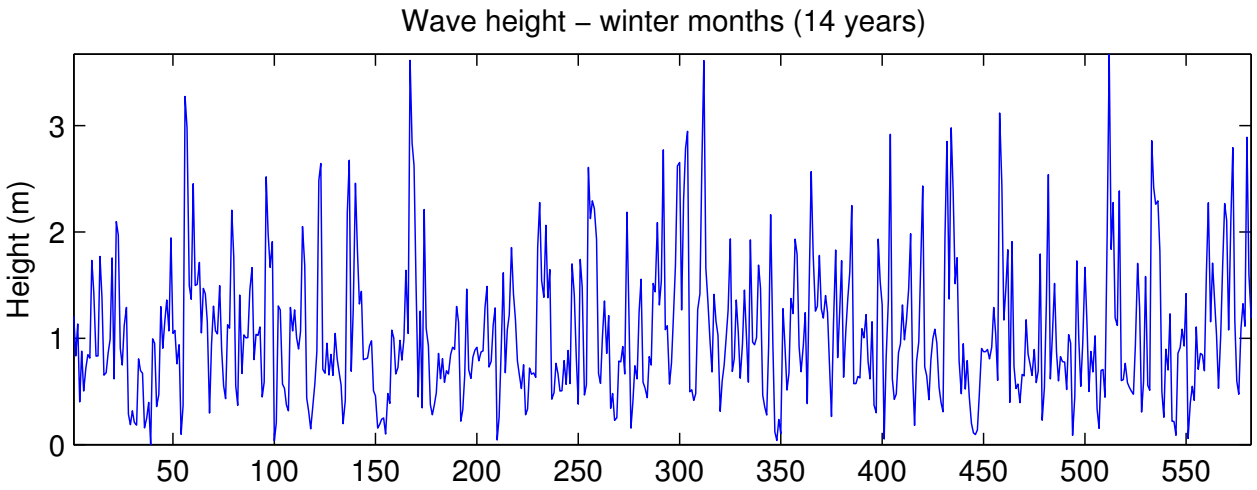
Statistics in Climate and Environment

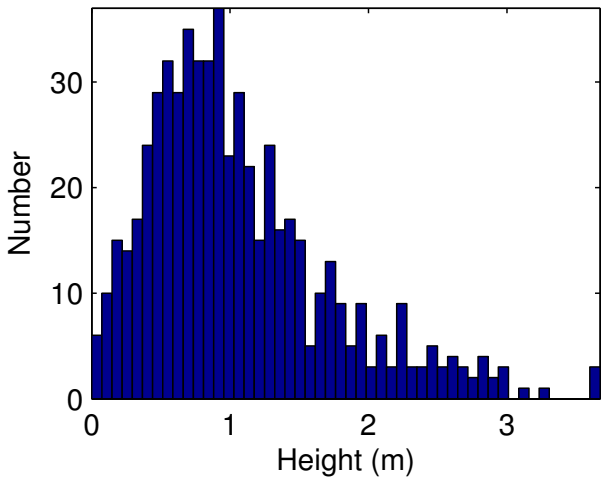
Johan Lindström

1 October 2014

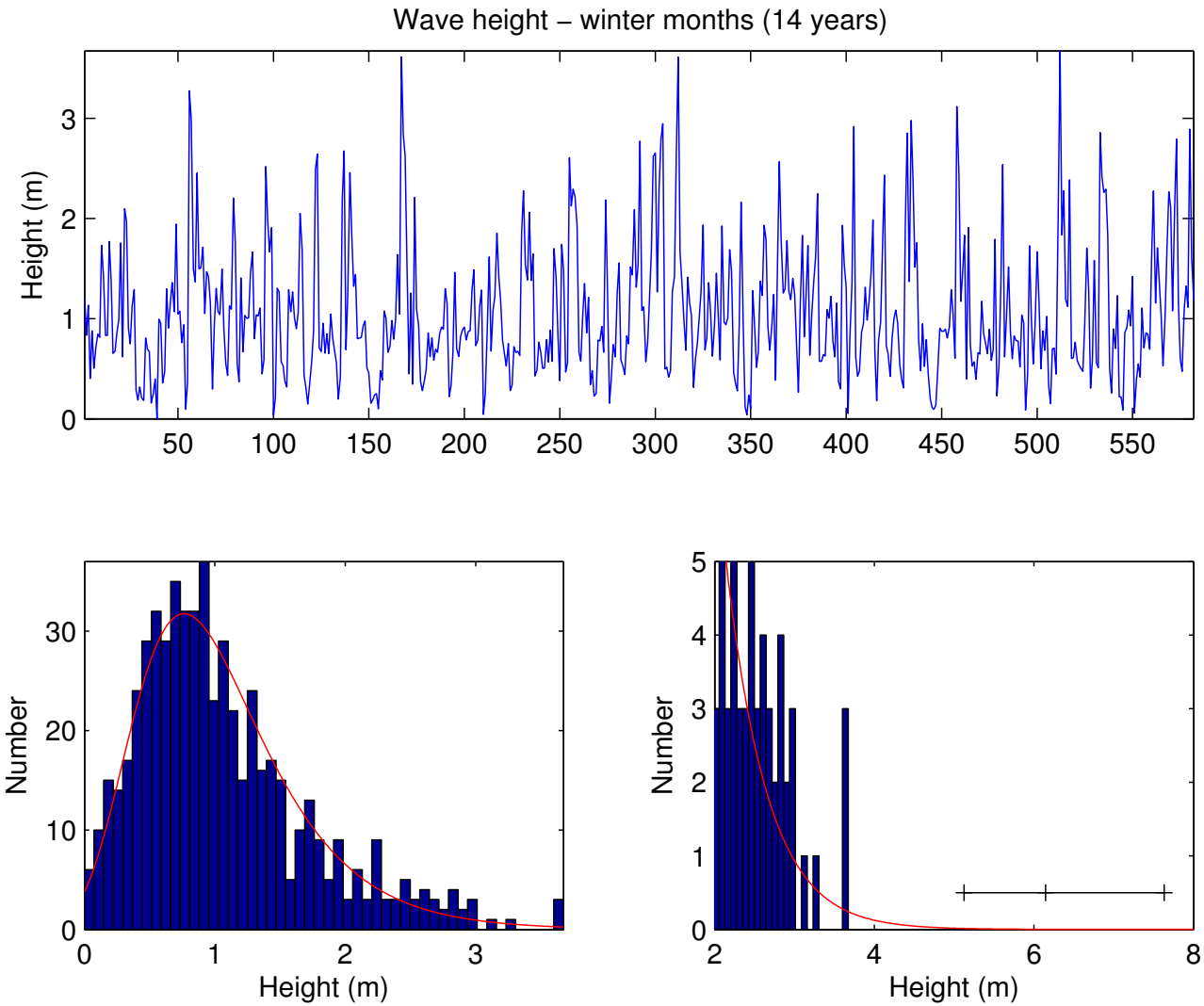
Goeree-Overflakkee, Netherlands, February 1953

How large waves should we expect?





