Model Skills and Uncertainties

An evaluation of model prediction performance against observations at eight LBA-MIP sites

- Latent Heat Flux -

Aaron Polhamus and Dr Joshua Fisher, NASA JPL
April 18, 2011

Acknowledgements

- *LBA-DMIP team*
- *Modeling teams*
- *Michel Muza*
- *Grayson Badgley*
- *Biosphere 2*

Data

- Hourly observations for multiple years of key meteorological/biophysical variables, with corresponding model predictions
- We average all data to the monthly scale

The challenge

- Identify which models perform well against the observations (assumption: perfect observations)
- Identify potentially problematic sites & models

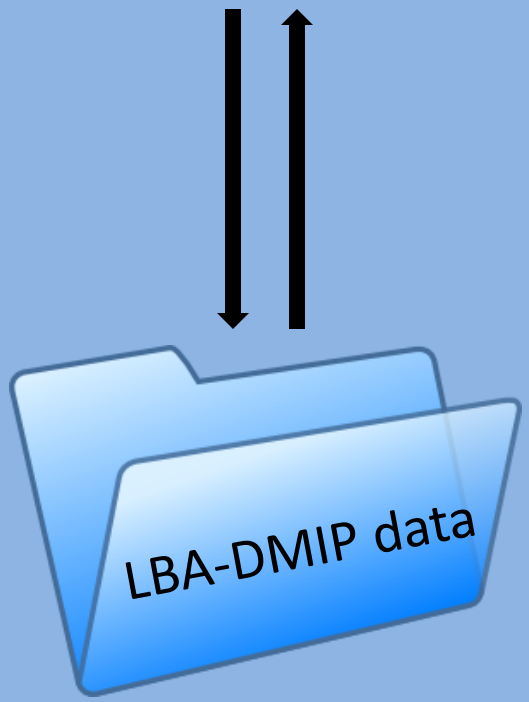
Method

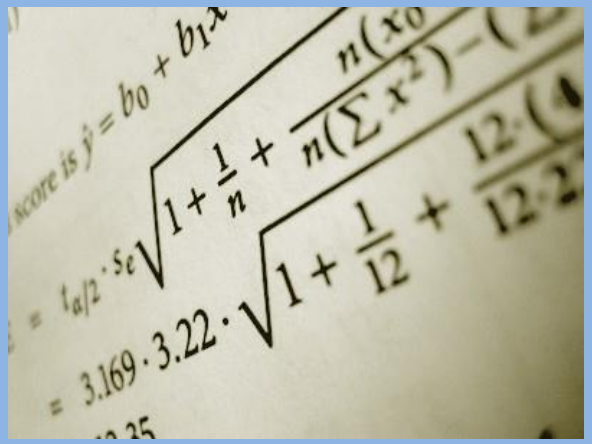
- Plot time series
- Bias (MBE), spread (RMSE), and agreement (R^2) on a per-site/per-model basis

Broad objective: a protocol for automated analysis

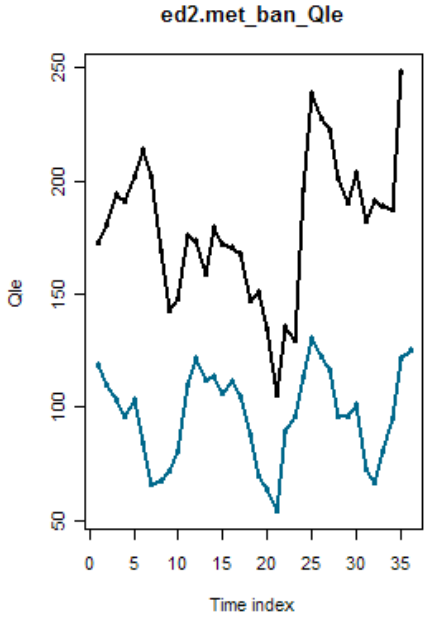
Select:

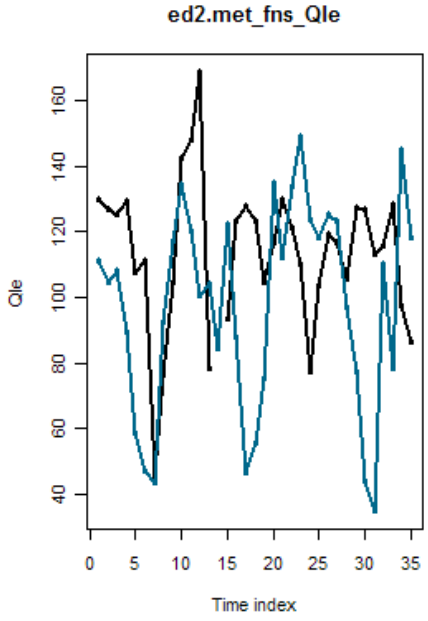
- Variables of interest and corresponding model output
- Whether daily or monthly averages of hourly data are desire
- Daytime or daily averages