How large waves should we expect?



Extreme Values — How large is the 1250–year wave?

Today the Netherlands uses an expected 1250–year wave to calculate the safety margin when building new dams.

For our data: 6.15 m (5.12,7.63)

Applications of extreme value theory.:

- Extreme weather
- Damages and loads
- Insurance industry

Course	Name	Points
MASM15	Stat. Model. of Extreme Values	7.5
MASM23	Stat. Model. of Multivariate Extremes	7.5

Extreme Values — How large is the 1250–year wave?

Today the Netherlands uses an expected 1250–year wave to calculate the safety margin when building new dams.
For our data: 6.15 m (5.12,7.63)

Applications of extreme value theory.:

- Extreme weather
- Damages and loads
- Insurance industry

Course	Name	Points
MASM15	Stat. Model. of Extreme Values	7.5
MASM23	Stat. Model. of Multivariate Extremes	7.5

Extreme Values — How large is the 1250–year wave?

Today the Netherlands uses an expected 1250–year wave to calculate the safety margin when building new dams.
For our data: 6.15 m (5.12,7.63)

Applications of extreme value theory.:

- Extreme weather
- Damages and loads
- Insurance industry

Course	Name	Points
MASM15	Stat. Model. of Extreme Values	7.5
MASM23	Stat. Model. of Multivariate Extremes	7.5

Extreme Values — How large is the 1250–year wave?

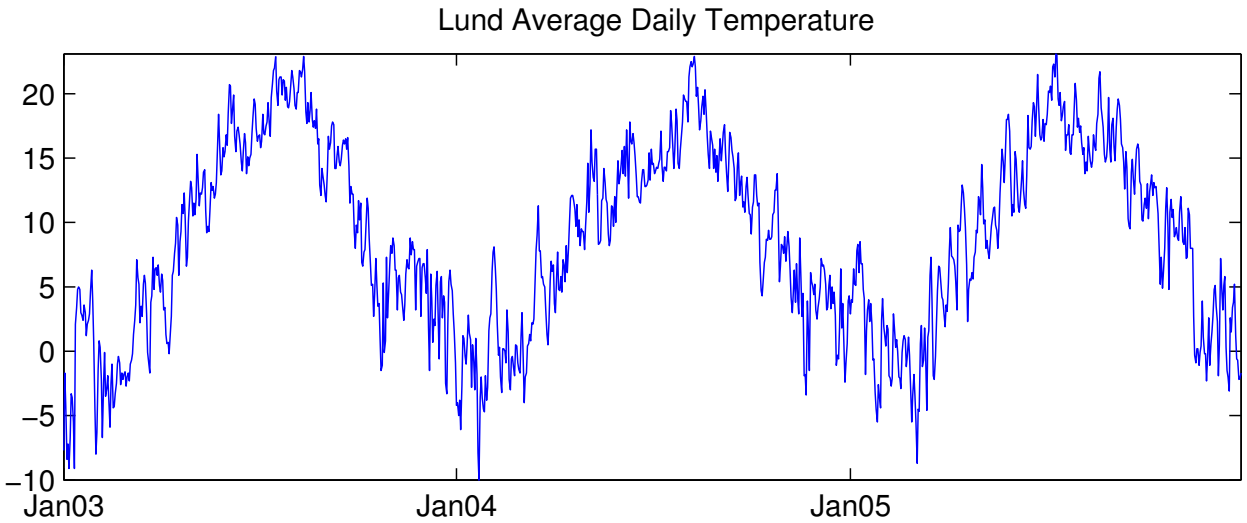
Today the Netherlands uses an expected 1250–year wave to calculate the safety margin when building new dams.
For our data: 6.15 m (5.12,7.63)

Applications of extreme value theory.:

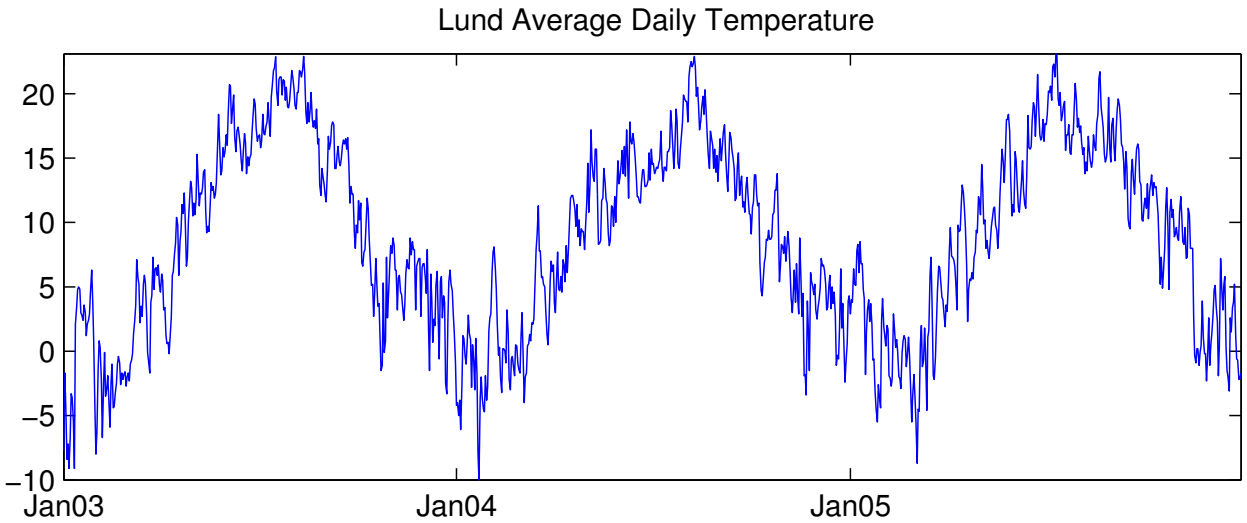
- Extreme weather
- Damages and loads
- Insurance industry