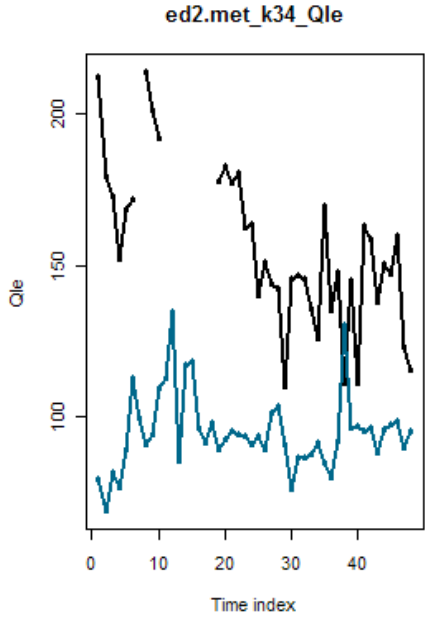


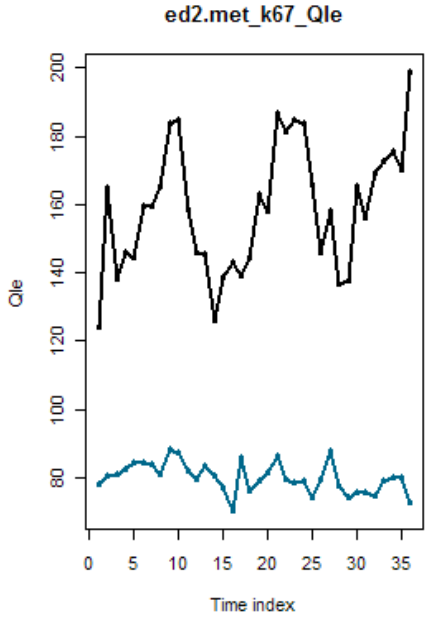


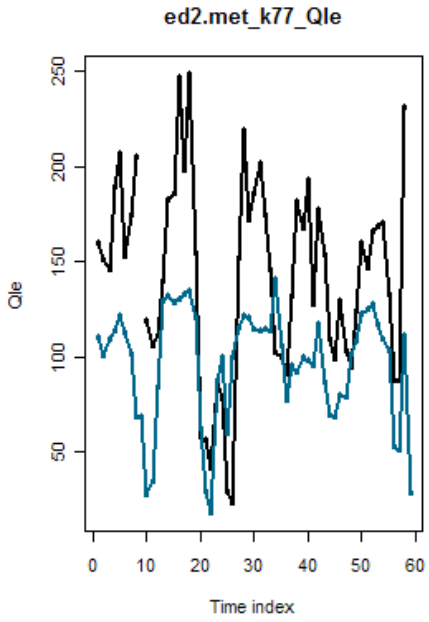
Monthly time series plots

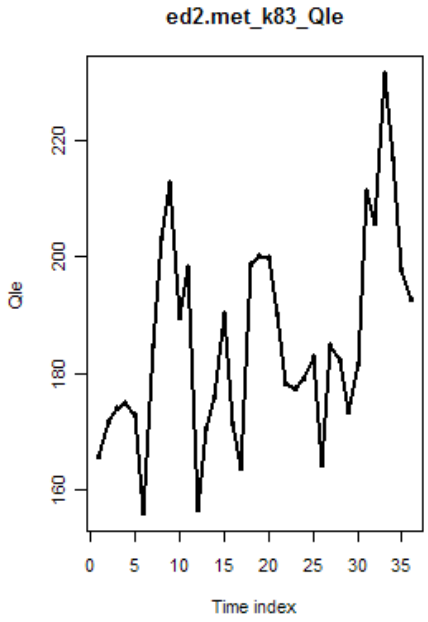

ed2.met_ban_Qle
Qle
Time index

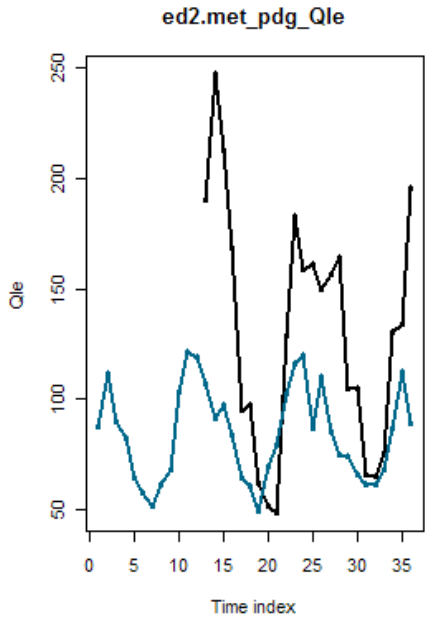

ed2.met_fns_Qle
Qle
Time index

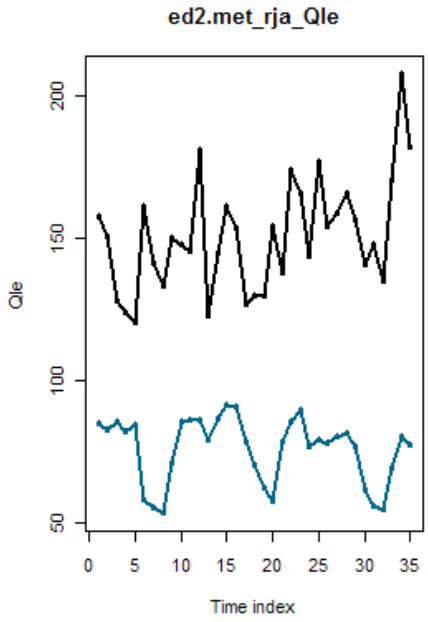

ed2.met_k34_Qle
Qle
Time index

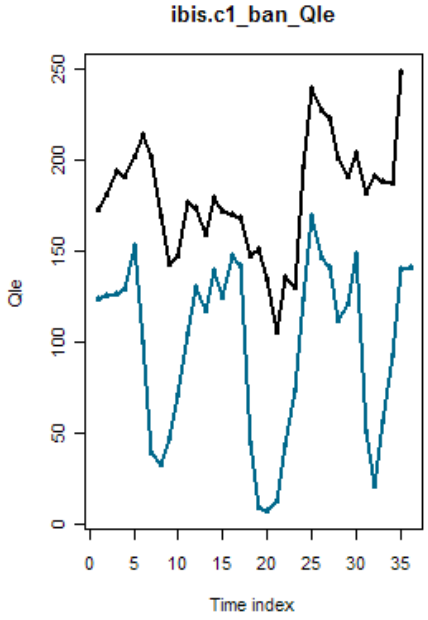

ed2.met_k67_Qle
Qle
Time index

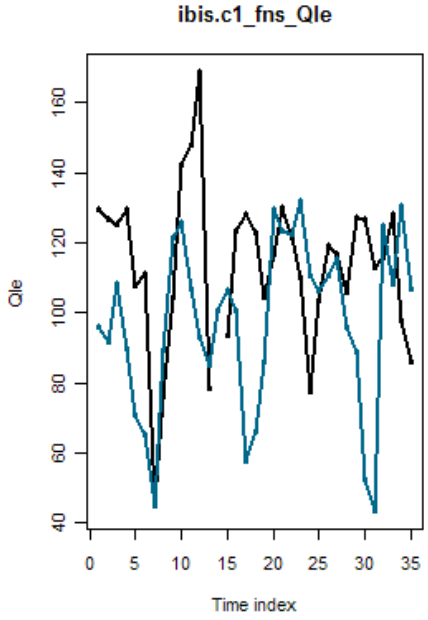

ed2.met_k77_Qle
Qle
Time index

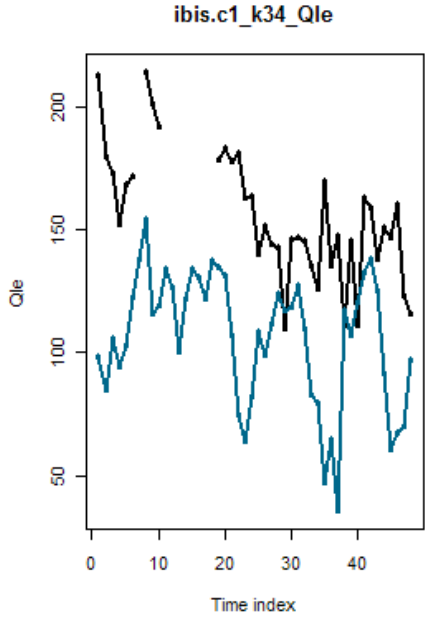

ed2.met_k83_Qle
Qle
Time index

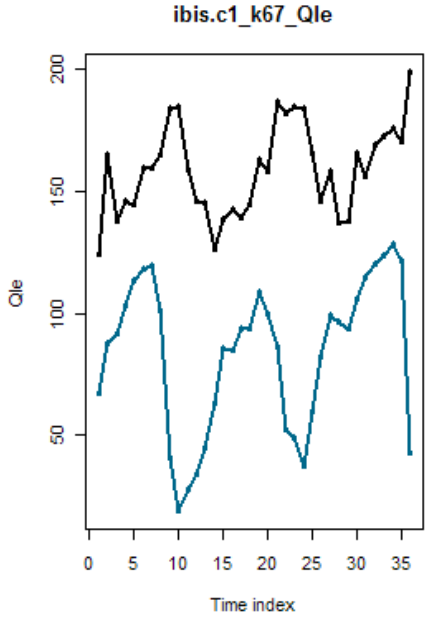

ed2.met_pdg_Qle
Qle
Time index

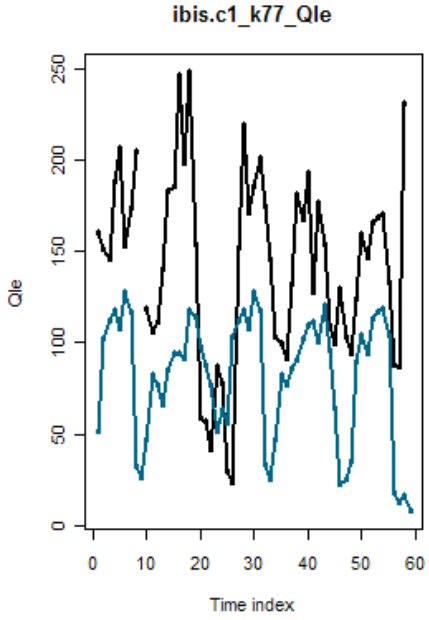

ed2.met_rja_Qle
Qle
Time index

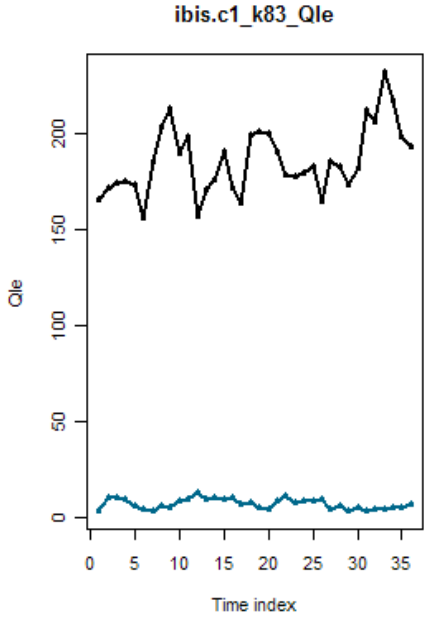

ibis.c1_ban_Qle
Qle
Time index

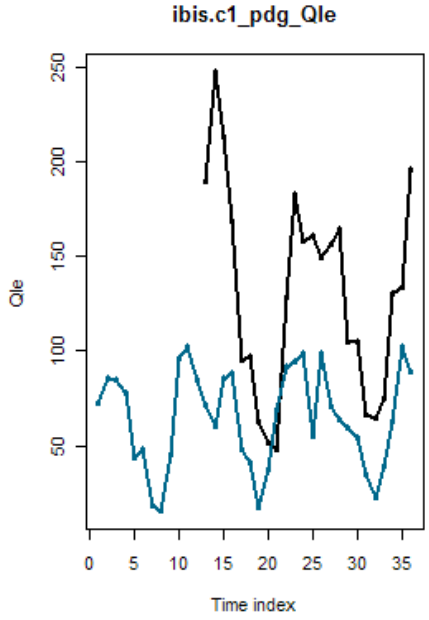

ibis.c1_fns_Qle
Qle
Time index

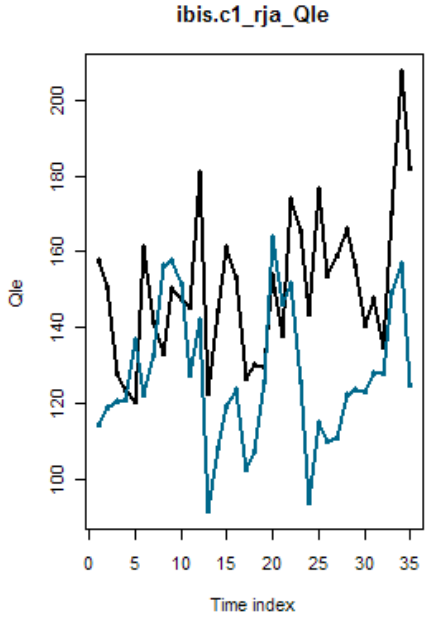

ibis.c1_k34_Qle
Qle
Time index

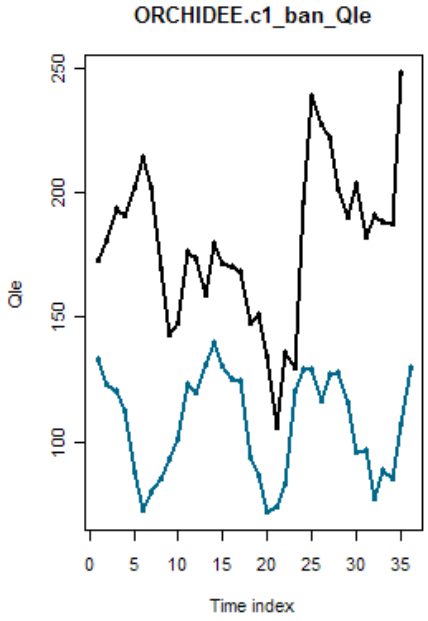

ibis.c1_k67_Qle
Qle
Time index

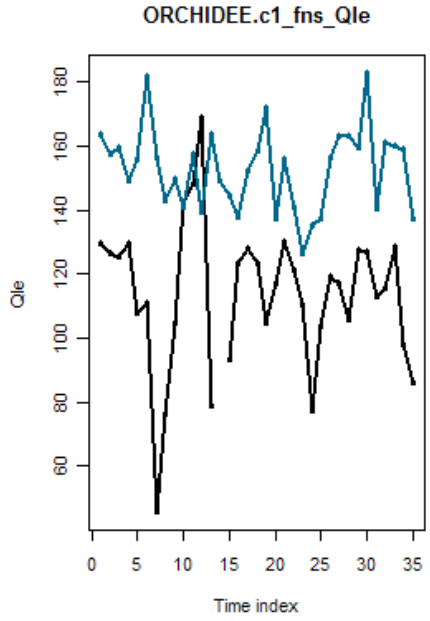

ibis.c1_k77_Qle
Qle
Time index

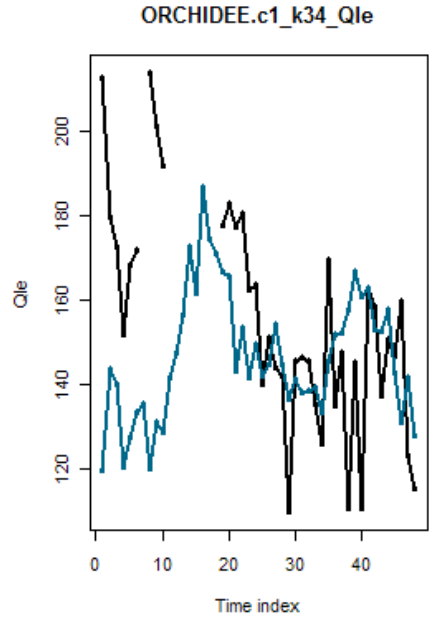

ibis.c1_k83_Qle
Qle
Time index

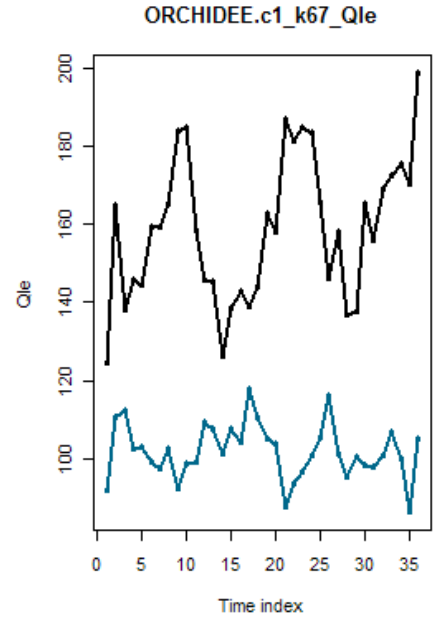

ibis.c1_pdg_Qle
Qle
Time index

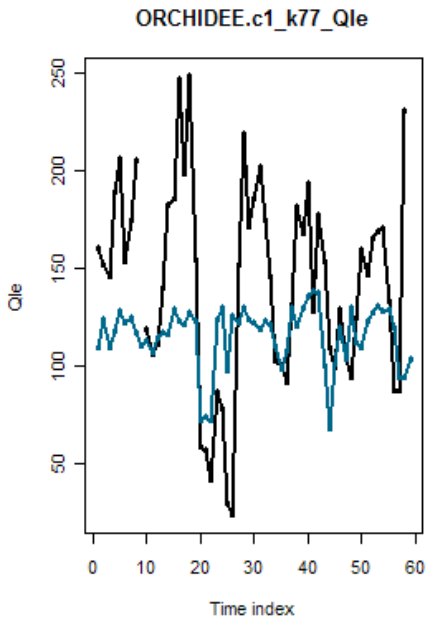

ibis.c1_rja_Qle
Qle
Time index

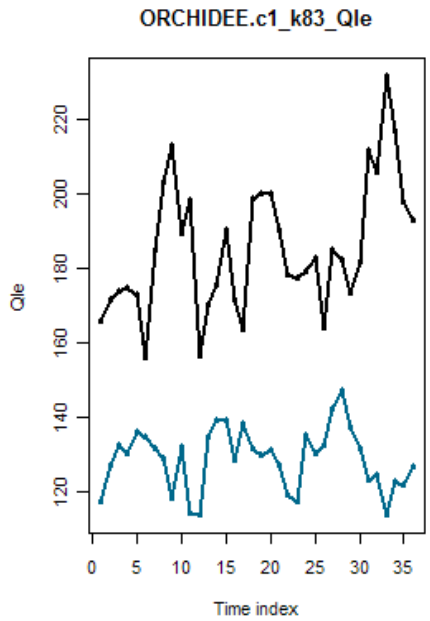

ORCHIDEE.c1_ban_Qle
Qle
Time index

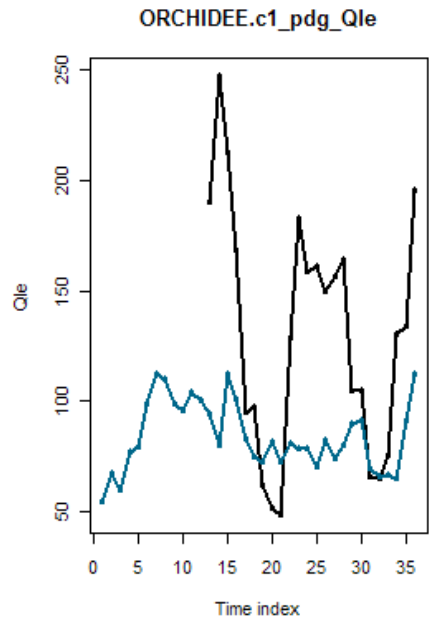

ORCHIDEE.c1_fns_Qle
Qle
Time index

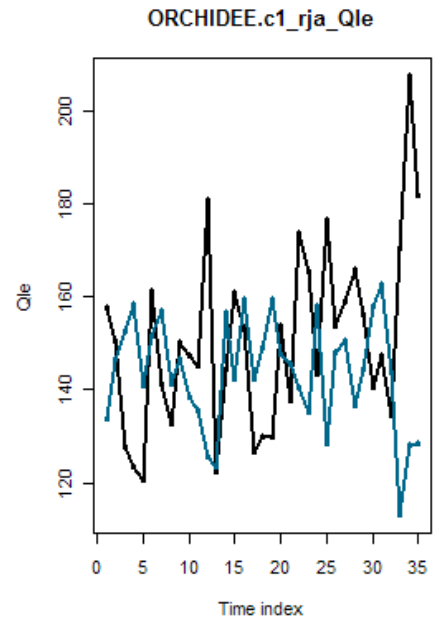

ORCHIDEE.c1_k34_Qle
Qle
Time index

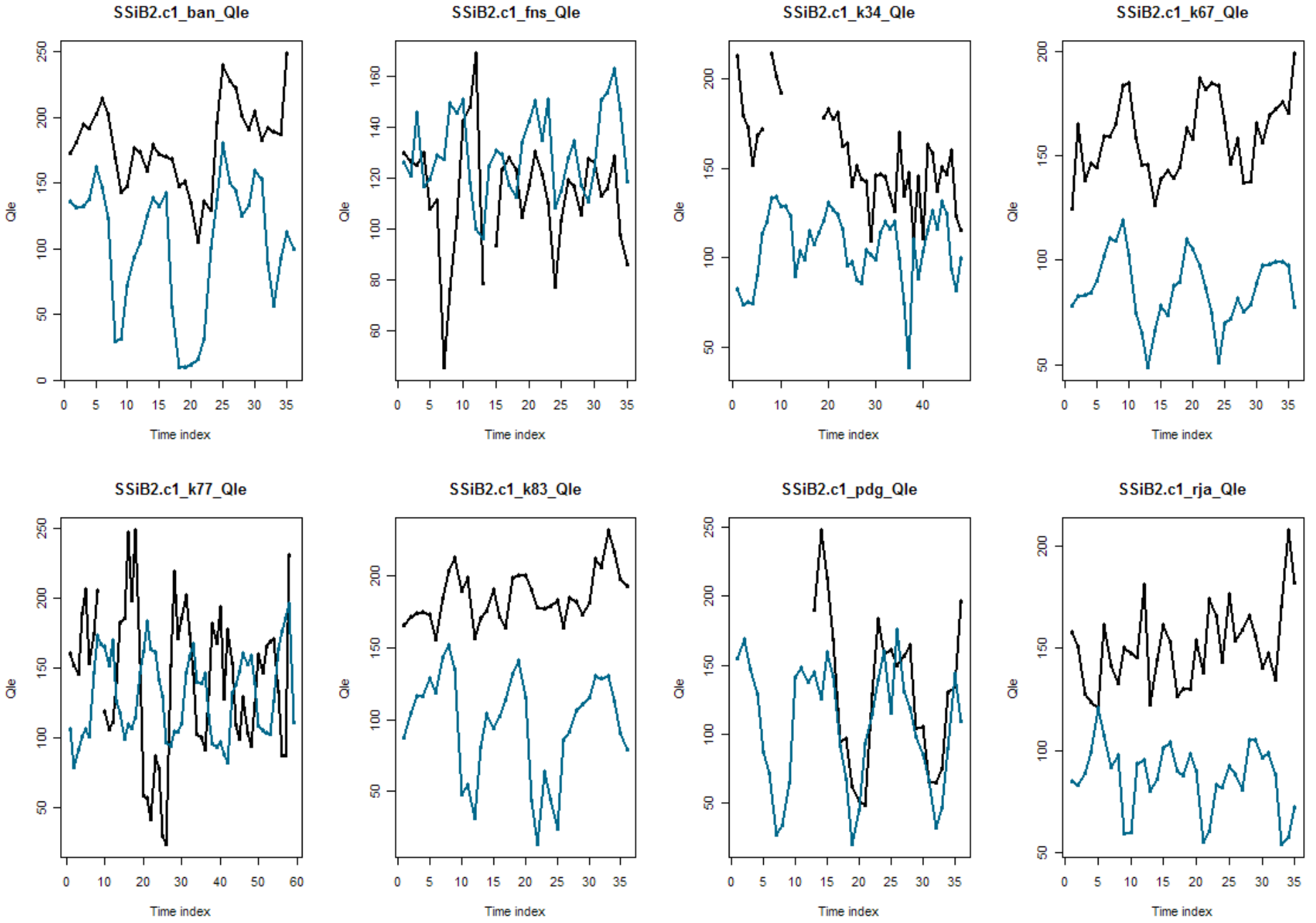

ORCHIDEE.c1_k67_Qle
Qle
Time index

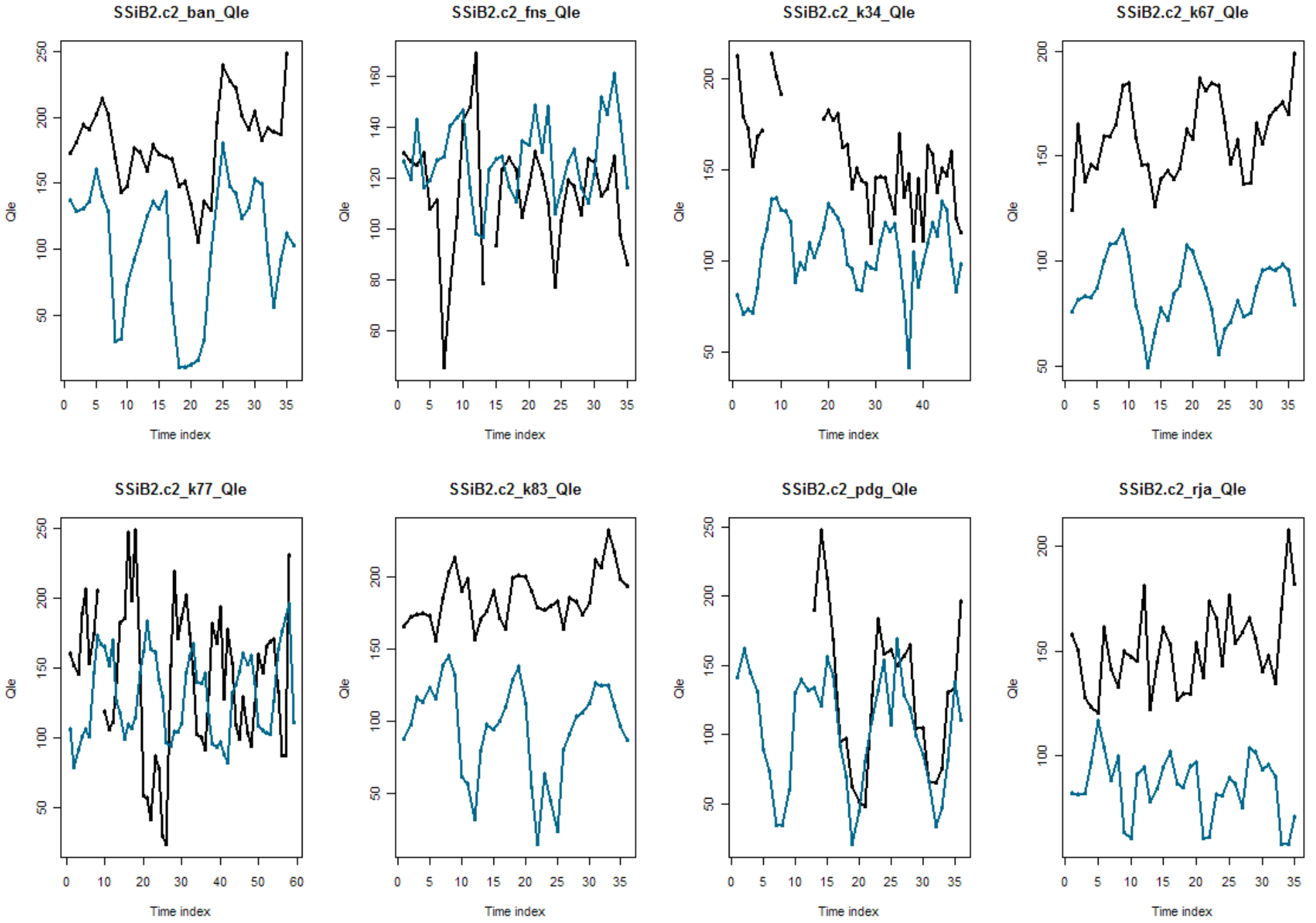

ORCHIDEE.c1_k77_Qle
Qle
Time index

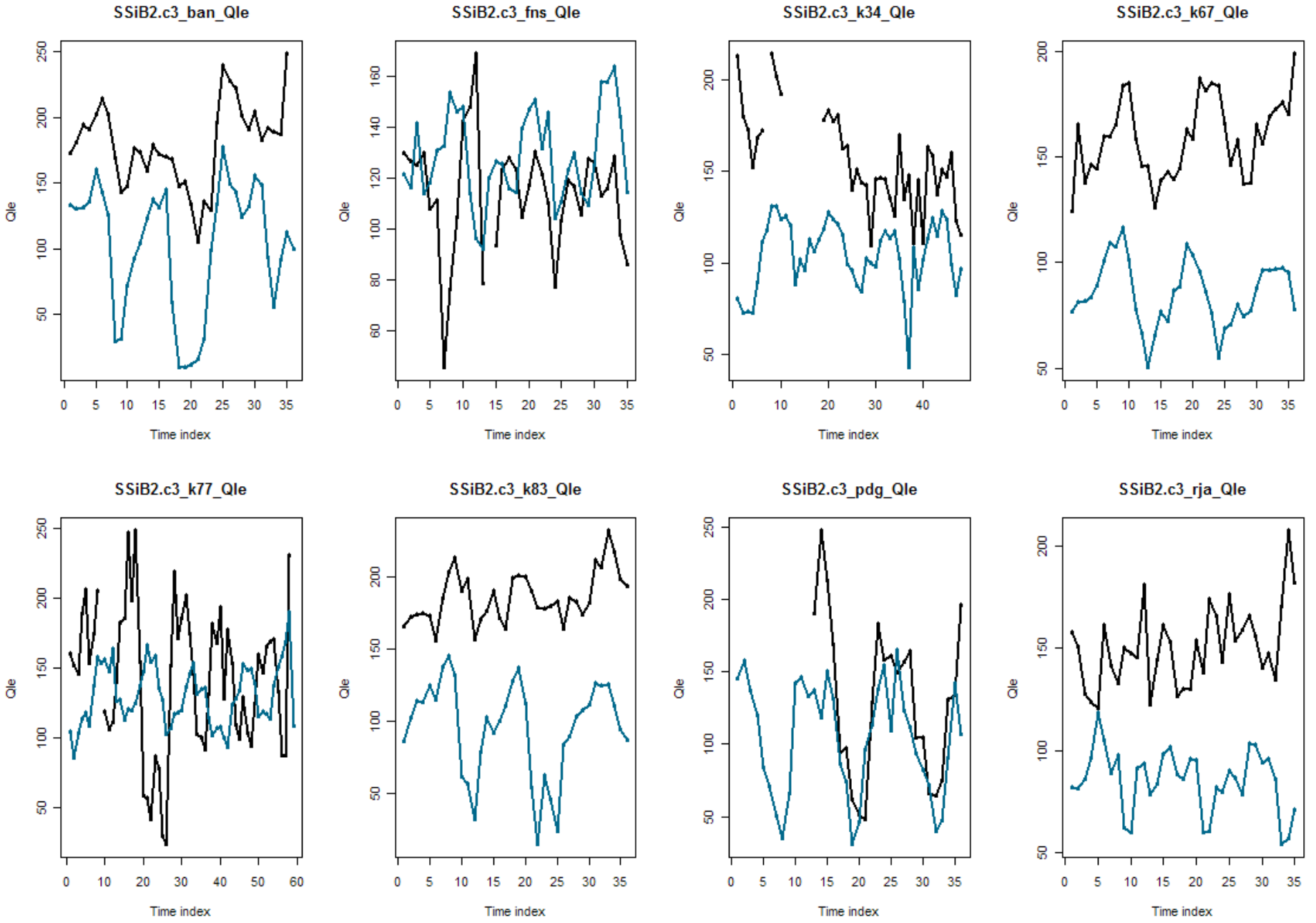

ORCHIDEE.c1_k83_Qle
Qle
Time index

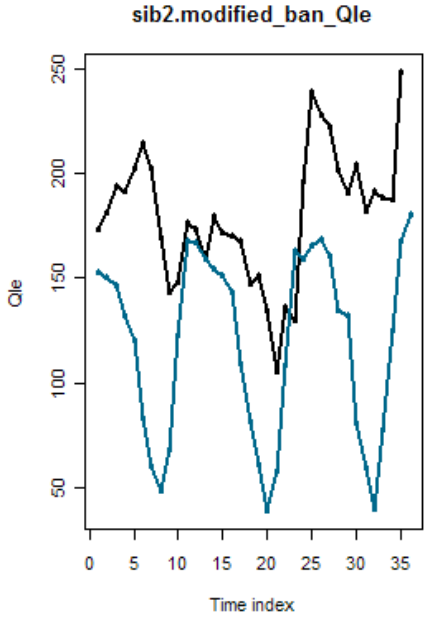

ORCHIDEE.c1_pdg_Qle
Qle
Time index

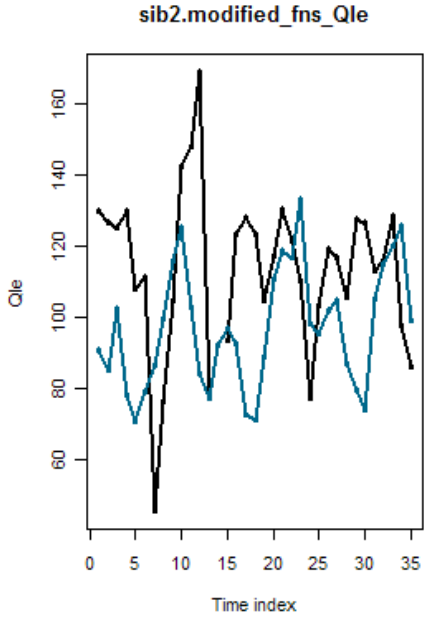

ORCHIDEE.c1_rja_Qle
Qle
Time index

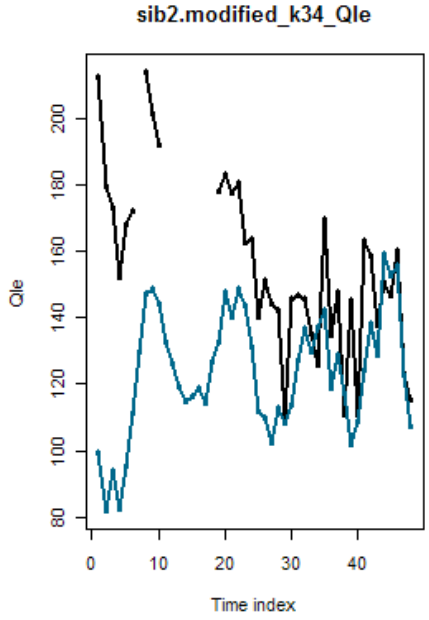

SSiB2.c1_ban_Qle
SSiB2.c1_fns_Qle
SSiB2.c1_k34_Qle
SSiB2.c1_k67_Qle
SSiB2.c1_k77_Qle
SSiB2.c1_k83_Qle
SSiB2.c1_pdg_Qle
SSiB2.c1_rja_Qle
Qle
Time index

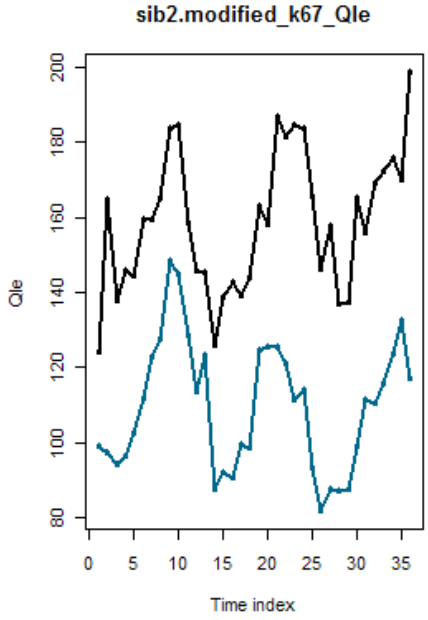

SSiB2.c2_ban_Qle
SSiB2.c2_fns_Qle
SSiB2.c2_k34_Qle
SSiB2.c2_k67_Qle
SSiB2.c2_k77_Qle
SSiB2.c2_k83_Qle
SSiB2.c2_pdg_Qle
SSiB2.c2_rja_Qle
Qle
Time index

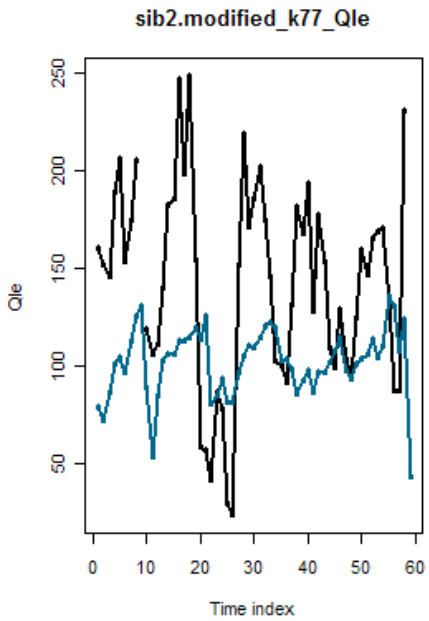

SSiB2.c3_ban_Qle
SSiB2.c3_fns_Qle
SSiB2.c3_k34_Qle
SSiB2.c3_k67_Qle
SSiB2.c3_k77_Qle
SSiB2.c3_k83_Qle
SSiB2.c3_pdg_Qle
SSiB2.c3_rja_Qle
Qle
Time index

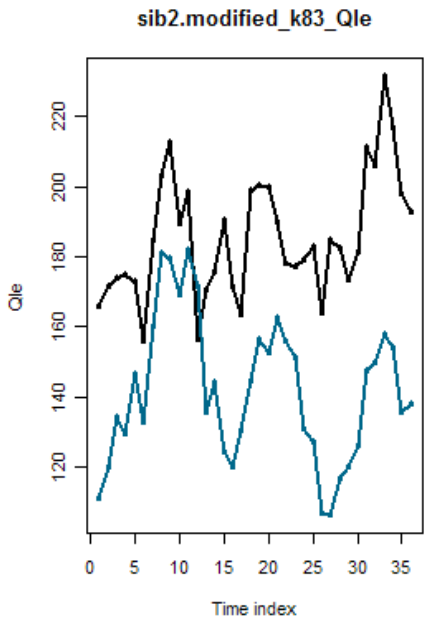

sib2.modified_ban_Qle
Qle
Time index

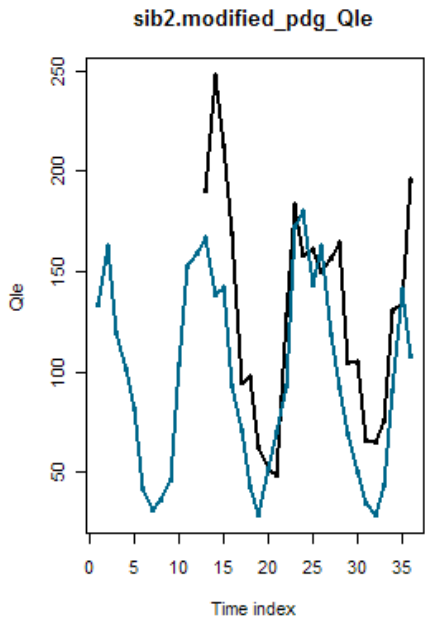

sib2.modified_fns_Qle
Qle
Time index

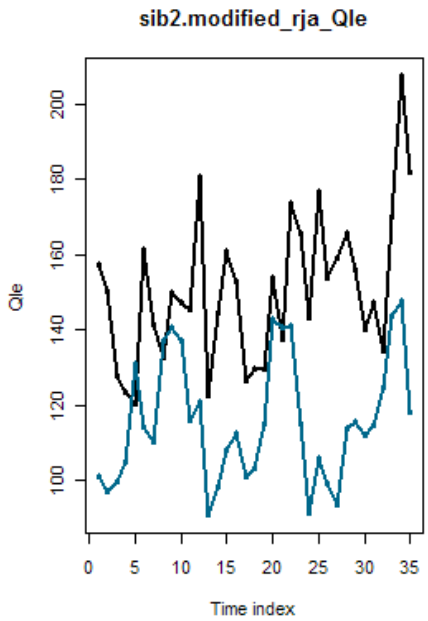

sib2.modified_k34_Qle
Qle
Time index

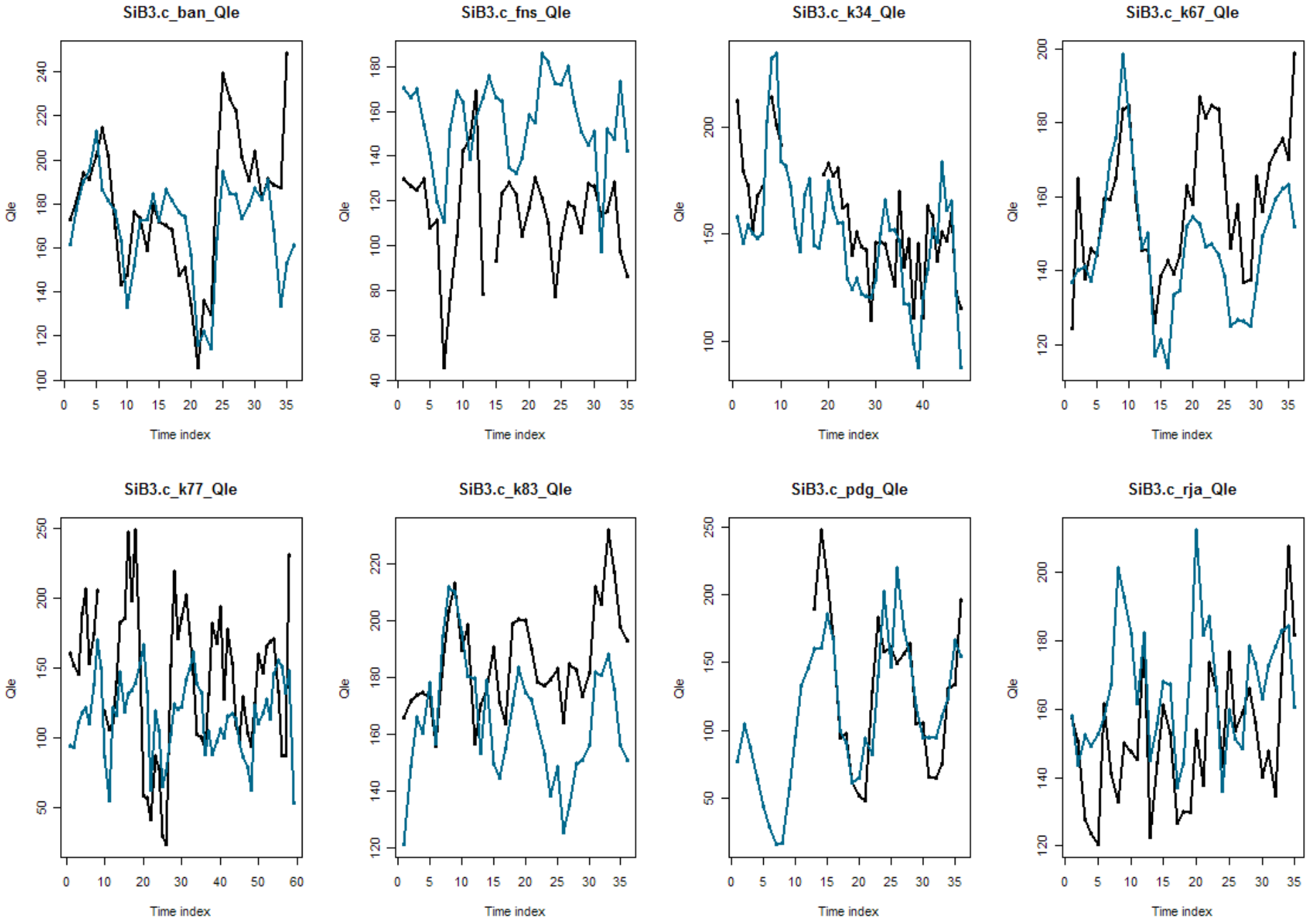

sib2.modified_k67_Qle
Qle
Time index

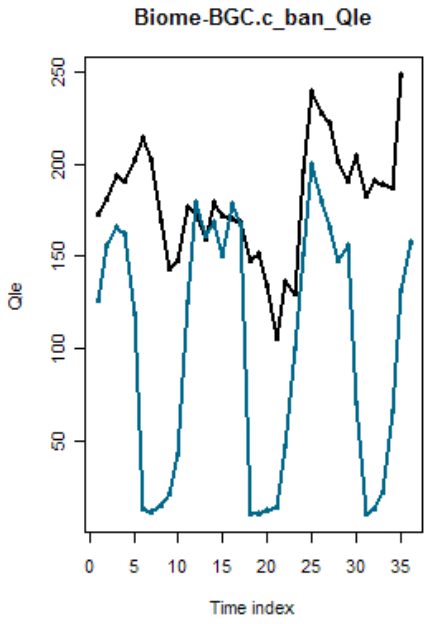

sib2.modified_k77_Qle
Qle
Time index

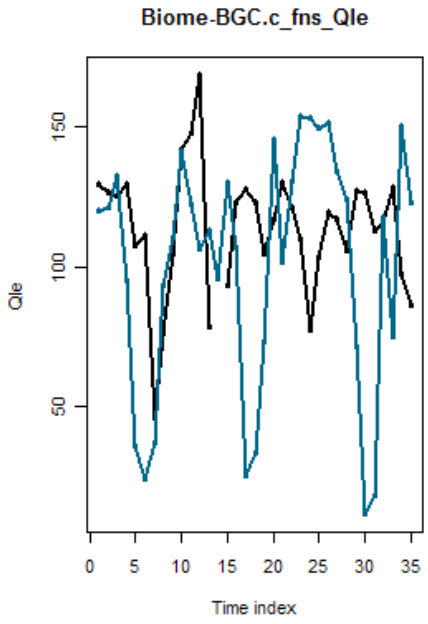

sib2.modified_k83_Qle
Qle
Time index

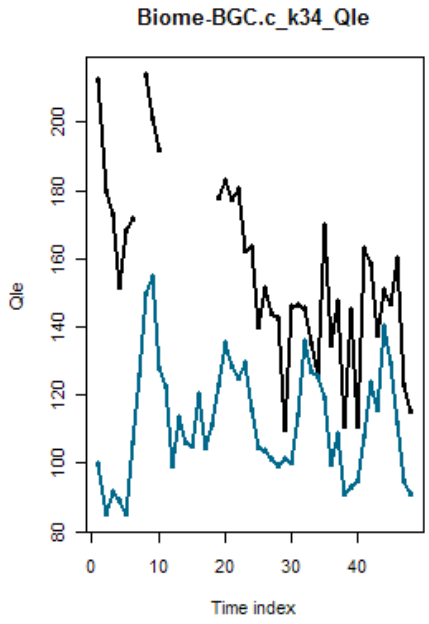

sib2.modified_pdg_Qle
Qle
Time index

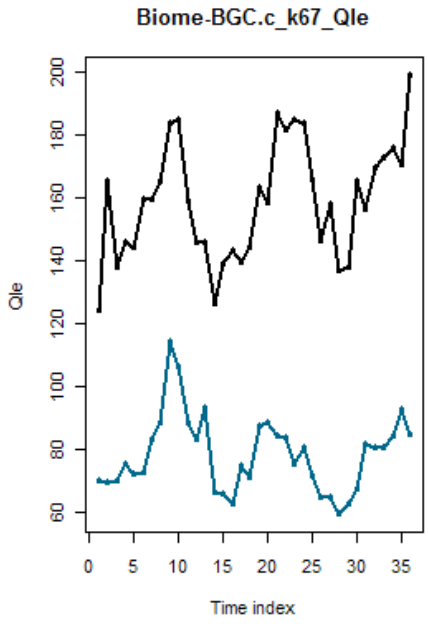

sib2.modified_rja_Qle
Qle
Time index

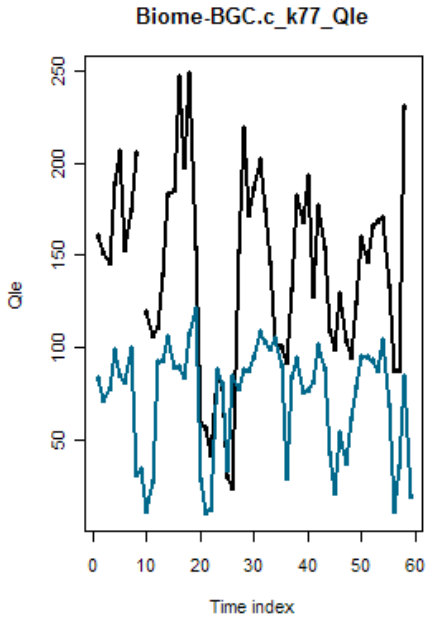

SiB3.c_ban_Qle
SiB3.c_fns_Qle
SiB3.c_k34_Qle
SiB3.c_k67_Qle
SiB3.c_k77_Qle
SiB3.c_k83_Qle
SiB3.c_pdg_Qle
SiB3.c_rja_Qle
Qle
Time index