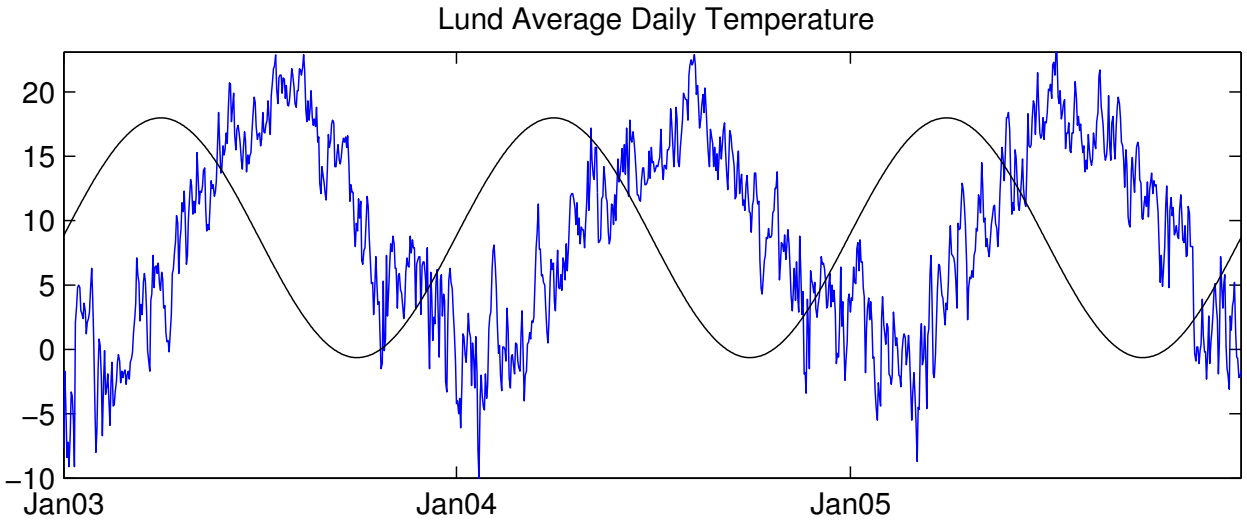
Course	Name	Points
MASM15	Stat. Model. of Extreme Values	7.5
MASM23	Stat. Model. of Multivariate Extremes	7.5