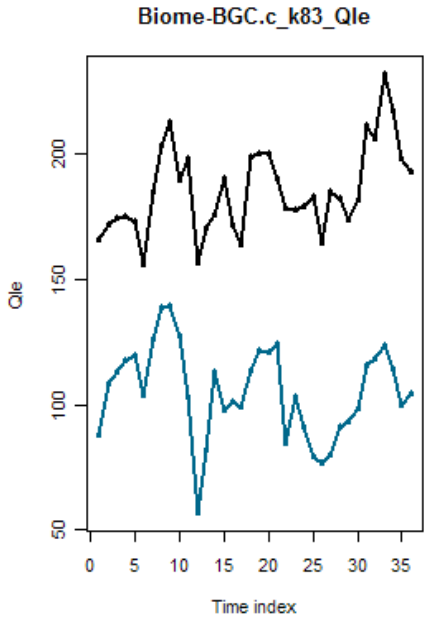

Biome-BGC.c_ban_Qle
Qle
Time index

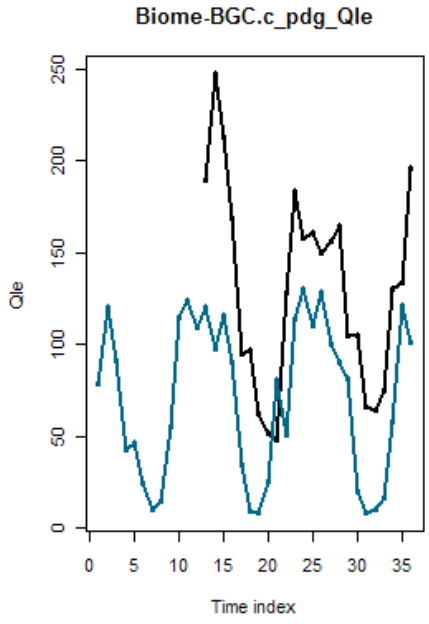

Biome-BGC.c_fns_Qle
Qle
Time index

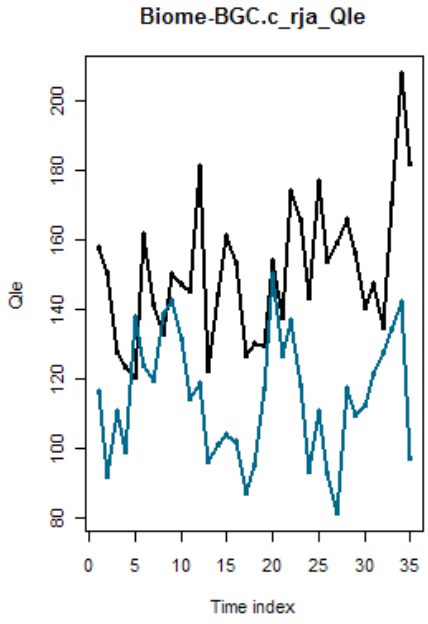

Biome-BGC.c_k34_Qle
Qle
Time index

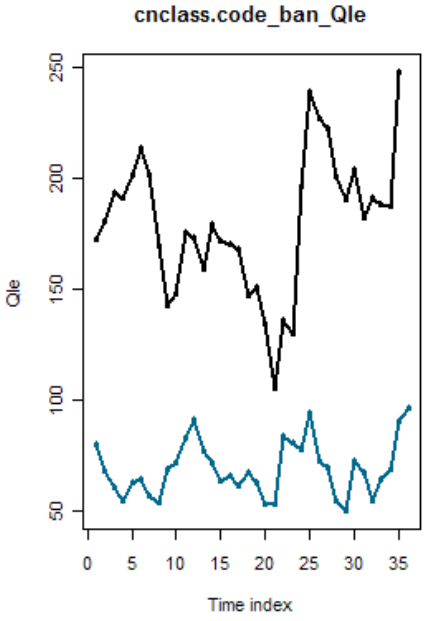

Biome-BGC.c_k67_Qle
Qle
Time index

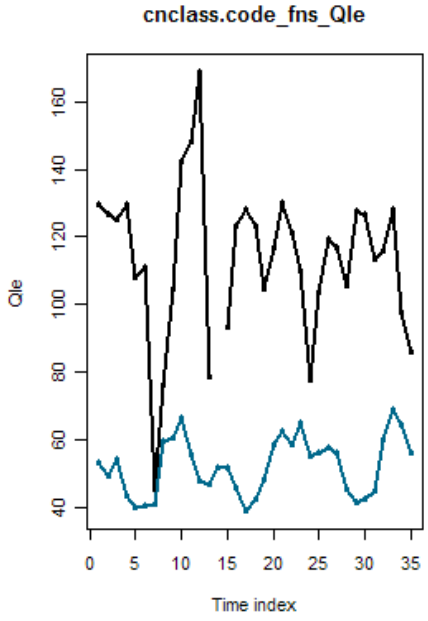

Biome-BGC.c_k77_Qle
Qle
Time index

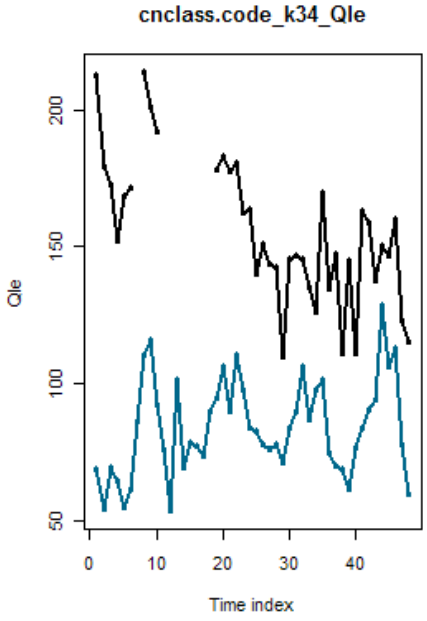

Biome-BGC.c_k83_Qle
Qle
Time index

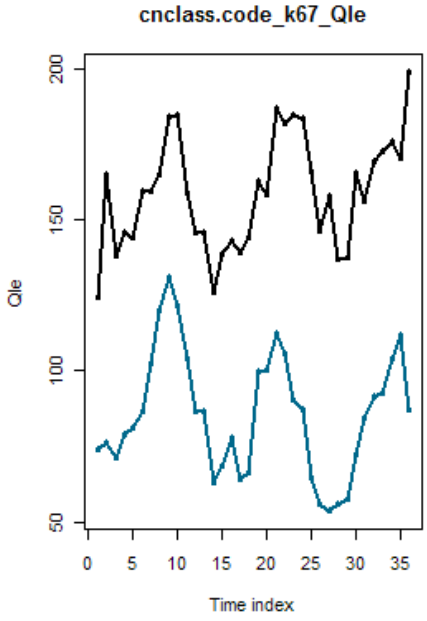

Biome-BGC.c_pdg_Qle
Qle
Time index

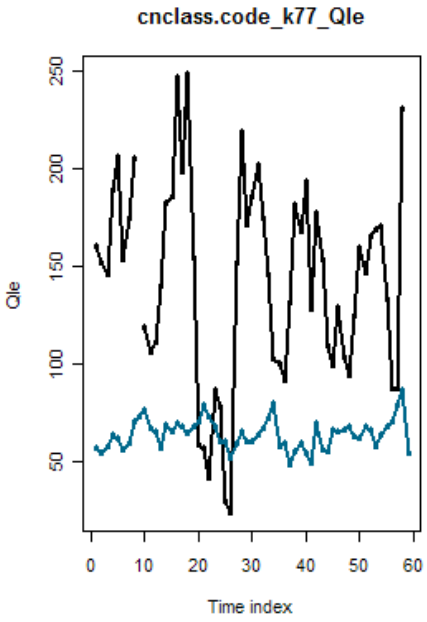

Biome-BGC.c_rja_Qle
Qle
Time index

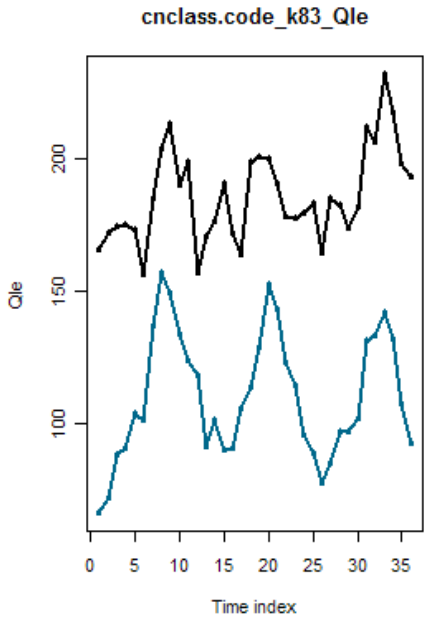

cnclass.code_ban_Qle
Qle
Time index

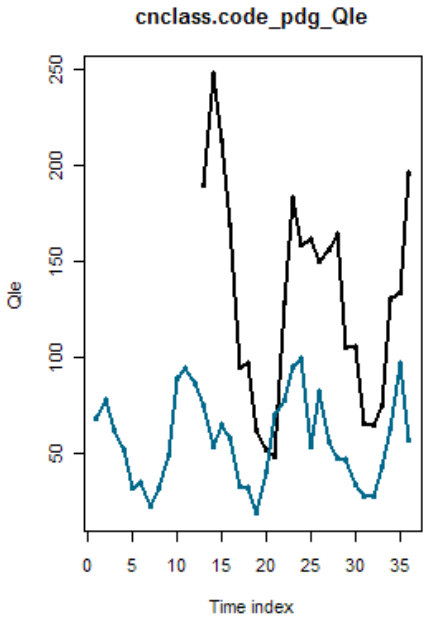

cnclass.code_fns_Qle
Qle
Time index

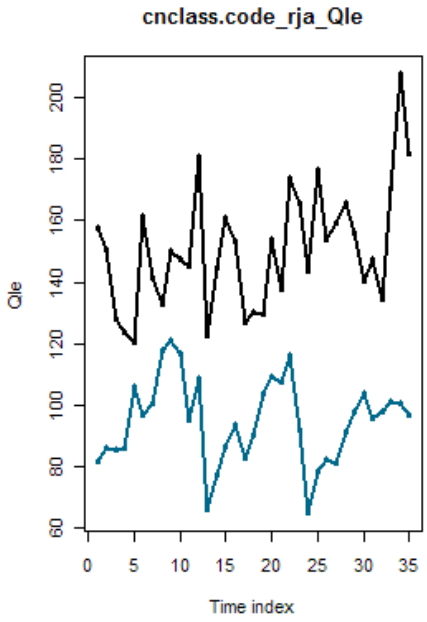

cnclass.code_k34_Qle
Qle
Time index

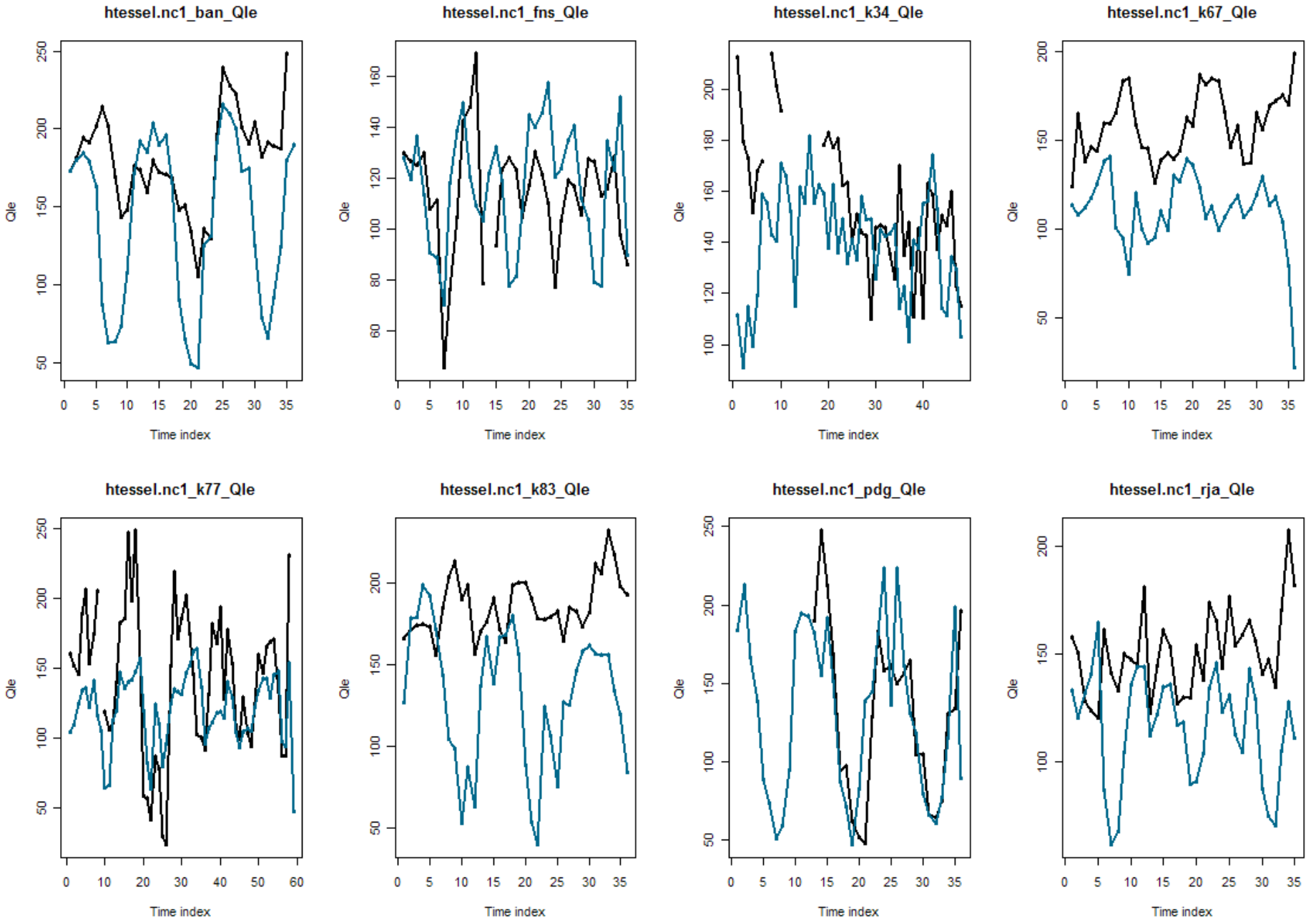

cnclass.code_k67_Qle
Qle
Time index

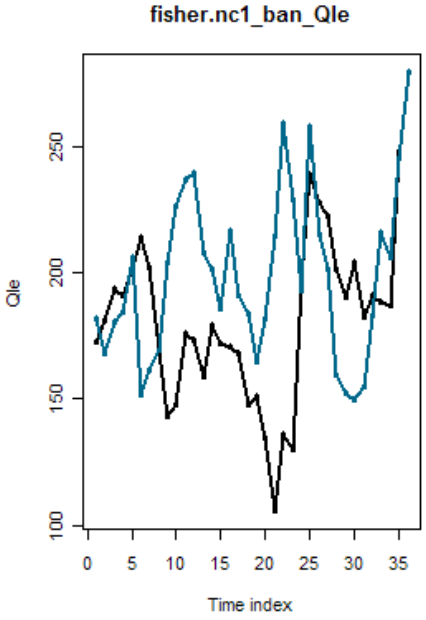

cnclass.code_k77_Qle
Qle
Time index

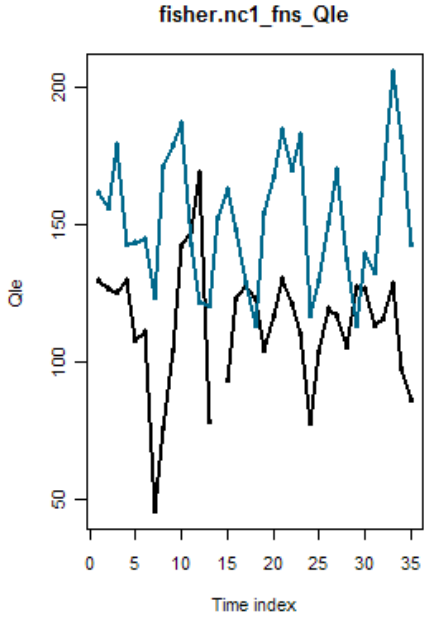

cnclass.code_k83_Qle
Qle
Time index

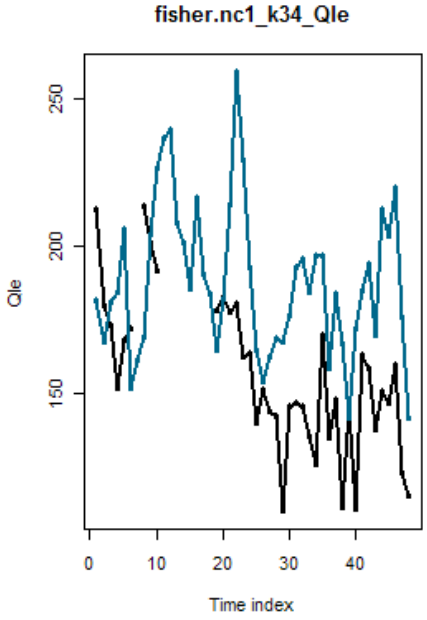

cnclass.code_pdg_Qle
Qle
Time index

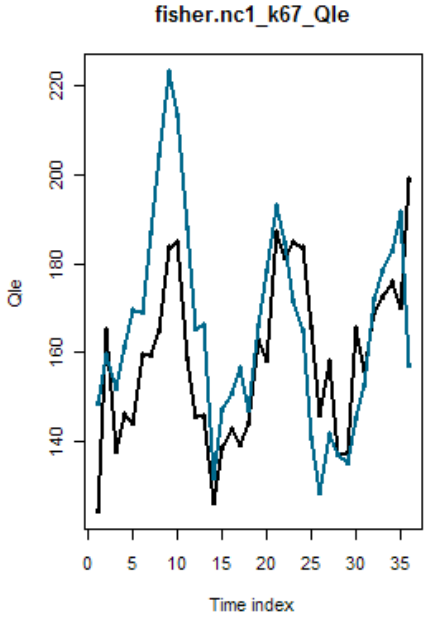

cnclass.code_rja_Qle
Qle
Time index

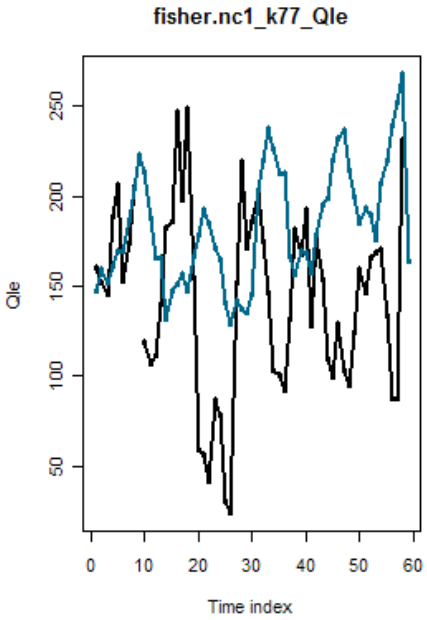

htessel.nc1_ban_Qle
htessel.nc1_fns_Qle
htessel.nc1_k34_Qle
htessel.nc1_k67_Qle
htessel.nc1_k77_Qle
htessel.nc1_k83_Qle
htessel.nc1_pdg_Qle
htessel.nc1_rja_Qle
Qle
Time index