Modelling of Temperature

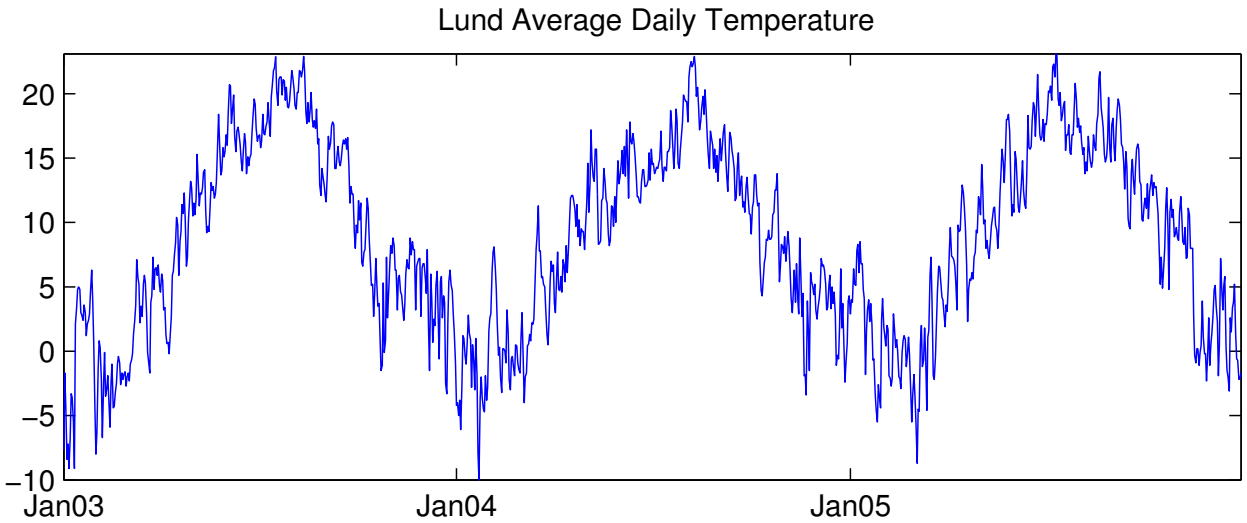


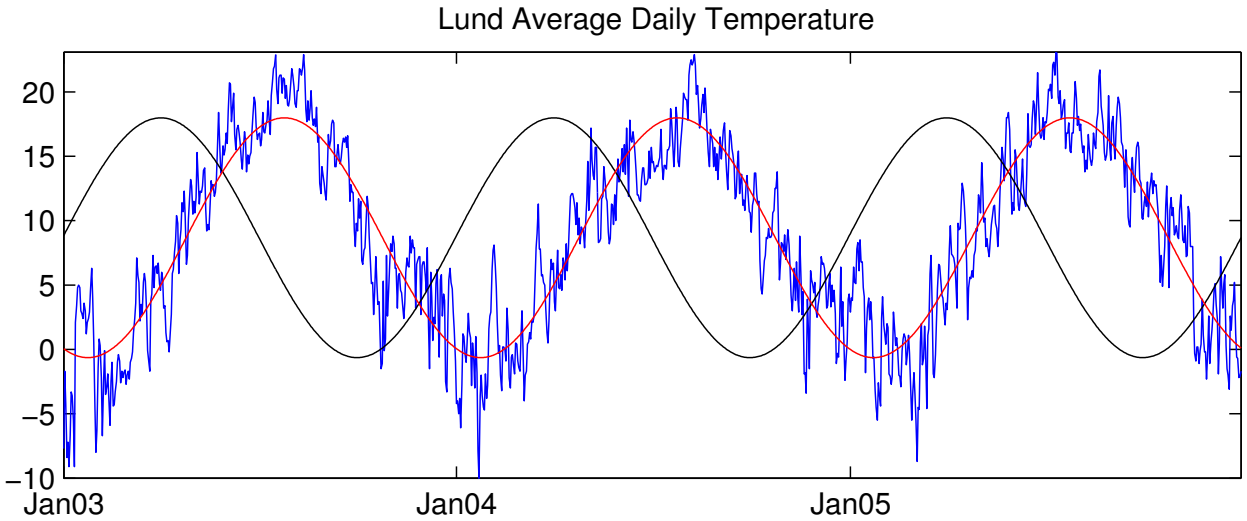
Modelling of Temperature



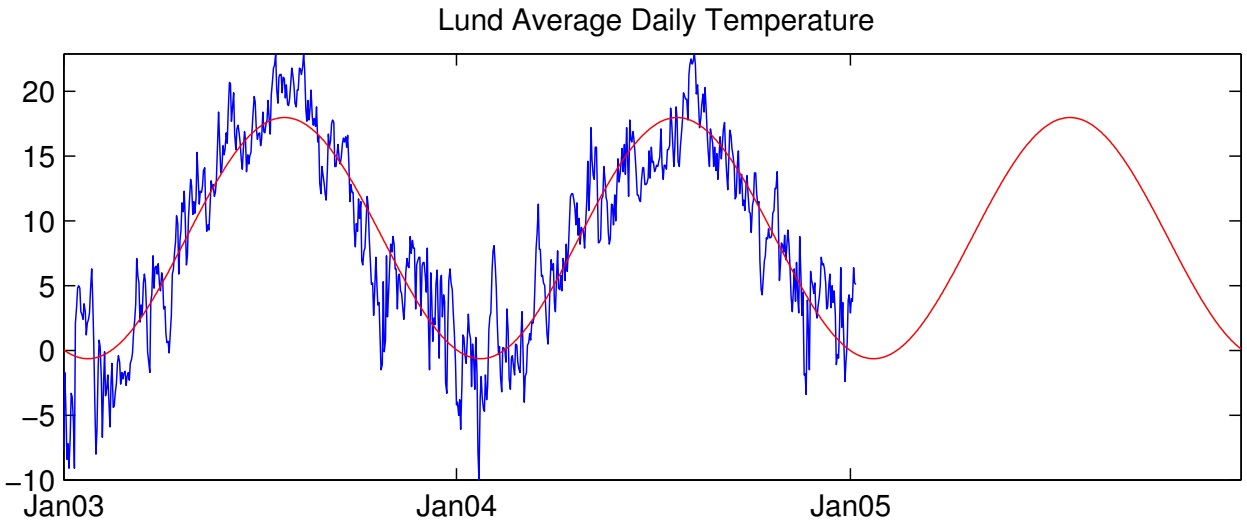


Modelling of Temperature

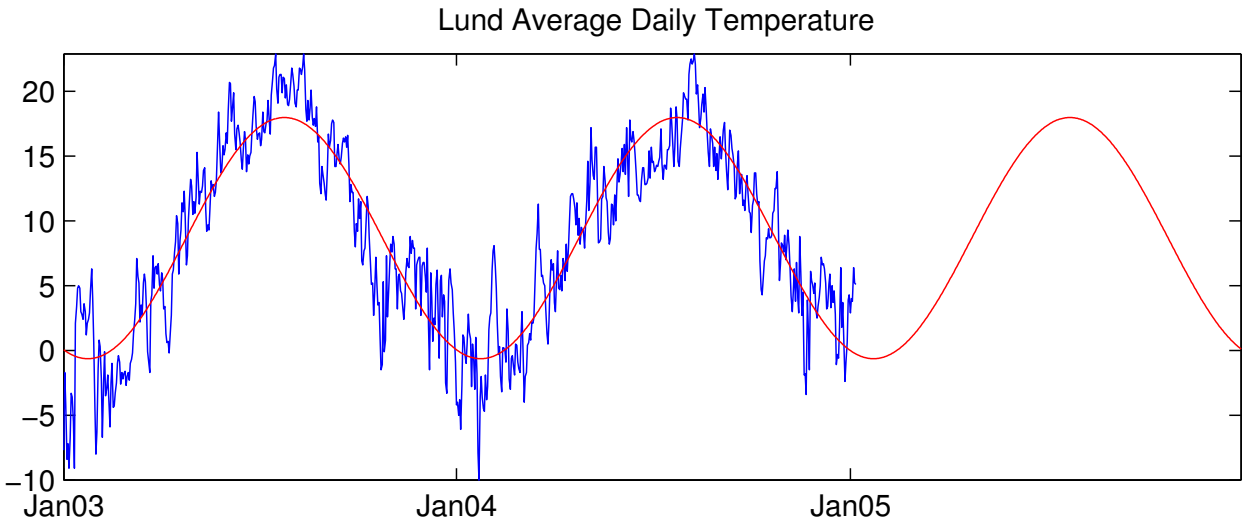


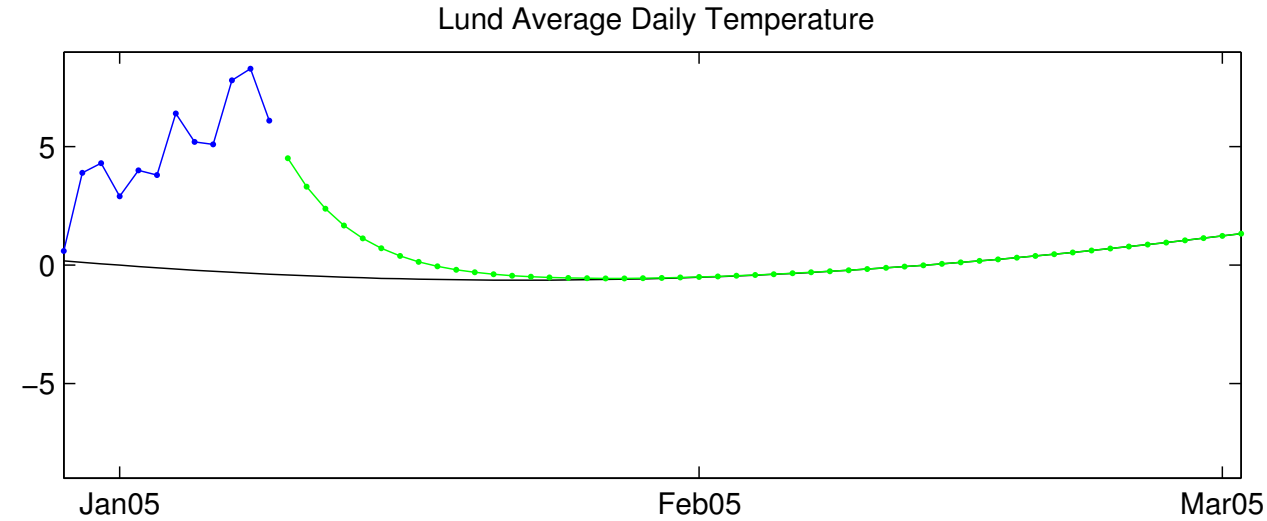


Temperature Prediction

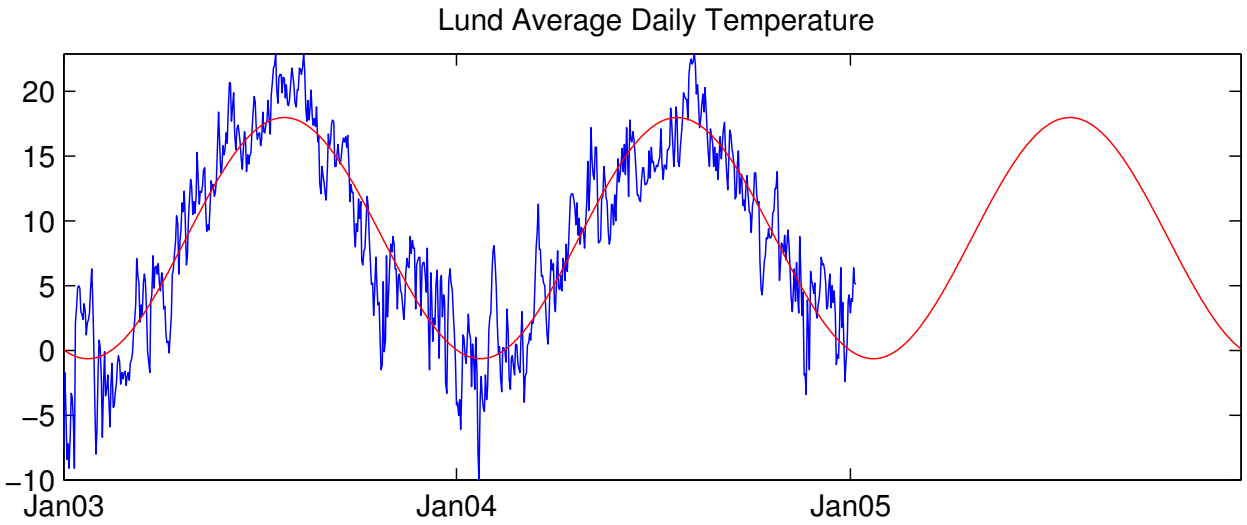


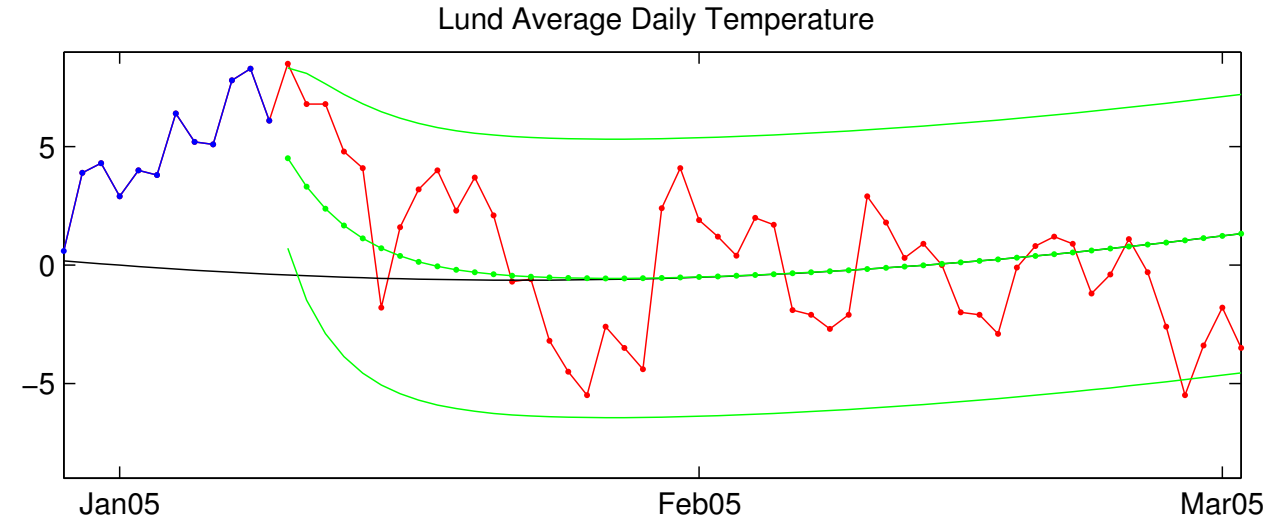
Temperature Prediction





Temperature Prediction





Time Series Data

Predictions of outdoor temperature can be used to estimate the need for district heating (fjärrvärme).

Time series data has many different applications:

- Weather data
- Consumption of electricity, district heating
- EKG-signals
- Financial data

Course	Name	Points
MASC04	Stationary Stochastic Processes	7.5
MASM17	Time Series Analysis	7.5
MASM26	Spectral Analysis[1]	7.5
MASM12	Non-linear Time Series Analysis[2]	7.5

[1] Given 2016

[2] Given 2014

Time Series Data

Predictions of outdoor temperature can be used to estimate the need for district heating (fjärrvärme).

Time series data has many different applications:

- Weather data
- Consumption of electricity, district heating
- EKG-signals
- Financial data

Course	Name	Points
MASC04	Stationary Stochastic Processes	7.5
MASM17	Time Series Analysis	7.5
MASM26	Spectral Analysis[1]	7.5
MASM12	Non-linear Time Series Analysis[2]	7.5

[1] Given 2016

[2] Given 2014

Time Series Data

Predictions of outdoor temperature can be used to estimate the need for district heating (fjärrvärme).

Time series data has many different applications:

- Weather data
- Consumption of electricity, district heating
- EKG-signals
- Financial data

Course	Name	Points
MASC04	Stationary Stochastic Processes	7.5
MASM17	Time Series Analysis	7.5
MASM26	Spectral Analysis[1]	7.5
MASM12	Non-linear Time Series Analysis[2]	7.5