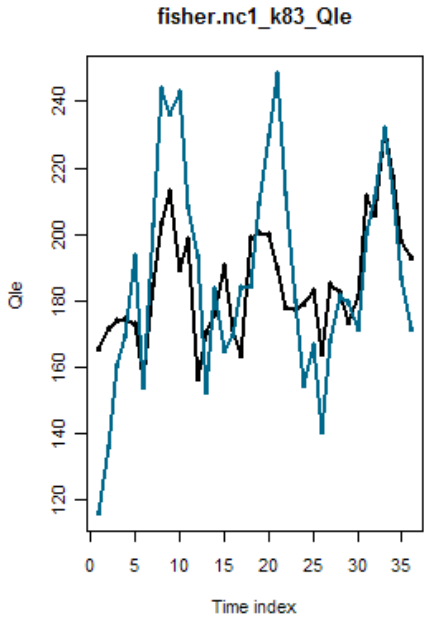

fisher.nc1_ban_Qle
Qle
Time index

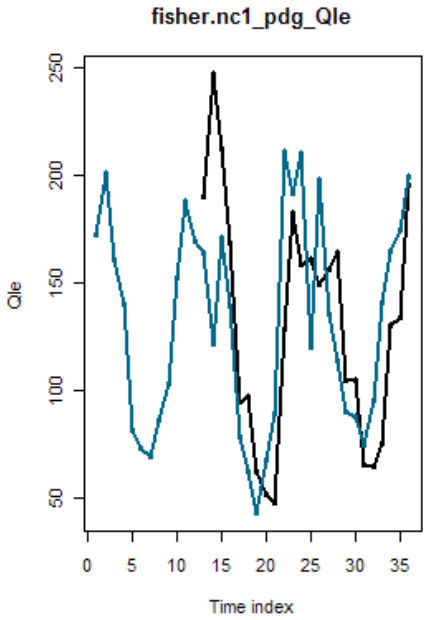

fisher.nc1_fns_Qle
Qle
Time index

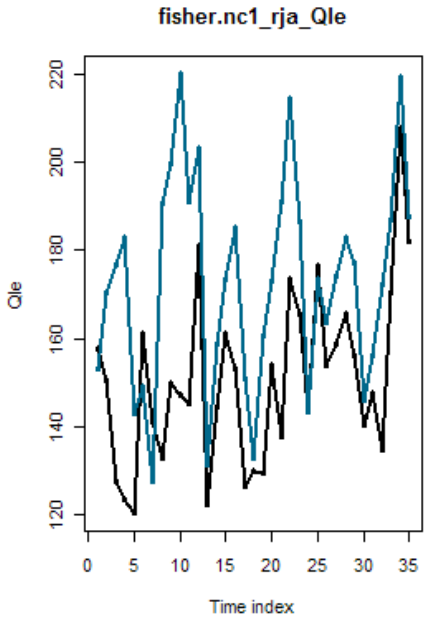

fisher.nc1_k34_Qle
Qle
Time index

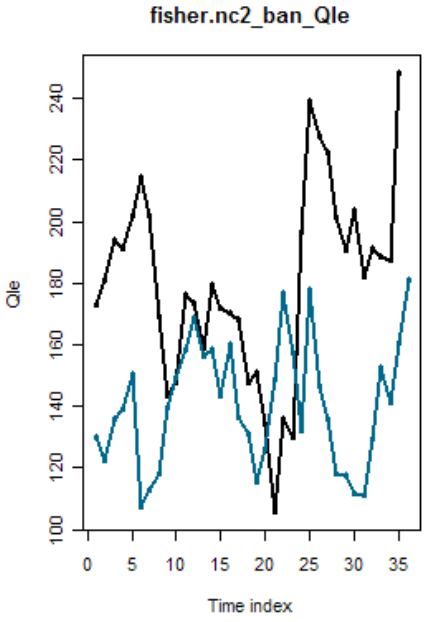

fisher.nc1_k67_Qle
Qle
Time index

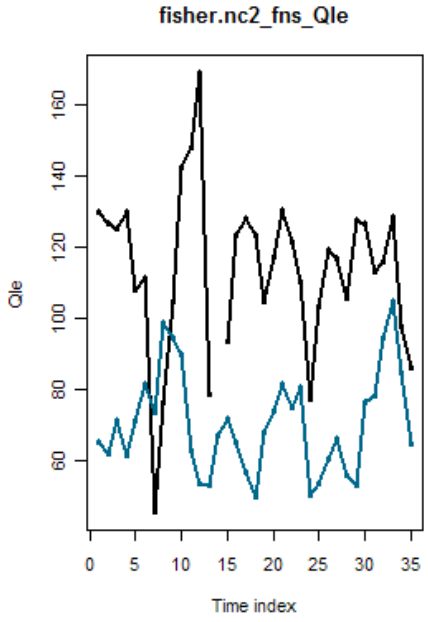

fisher.nc1_k77_Qle
Qle
Time index

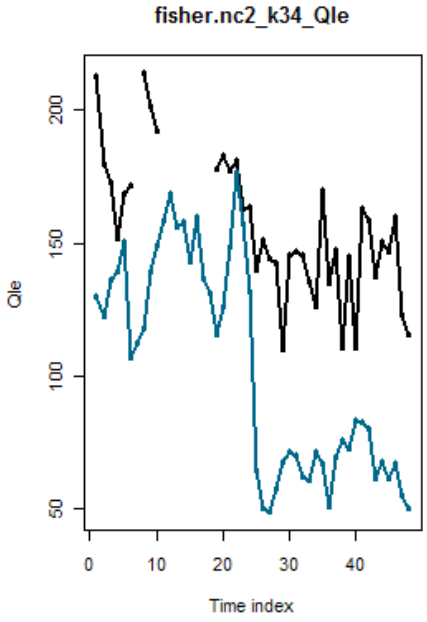

fisher.nc1_k83_Qle
Qle
Time index

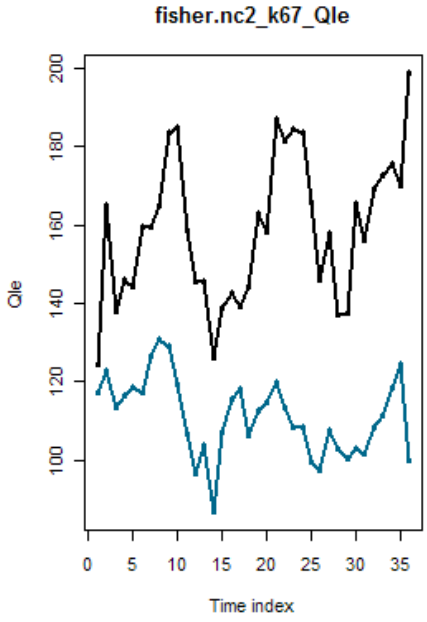

fisher.nc1_pdg_Qle
Qle
Time index

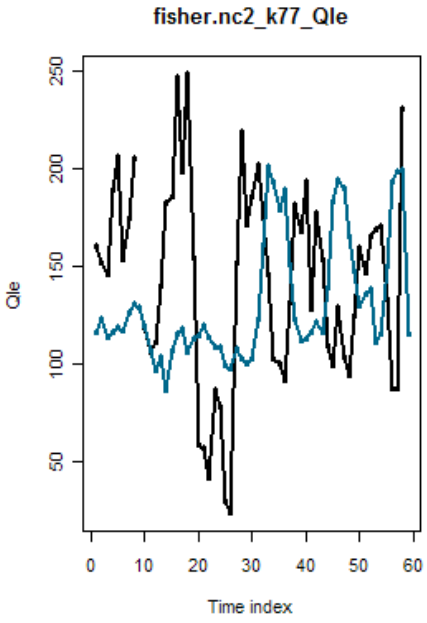

fisher.nc1_rja_Qle
Qle
Time index

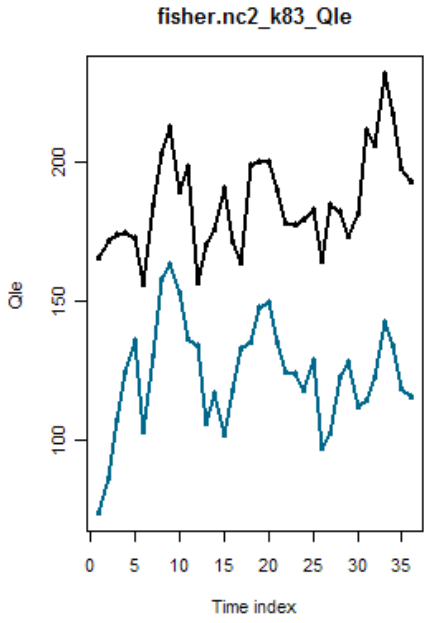

fisher.nc2_ban_Qle
Qle
Time index

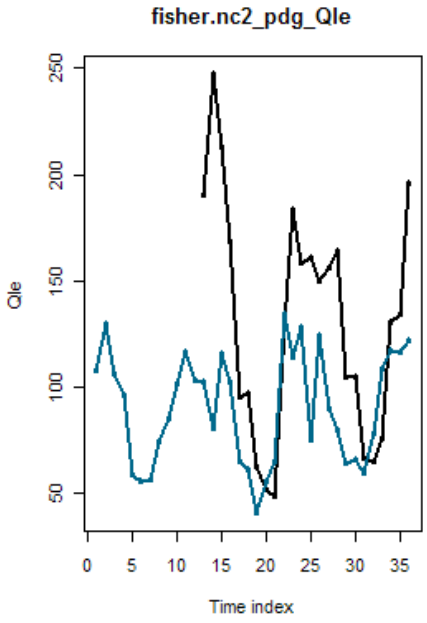

fisher.nc2_fns_Qle
Qle
Time index

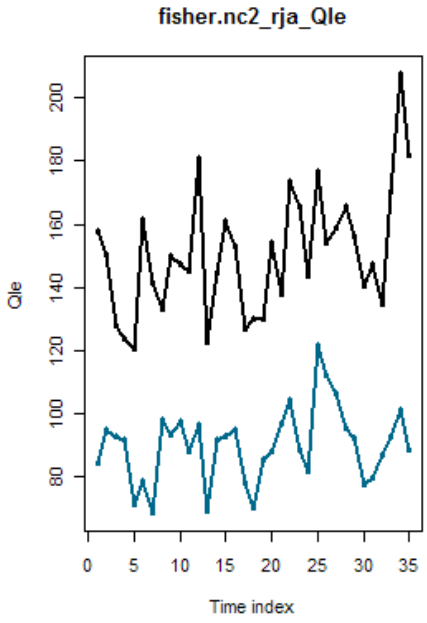

fisher.nc2_k34_Qle
Qle
Time index

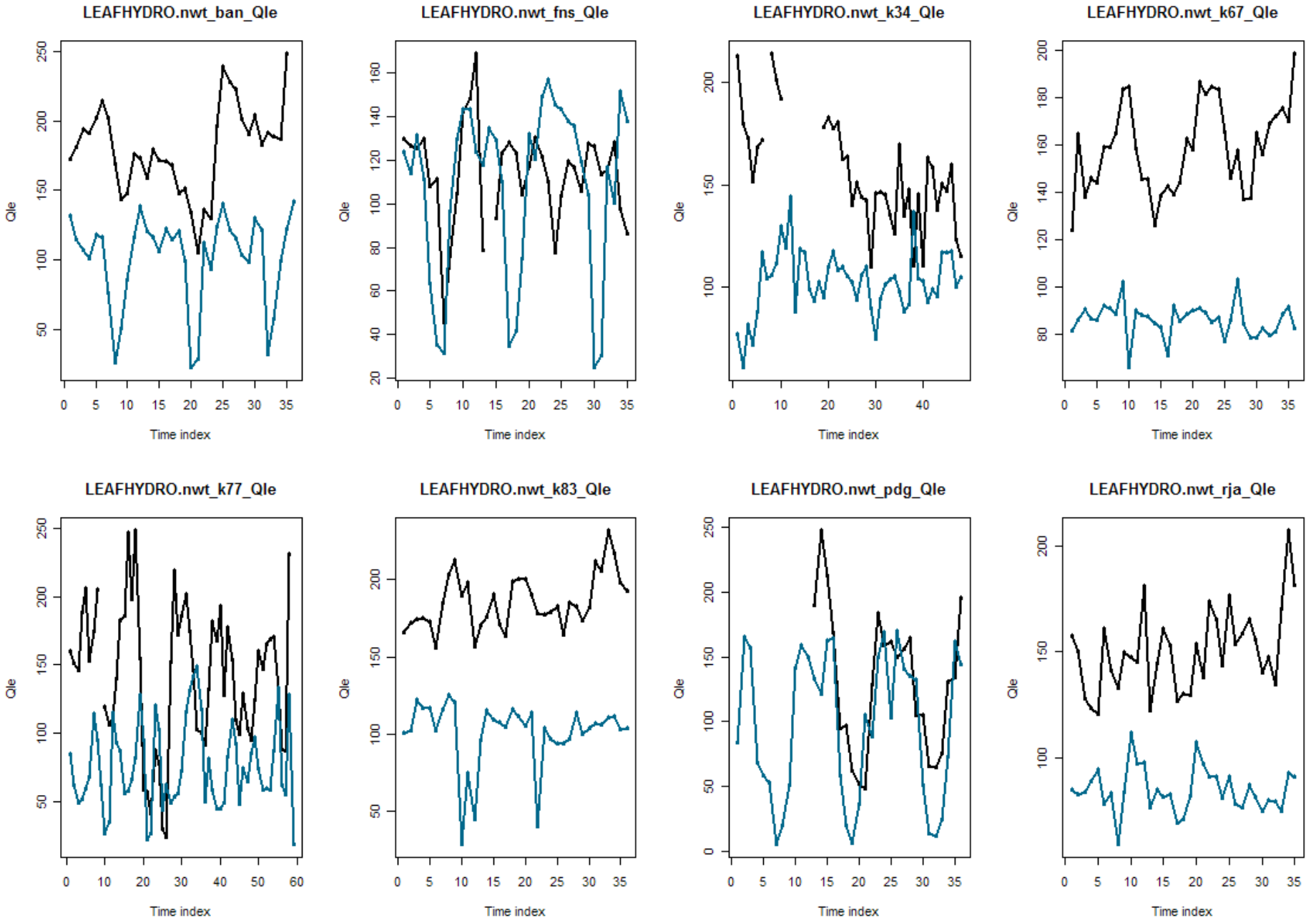

fisher.nc2_k67_Qle
Qle
Time index

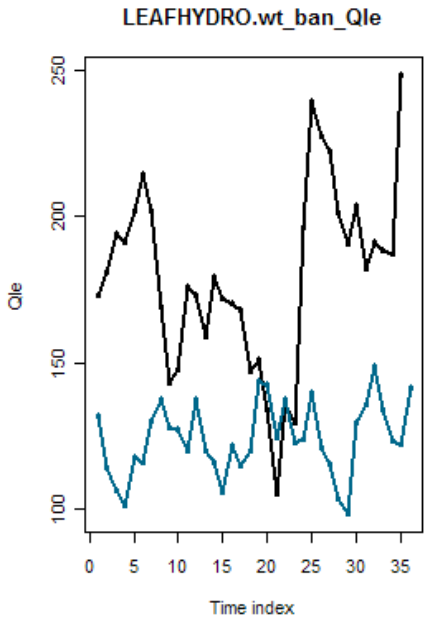

fisher.nc2_k77_Qle
Qle
Time index

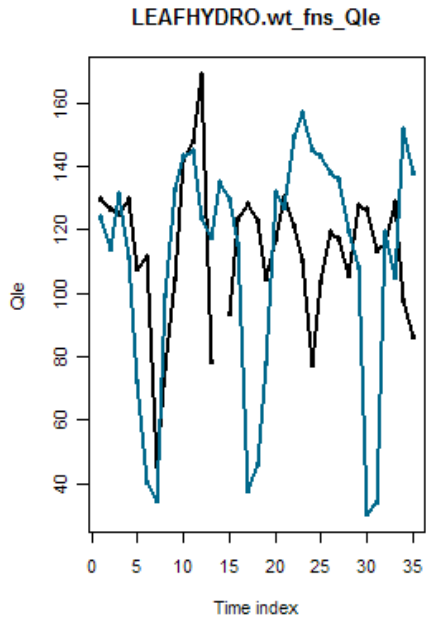

fisher.nc2_k83_Qle
Qle
Time index

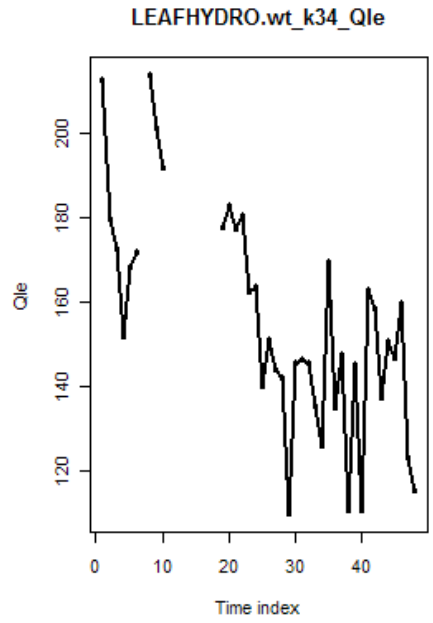

fisher.nc2_pdg_Qle
Qle
Time index

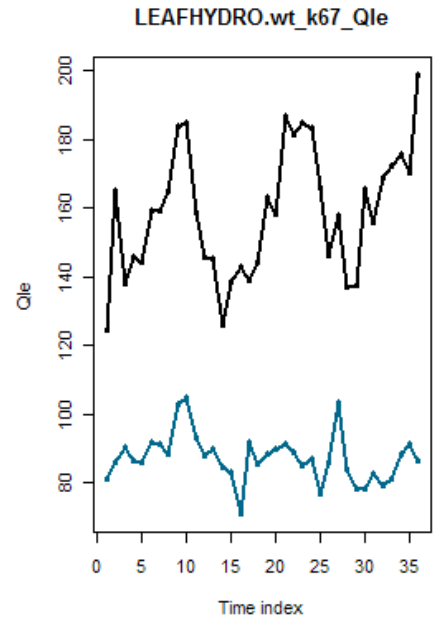

fisher.nc2_rja_Qle
Qle
Time index

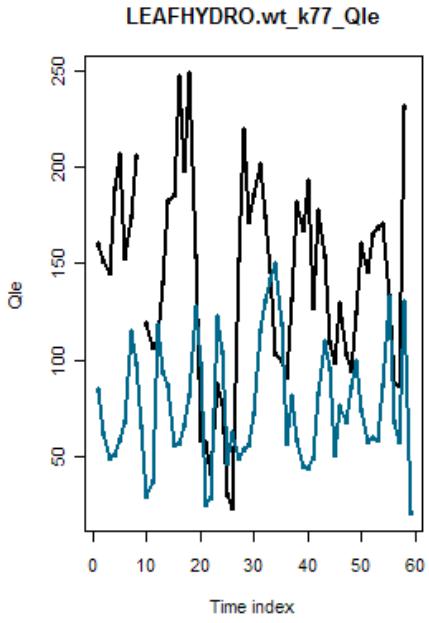

LEAFHYDRO.nwt_ban_Qle
LEAFHYDRO.nwt_fns_Qle
LEAFHYDRO.nwt_k34_Qle
LEAFHYDRO.nwt_k67_Qle
LEAFHYDRO.nwt_k77_Qle
LEAFHYDRO.nwt_k83_Qle
LEAFHYDRO.nwt_pdg_Qle
LEAFHYDRO.nwt_rja_Qle
Qle
Time index