[1] Given 2016

[2] Given 2014

Multi-Ethnic Study of Atherosclerosis — Air pollution

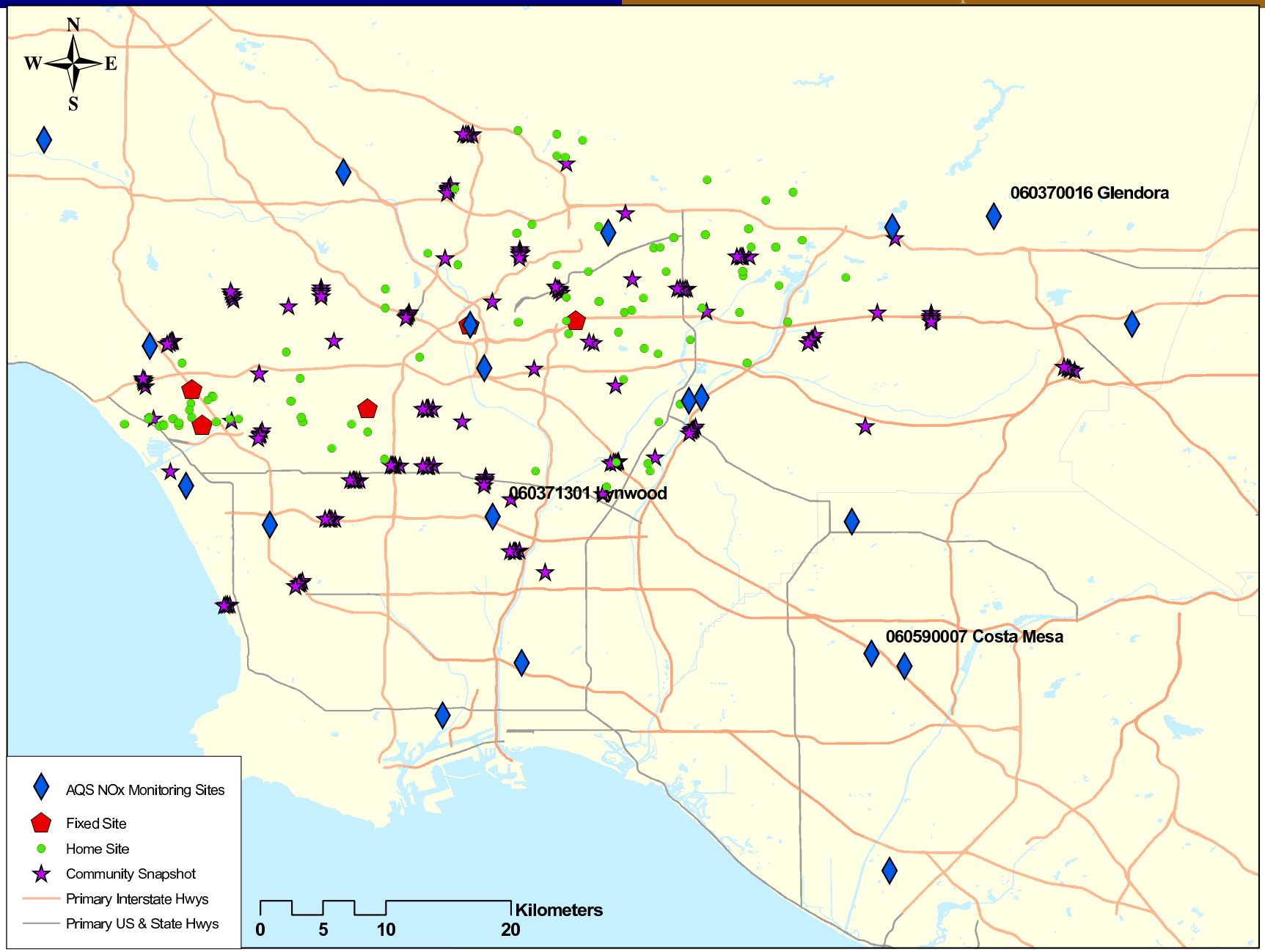
"The purpose of the MESA Air Pollution study is to relate how the amount of air pollution you breathe may be related to early stages of heart diseases and diseases of the blood vessels and lung."

- EPA funded study to investigate the relationship between **air pollution** and **cardiovascular disease** .
- $>$ 6 000 people studied for 10+ years.
- 6 metropolitan areas (Baltimore, Chicago, Los Angeles, Minneapolis–St. Paul, New York, Winston–Salem).
- Main concern $PM_{2.5}$ and NO_x.

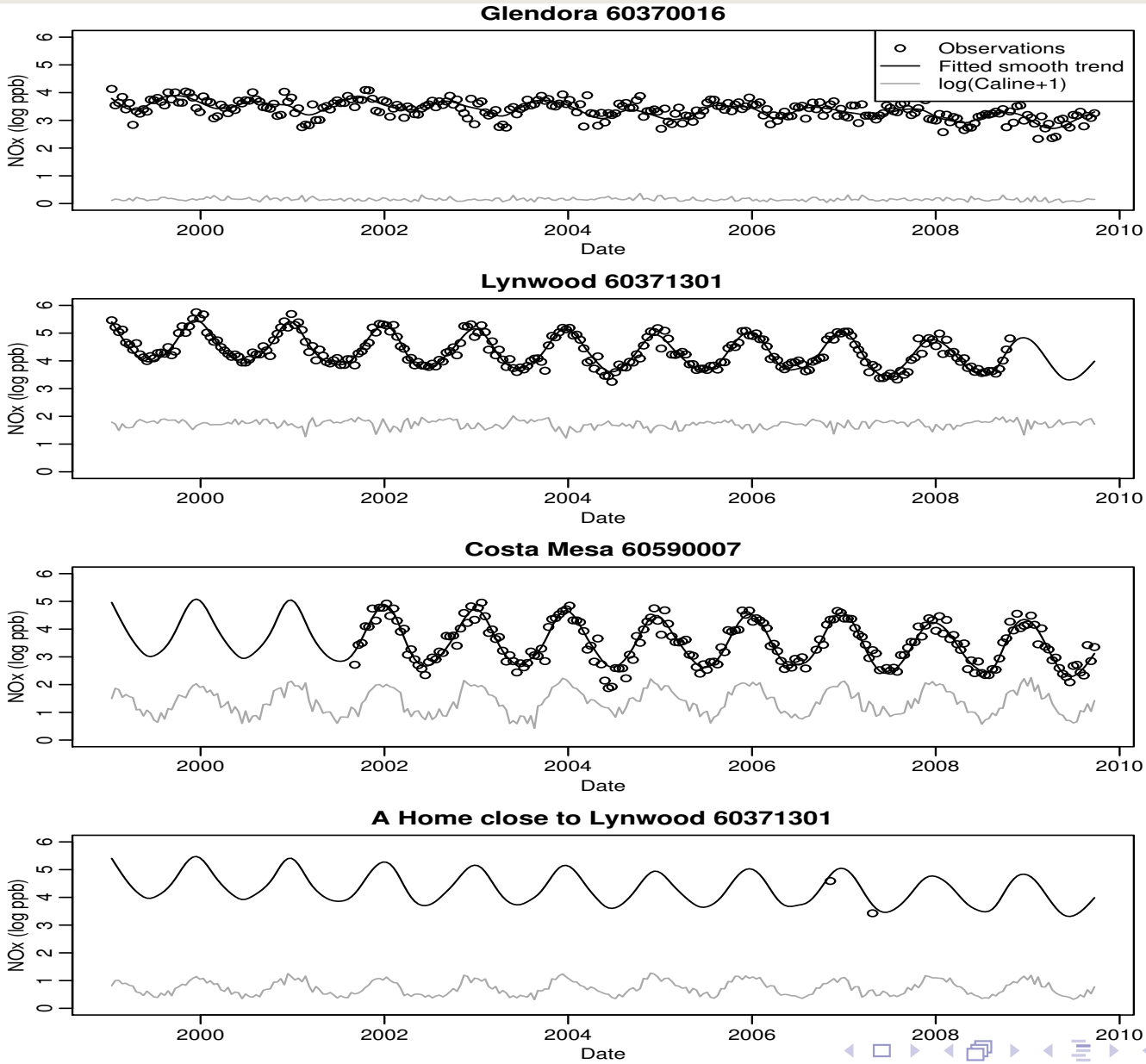
Multi-Ethnic Study of Atherosclerosis — Air pollution

"The purpose of the MESA Air Pollution study is to relate how the amount of air pollution you breathe may be related to early stages of heart diseases and diseases of the blood vessels and lung."

- EPA funded study to investigate the relationship between **air pollution** and **cardiovascular disease** .
- $>$ 6 000 people studied for 10+ years.
- 6 metropolitan areas (Baltimore, Chicago, Los Angeles, Minneapolis–St. Paul, New York, Winston–Salem).
- Main concern $PM_{2.5}$ and NO_x.


N
W
E
S
060370016 Glendora
060371301 Lynwood
060590007 Costa Mesa
AQS NOx Monitoring Sites
Fixed Site
Home Site
Community Snapshot
Primary Interstate Hwys
Primary US & State Hwys
0
5
10
20
Kilometers

Data — Los Angeles



Data — Los Angeles

A number of features that a model needs to capture:

- Mean level varies among locations.
- Seasonal structure has different amplitude (differs between coastal and inland locations).
- Decreasing trend during the 10 years.

A large number of possible explanatory variables are available:

- Distance to coast
- Distance to major roads
- Population density

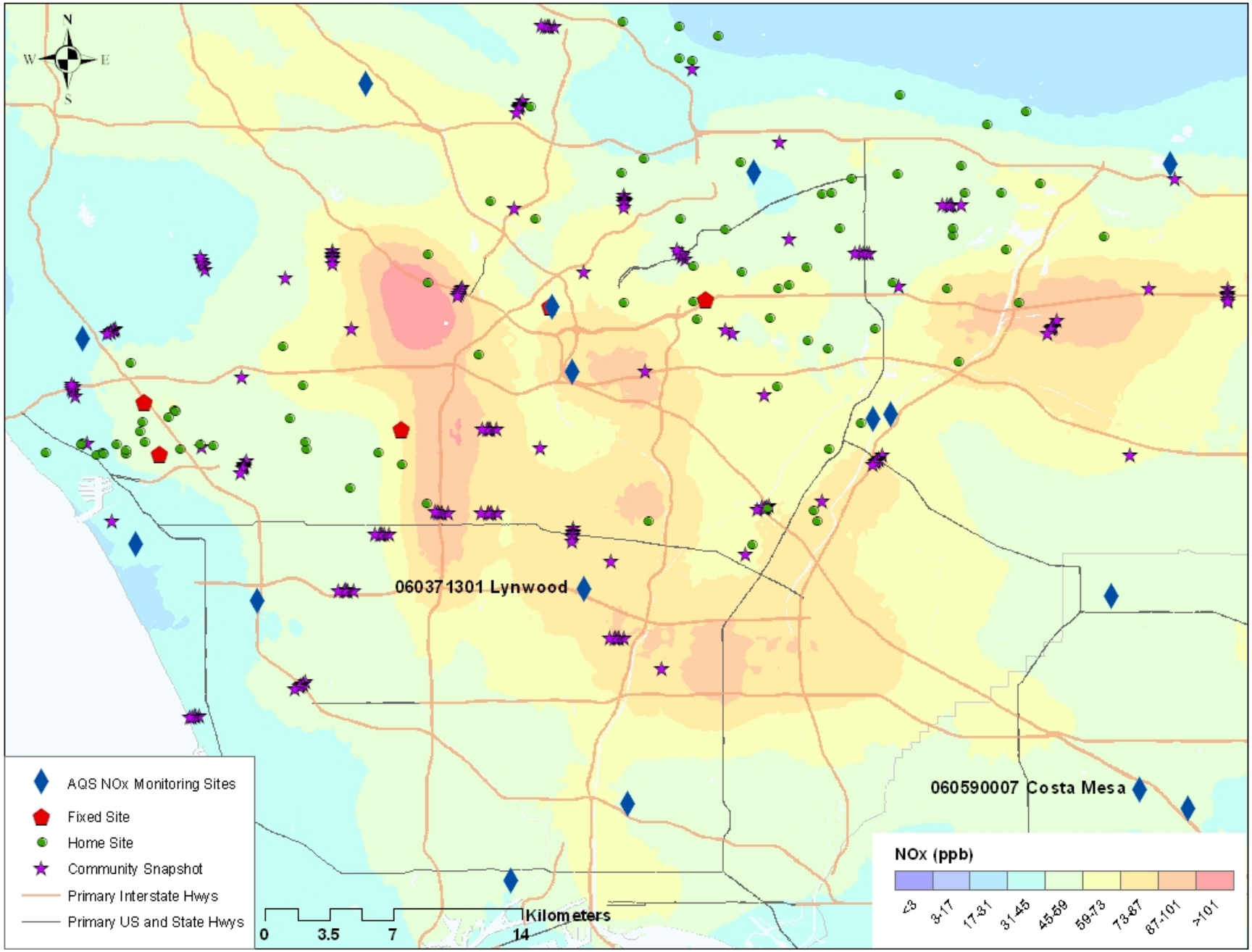
Data — Los Angeles

A number of features that a model needs to capture:

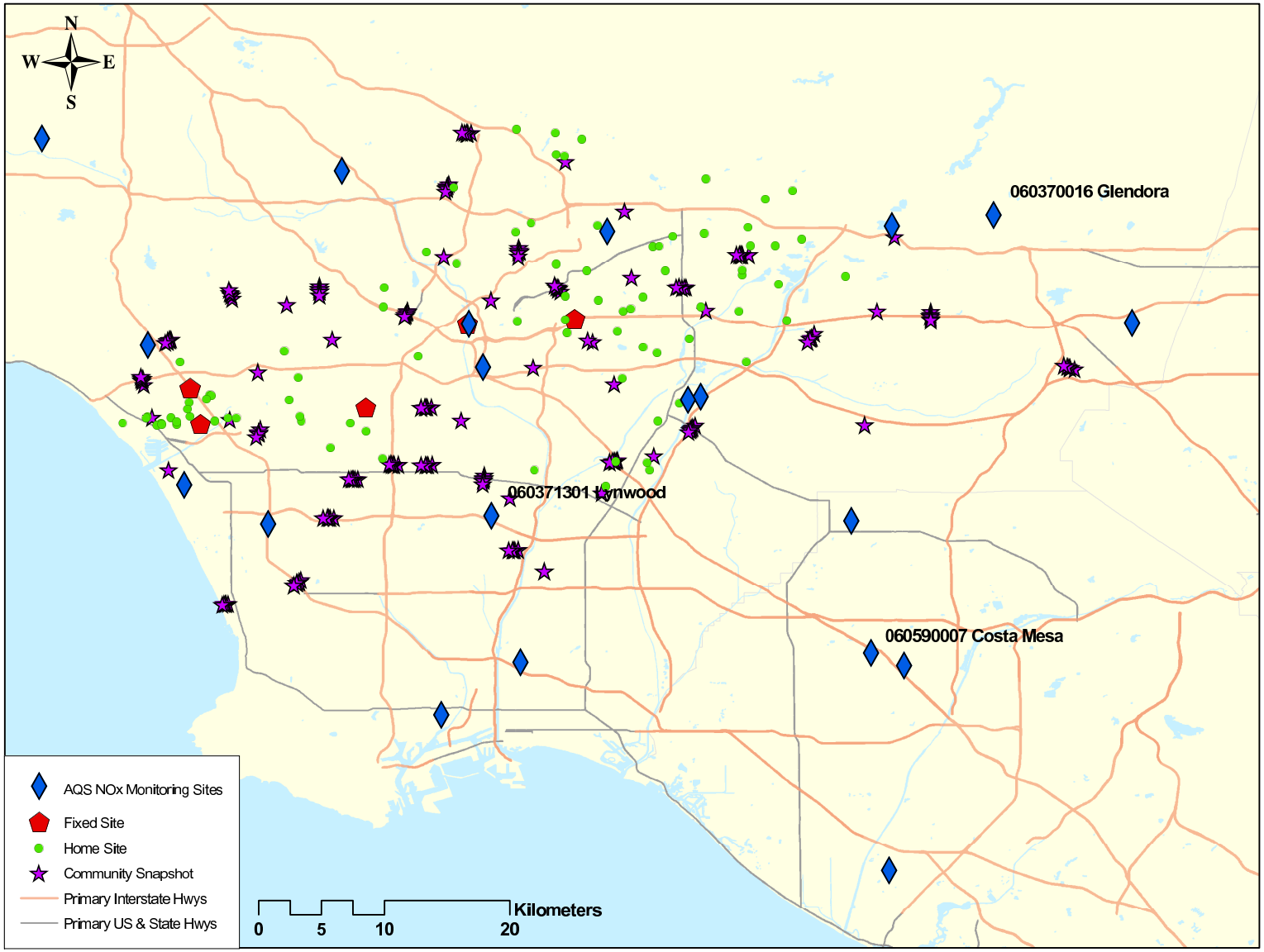
- Mean level varies among locations.
- Seasonal structure has different amplitude (differs between coastal and inland locations).
- Decreasing trend during the 10 years.