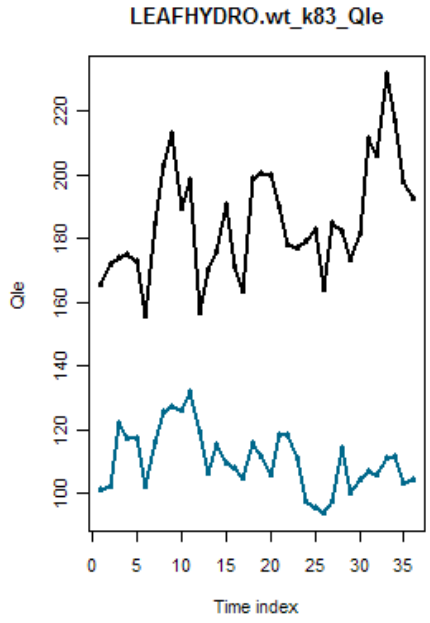

LEAFHYDRO.wt_ban_Qle
Qle
Time index

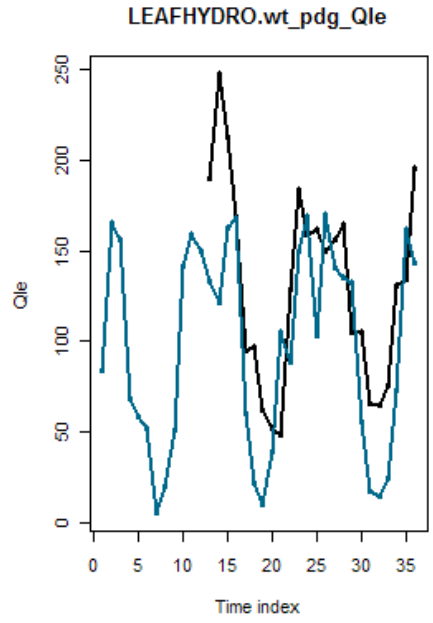

LEAFHYDRO.wt_fns_Qle
Qle
Time index

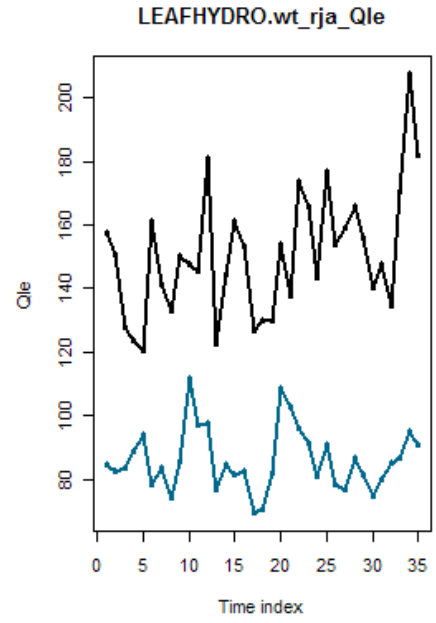

LEAFHYDRO.wt_k34_Qle
Qle
Time index

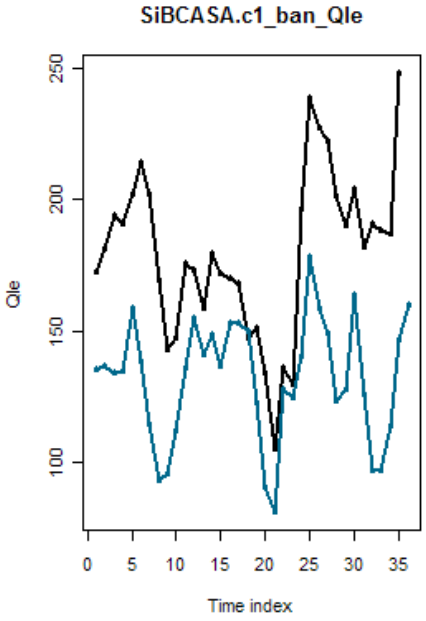

LEAFHYDRO.wt_k67_Qle
Qle
Time index

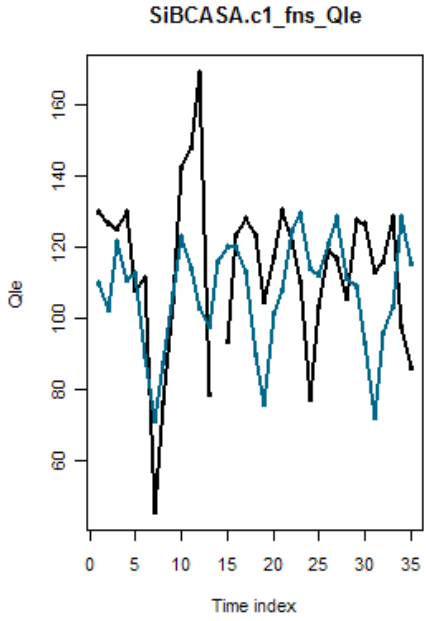

LEAFHYDRO.wt_k77_Qle
Qle
Time index

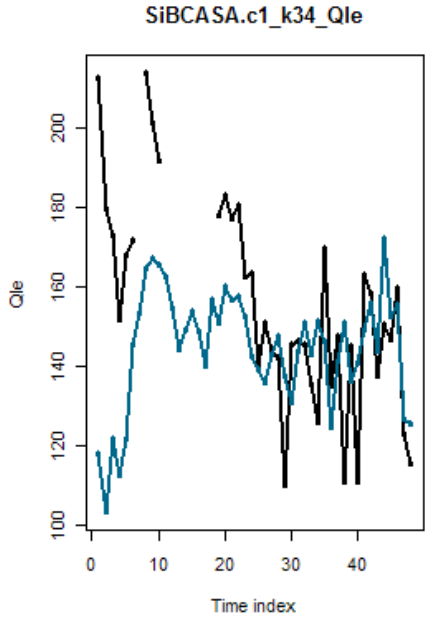

LEAFHYDRO.wt_k83_Qle
Qle
Time index

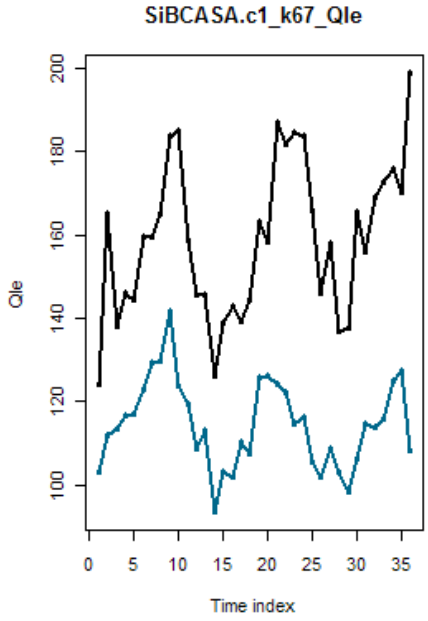

LEAFHYDRO.wt_pdg_Qle
Qle
Time index

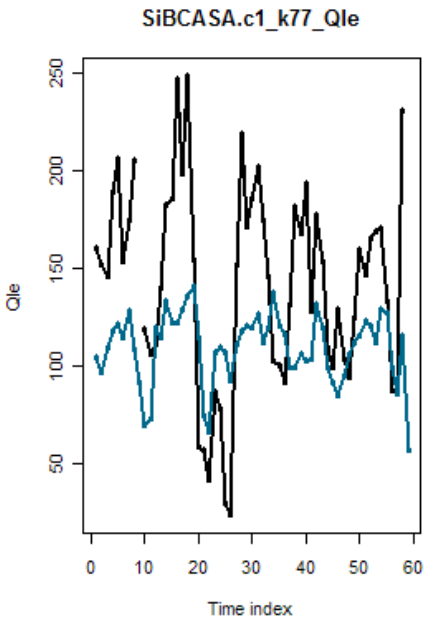

LEAFHYDRO.wt_rja_Qle
Qle
Time index

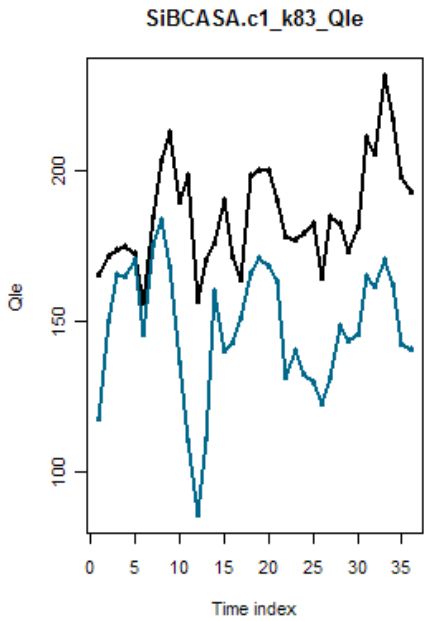

SiBCASA.c1_ban_Qle
Qle
Time index

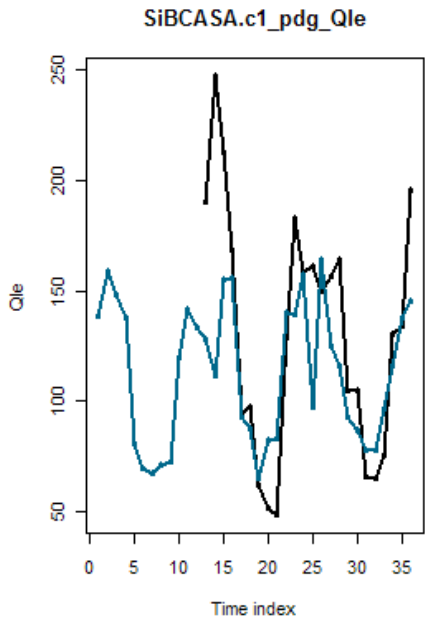

SiBCASA.c1_fns_Qle
Qle
Time index

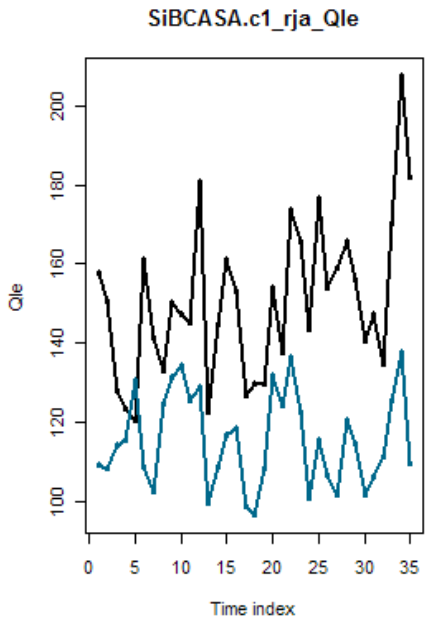

SiBCASA.c1_k34_Qle
Qle
Time index

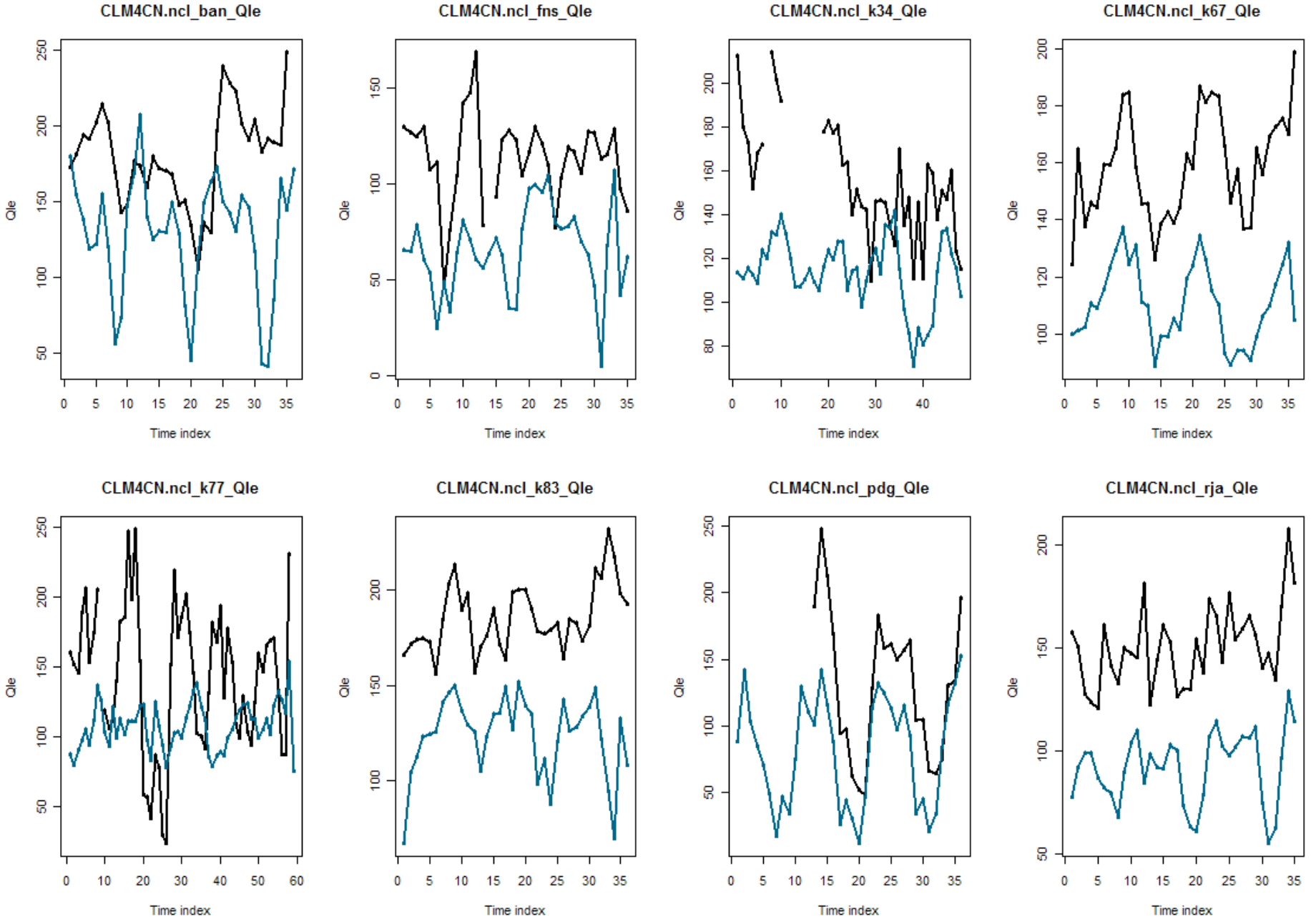

SiBCASA.c1_k67_Qle
Qle
Time index

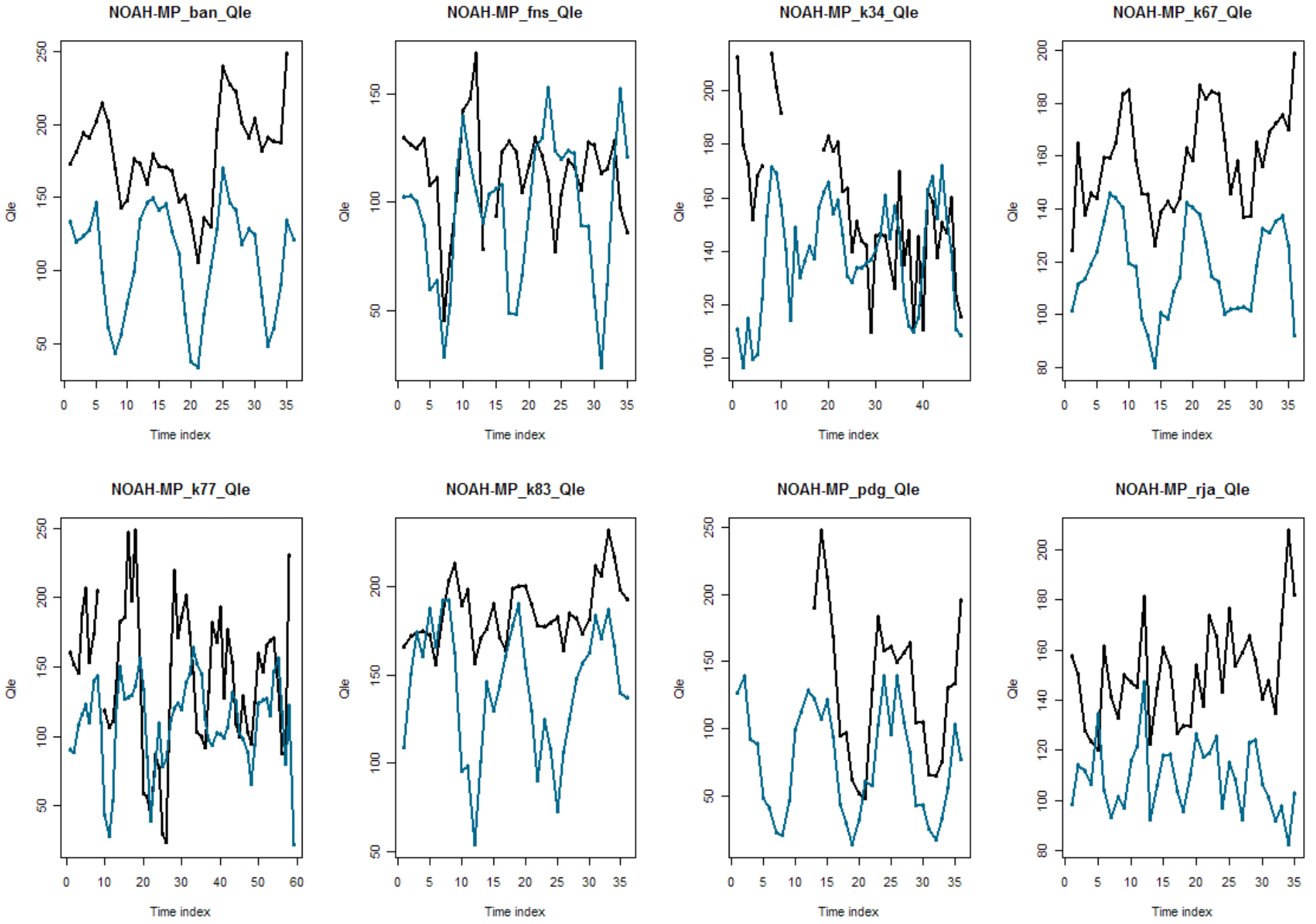

SiBCASA.c1_k77_Qle
Qle
Time index

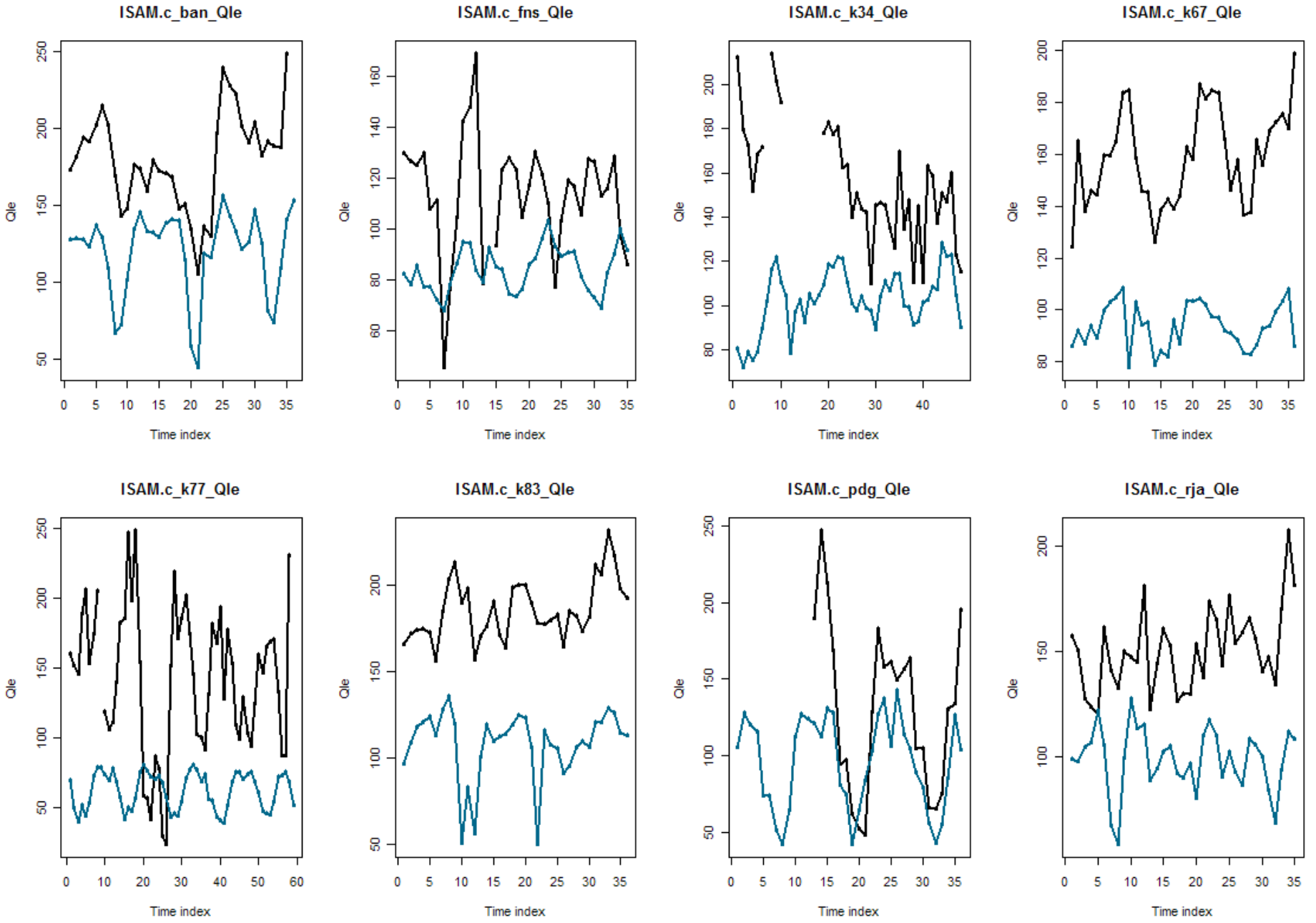

SiBCASA.c1_k83_Qle
Qle
Time index

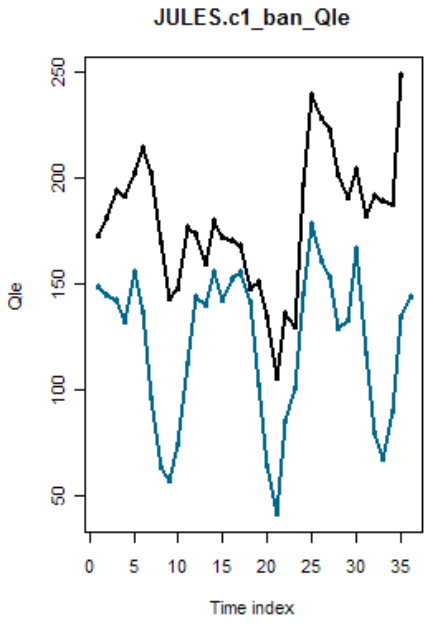

SiBCASA.c1_pdg_Qle
Qle
Time index

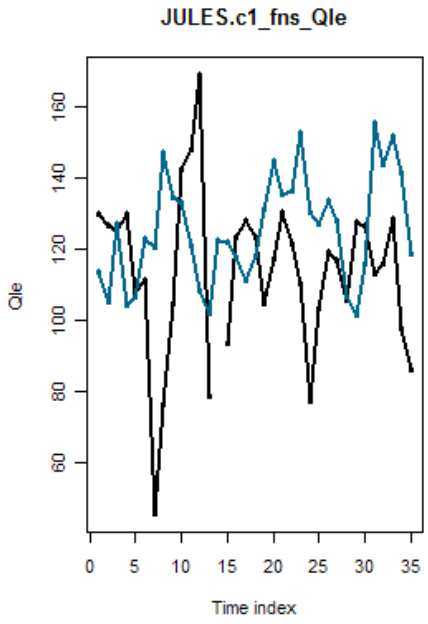

SiBCASA.c1_rja_Qle
Qle
Time index

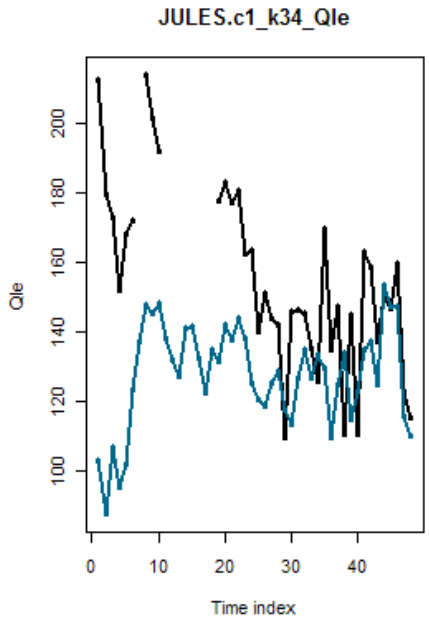

CLM4CN.ncl_ban_Qle
CLM4CN.ncl_fns_Qle
CLM4CN.ncl_k34_Qle
CLM4CN.ncl_k67_Qle
CLM4CN.ncl_k77_Qle
CLM4CN.ncl_k83_Qle
CLM4CN.ncl_pdg_Qle
CLM4CN.ncl_rja_Qle
Qle
Time index

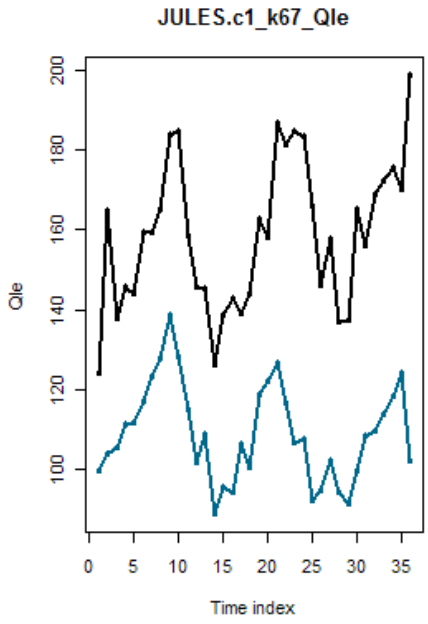

NOAH-MP_ban_Qle
NOAH-MP_fns_Qle
NOAH-MP_k34_Qle
NOAH-MP_k67_Qle
NOAH-MP_k77_Qle
NOAH-MP_k83_Qle
NOAH-MP_pdg_Qle
NOAH-MP_rja_Qle
Qle
Time index

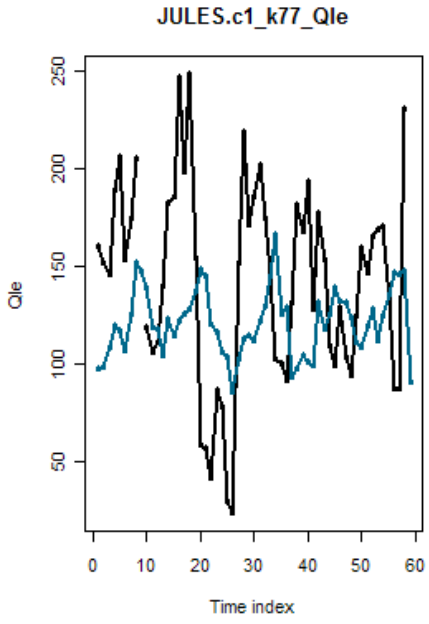

ISAM.c_ban_Qle
ISAM.c_fns_Qle
ISAM.c_k34_Qle
ISAM.c_k67_Qle
ISAM.c_k77_Qle
ISAM.c_k83_Qle
ISAM.c_pdg_Qle
ISAM.c_rja_Qle
Qle
Time index

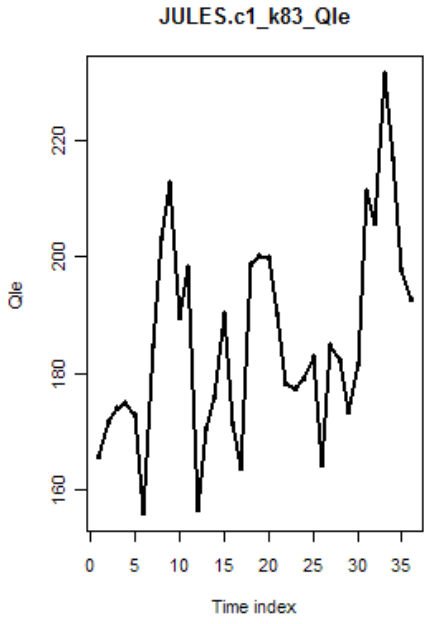

JULES.c1_ban_Qle
Qle
Time index

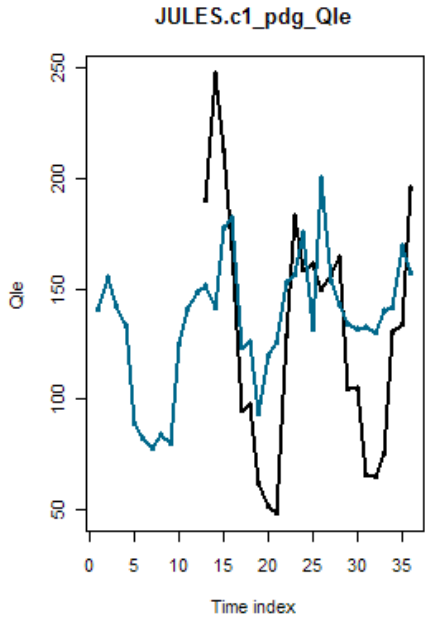

JULES.c1_fns_Qle
Qle
Time index

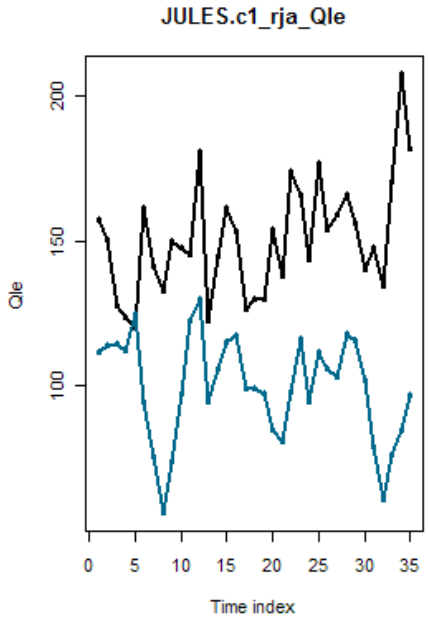

JULES.c1_k34_Qle
Qle
Time index

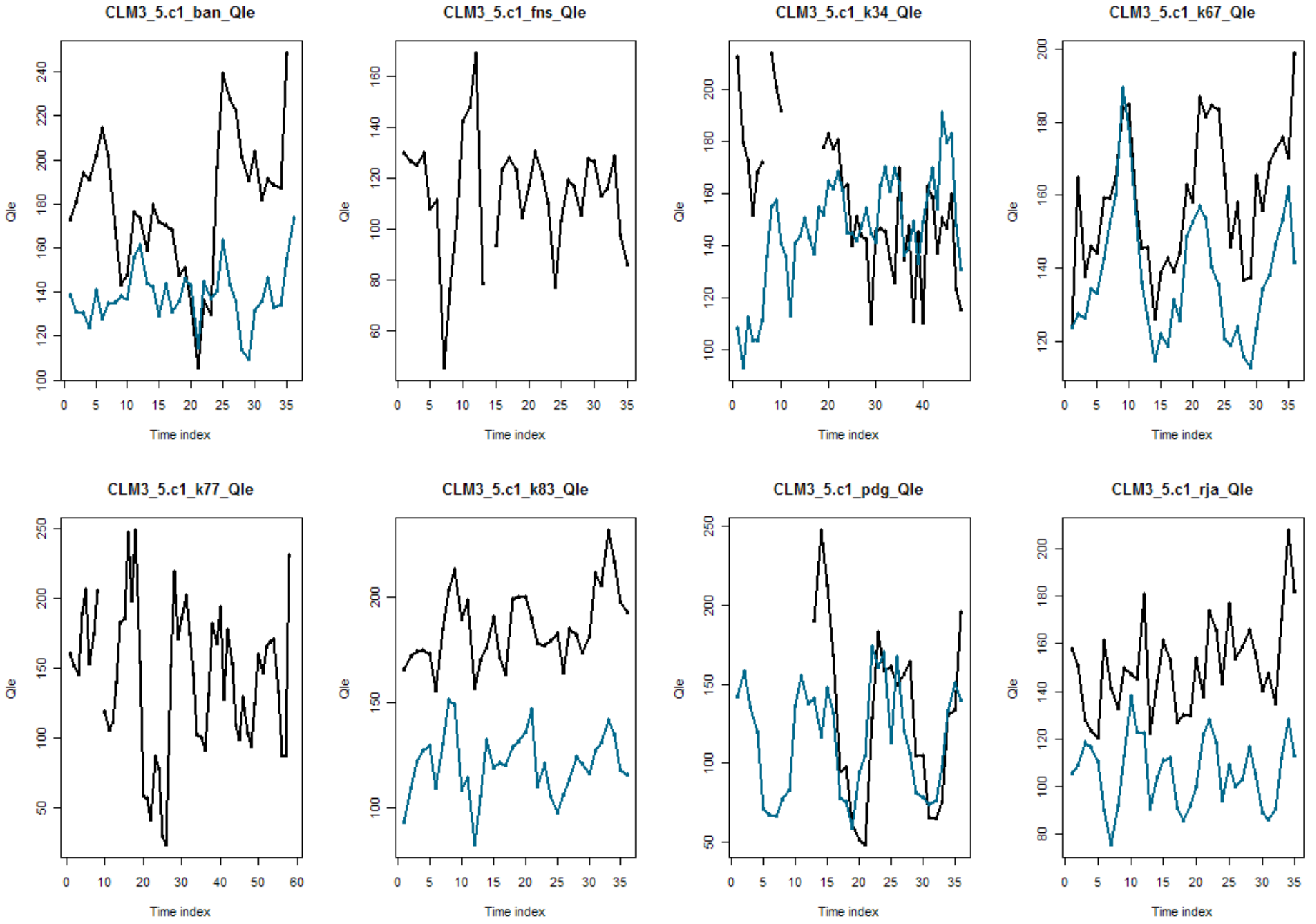

JULES.c1_k67_Qle
Qle
Time index

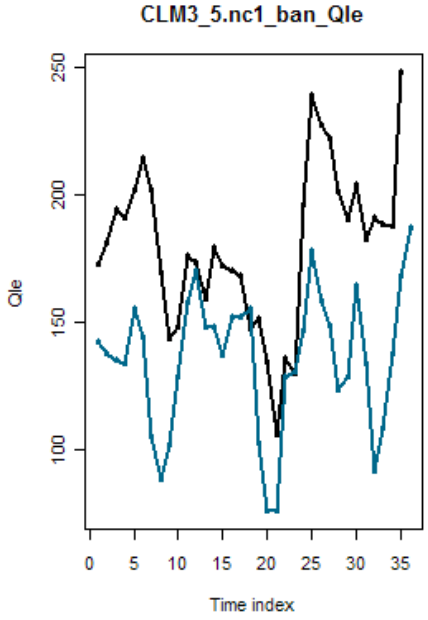

JULES.c1_k77_Qle
Qle
Time index

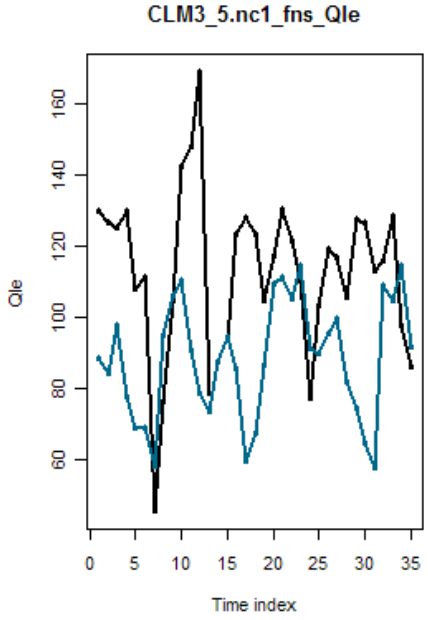

JULES.c1_k83_Qle
Qle
Time index

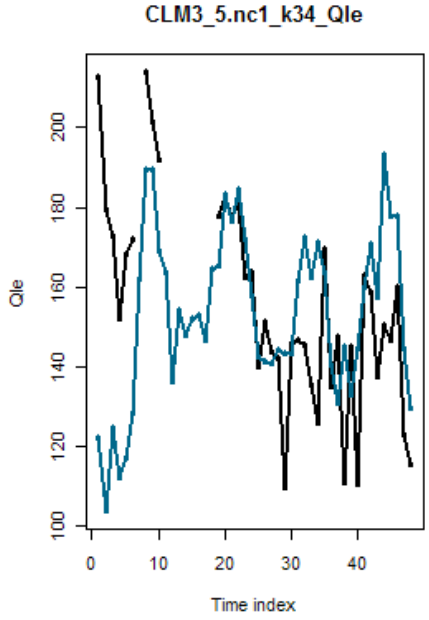

JULES.c1_pdg_Qle
Qle
Time index

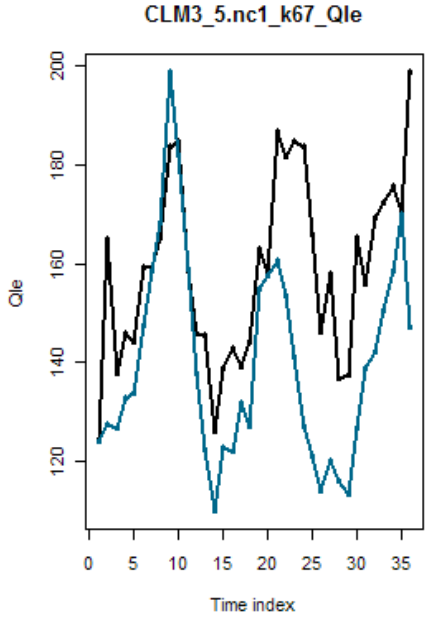

JULES.c1_rja_Qle
Qle
Time index

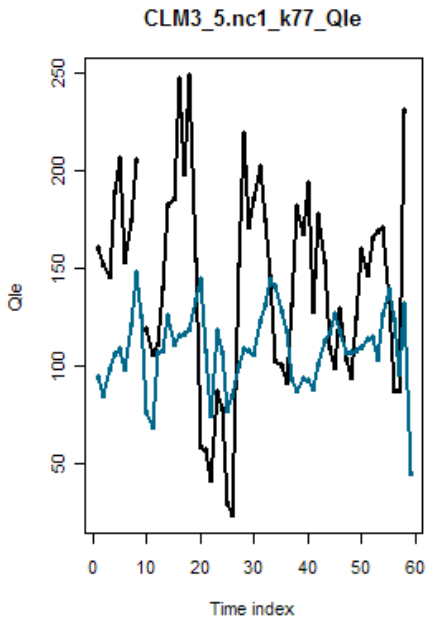

CLM3_5.c1_ban_Qle
CLM3_5.c1_fns_Qle
CLM3_5.c1_k34_Qle
CLM3_5.c1_k67_Qle
CLM3_5.c1_k77_Qle
CLM3_5.c1_k83_Qle
CLM3_5.c1_pdg_Qle
CLM3_5.c1_rja_Qle
Qle
Time index