A large number of possible explanatory variables are available:

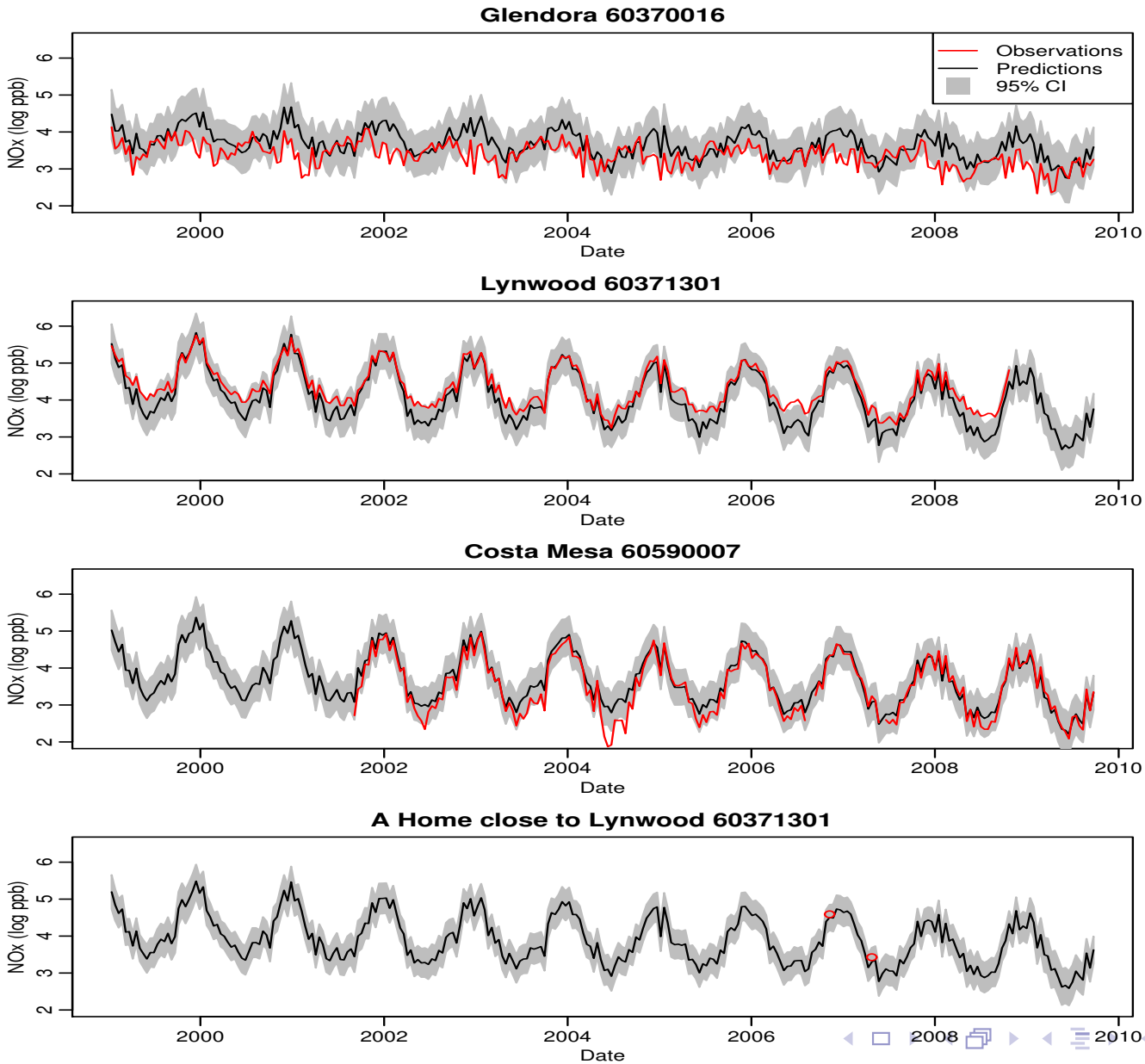
- Distance to coast
- Distance to major roads
- Population density


N
W
E
S
060371301 Lynwood
060590007 Costa Mesa
AQS NOx Monitoring Sites
Fixed Site
Home Site
Community Snapshot
Primary Interstate Hwys
Primary US and State Hwys
0
3.5
7
14
Kilometers
NOx (ppb)
<3
3-17
17-31
31-45
45-59
59-73
73-87
87-101
>101

Model Validation



Model Validation



Data in space

- Climate and Environmental data
- Satellite images
- Medical imaging

Course	Name	Points
MASM25	Spatial Statistics with Image Analysis	7.5
MASM11	Monte Carlo Methods for Inference	7.5

Data in space

- Climate and Environmental data
- Satellite images
- Medical imaging

Course	Name	Points
MASM25	Spatial Statistics with Image Analysis	7.5
MASM11	Monte Carlo Methods for Inference	7.5

Questions?