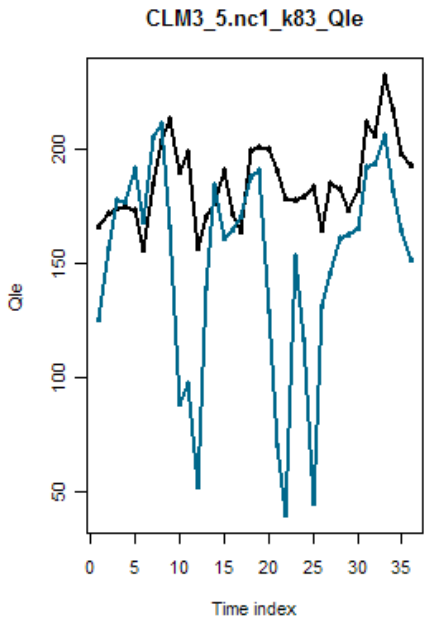

CLM3_5.nc1_ban_Qle
Qle
Time index

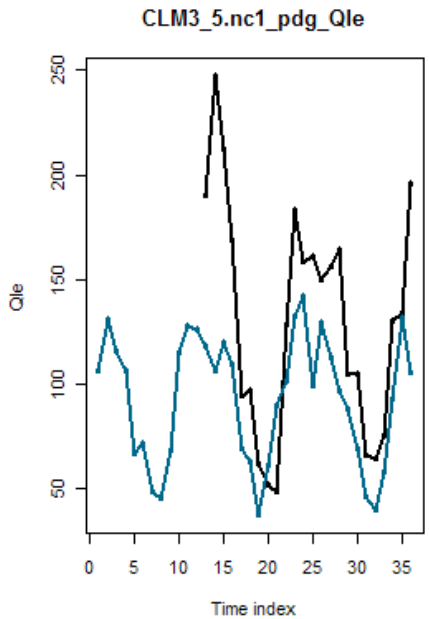

CLM3_5.nc1_fns_Qle
Qle
Time index

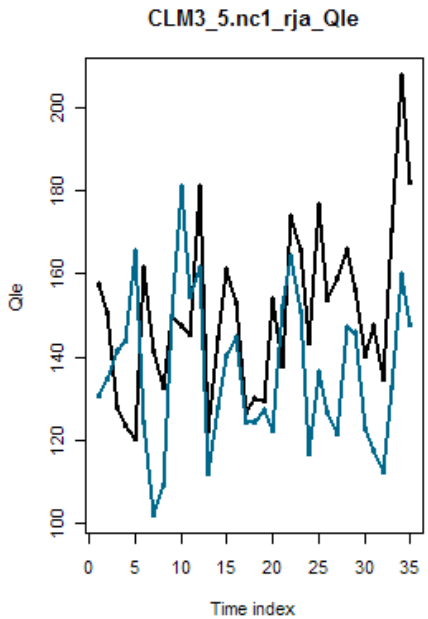

CLM3_5.nc1_k34_Qle
Qle
Time index

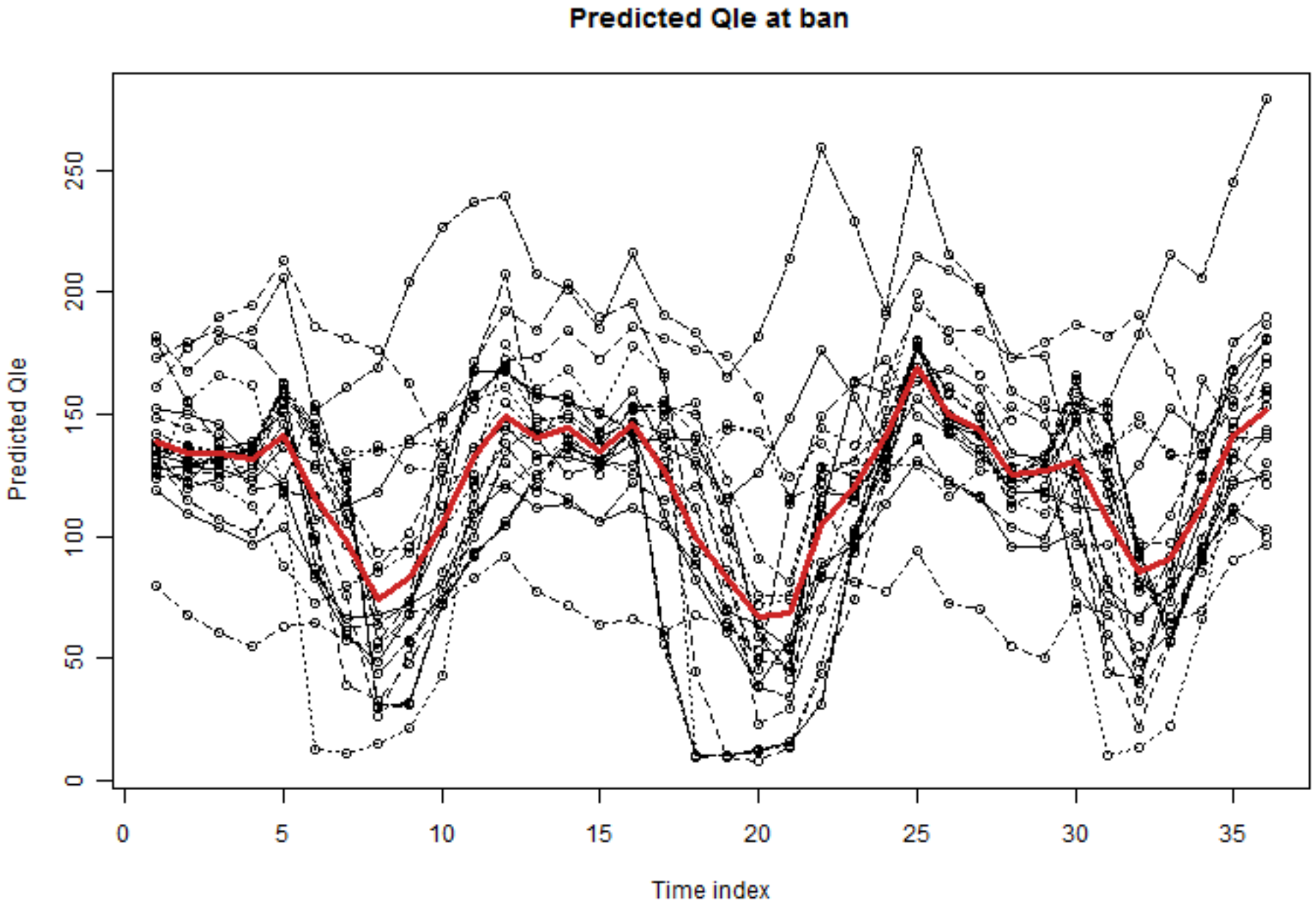

CLM3_5.nc1_k67_Qle
Qle
Time index

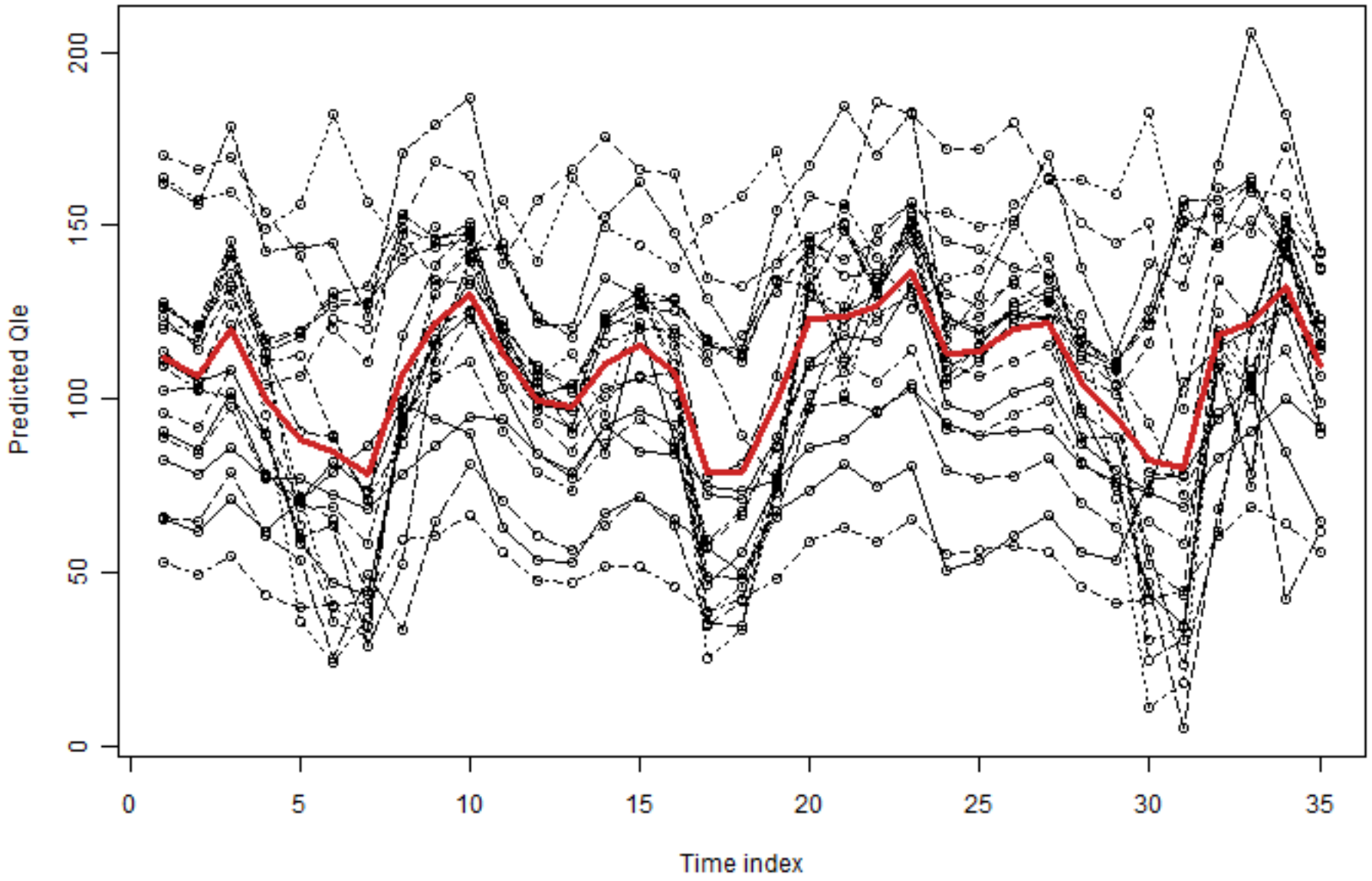

CLM3_5.nc1_k77_Qle
Qle
Time index

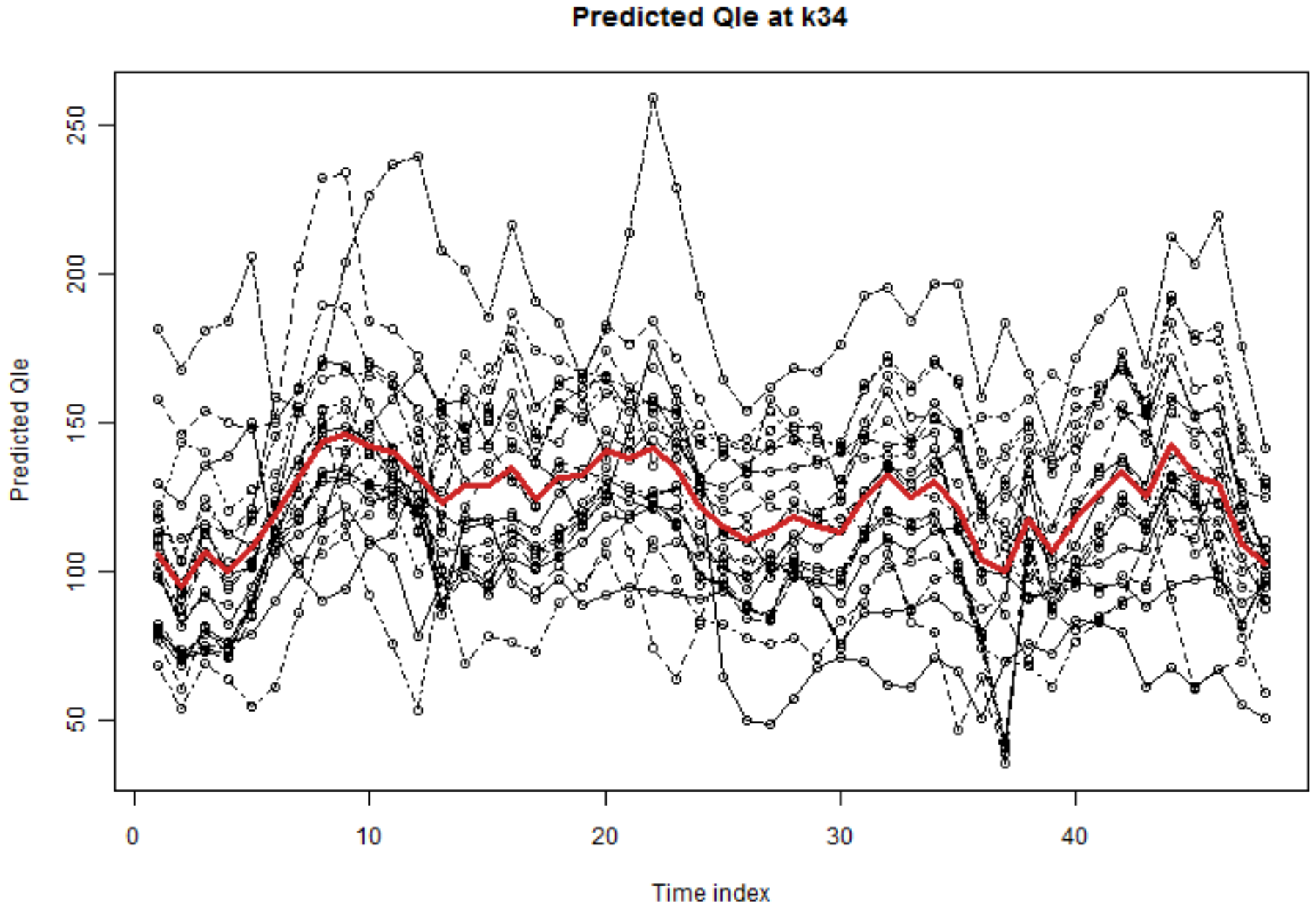

CLM3_5.nc1_k83_Qle
Qle
Time index

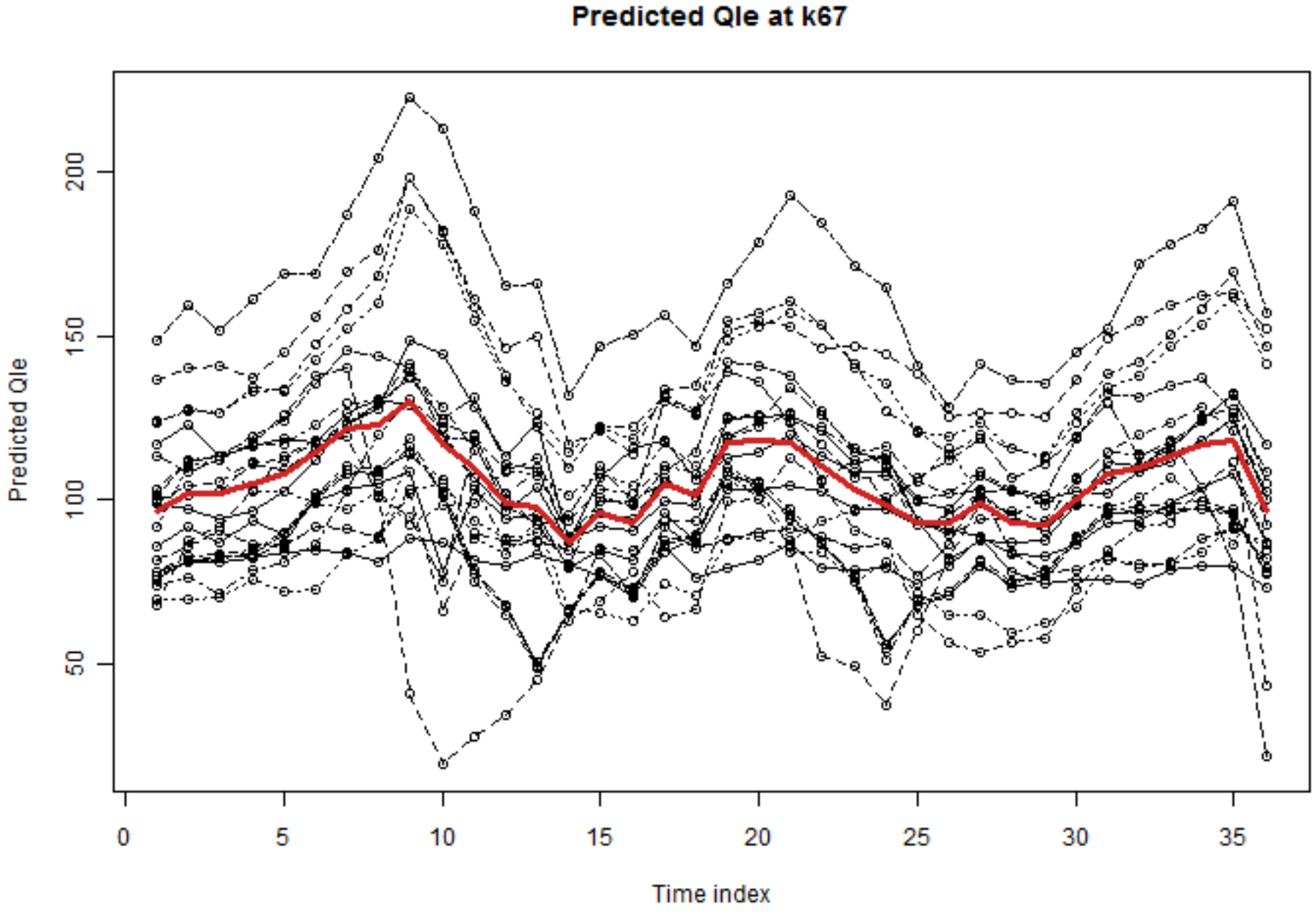

CLM3_5.nc1_pdg_Qle
Qle
Time index

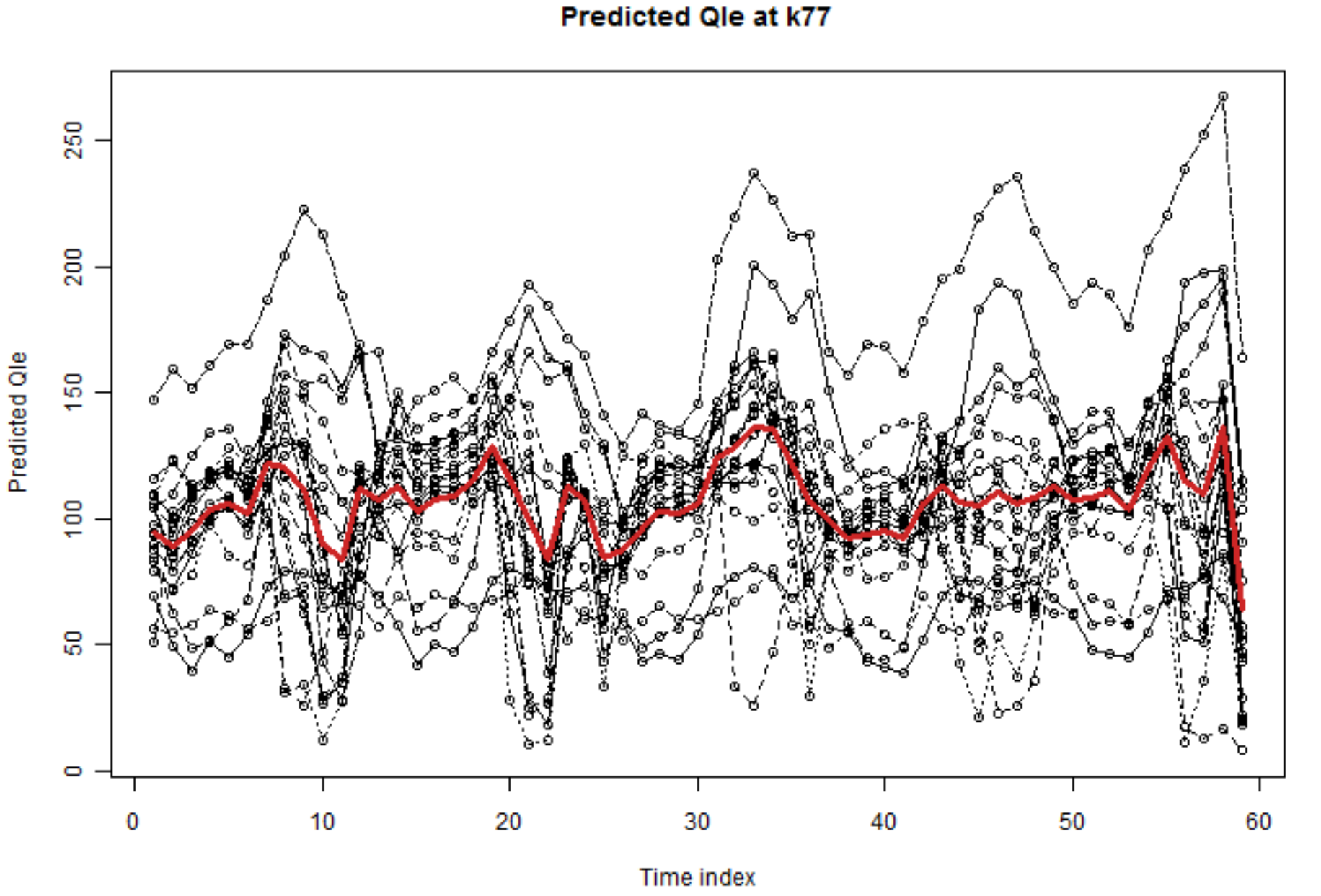

CLM3_5.nc1_rja_Qle
Qle
Time index

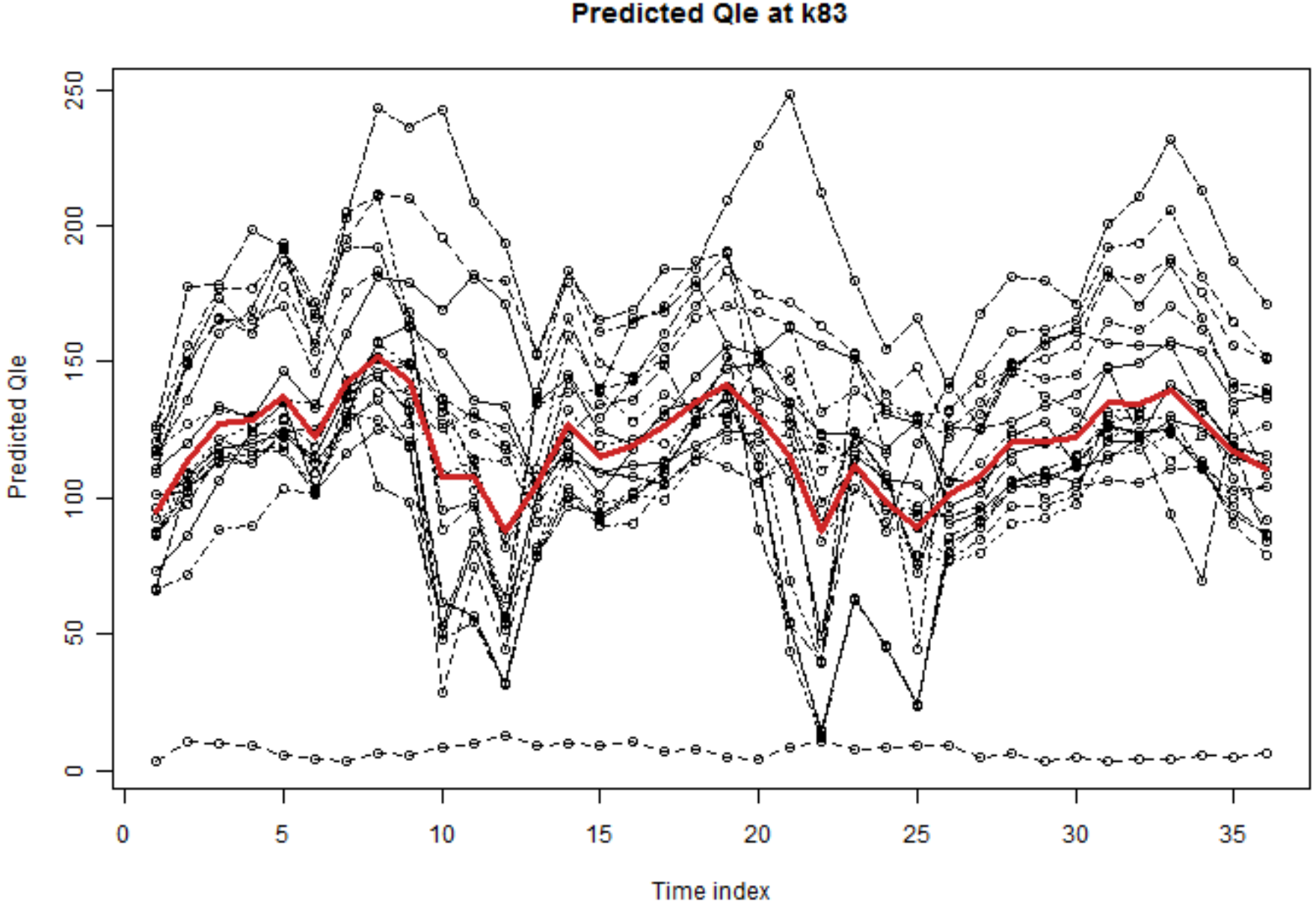

Predicted Qle at ban
Predicted Qle
0
50
100
150
200
250
0
5
10
15
20
25
30
35
Time index

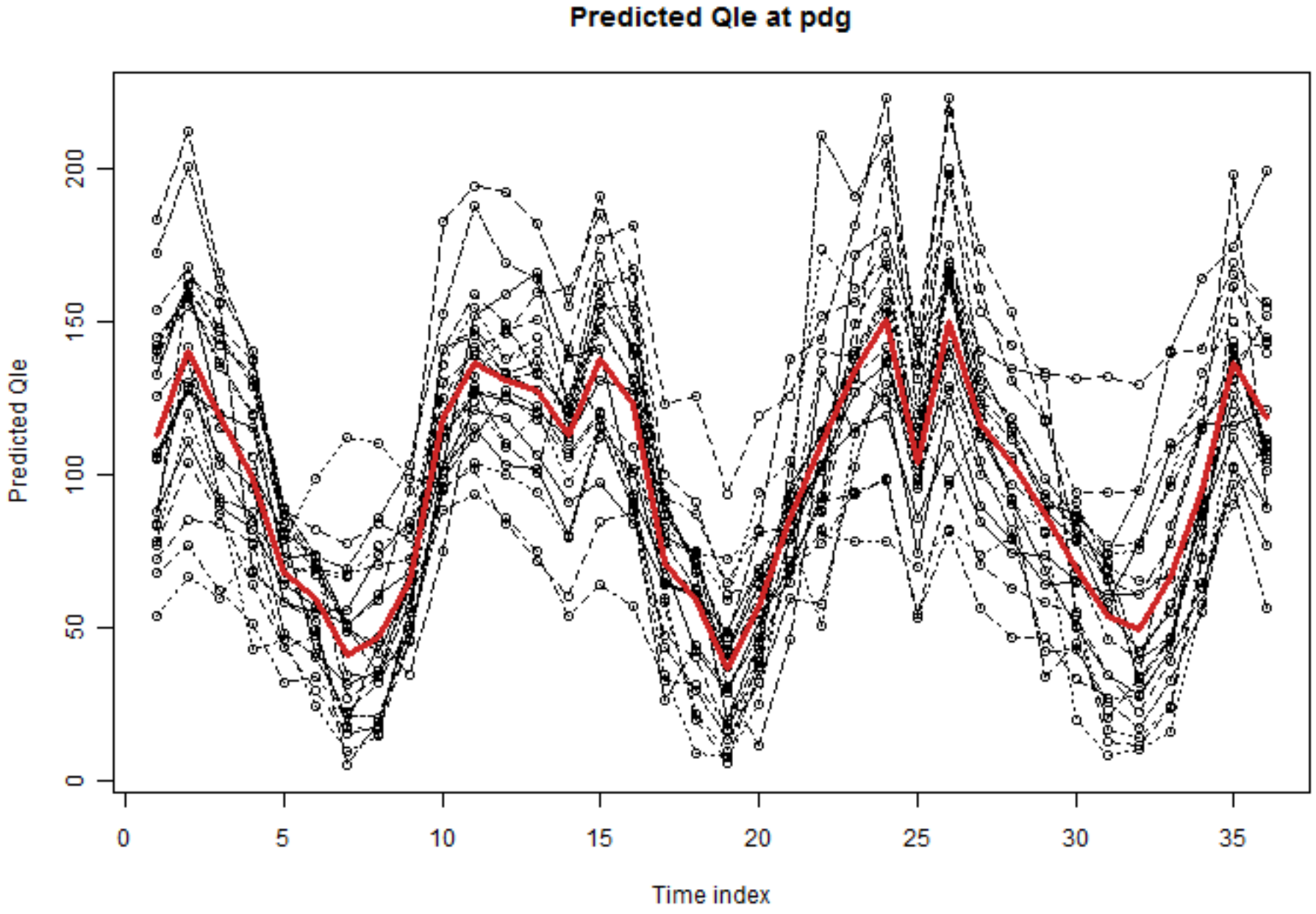

Predicted Qle at fns
Predicted Qle
0
50
100
150
200
0
5
10
15
20
25
30
35
Time index

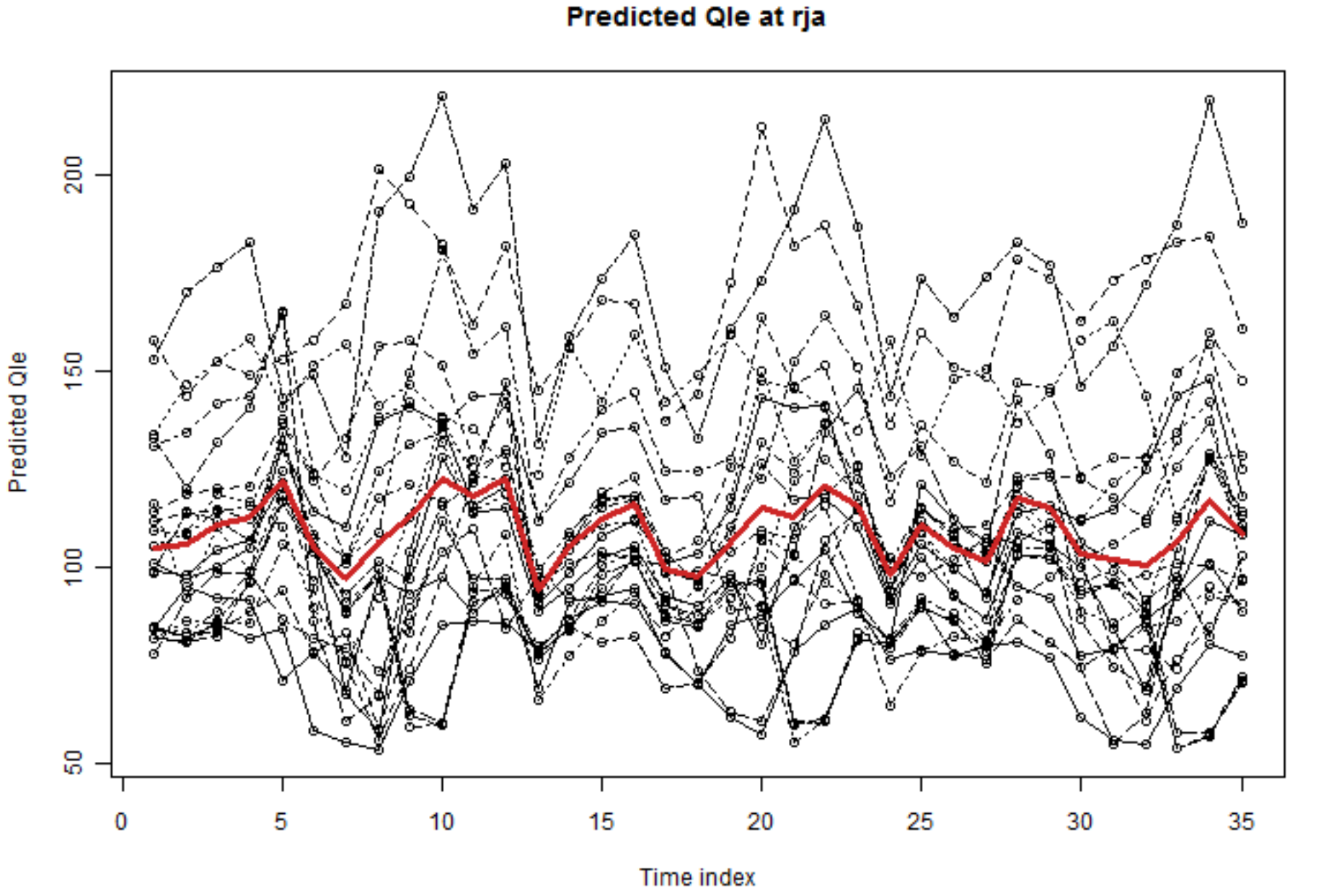

Predicted Qle at k34
Predicted Qle
50
100
150
200
250
0
10
20
30
40
Time index


Predicted Qle at k67
Predicted Qle
50
100
150
200
0
5
10
15
20
25
30
35
Time index


Predicted Qle at k77
Predicted Qle
0
50
100
150
200
250
0
10
20
30
40
50
60
Time index


Predicted Qle at k83
Predicted Qle
Time index
0
50
100
150
200
250
0
5
10
15
20
25
30
35


Predicted Qle at pdg
Predicted Qle
0
50
100
150
200
0
5
10
15
20
25
30
35
Time index


Predicted Qle at rja
Predicted Qle
50
100
150
200
0
5
10
15
20
25
30
35
Time index

Statistical output

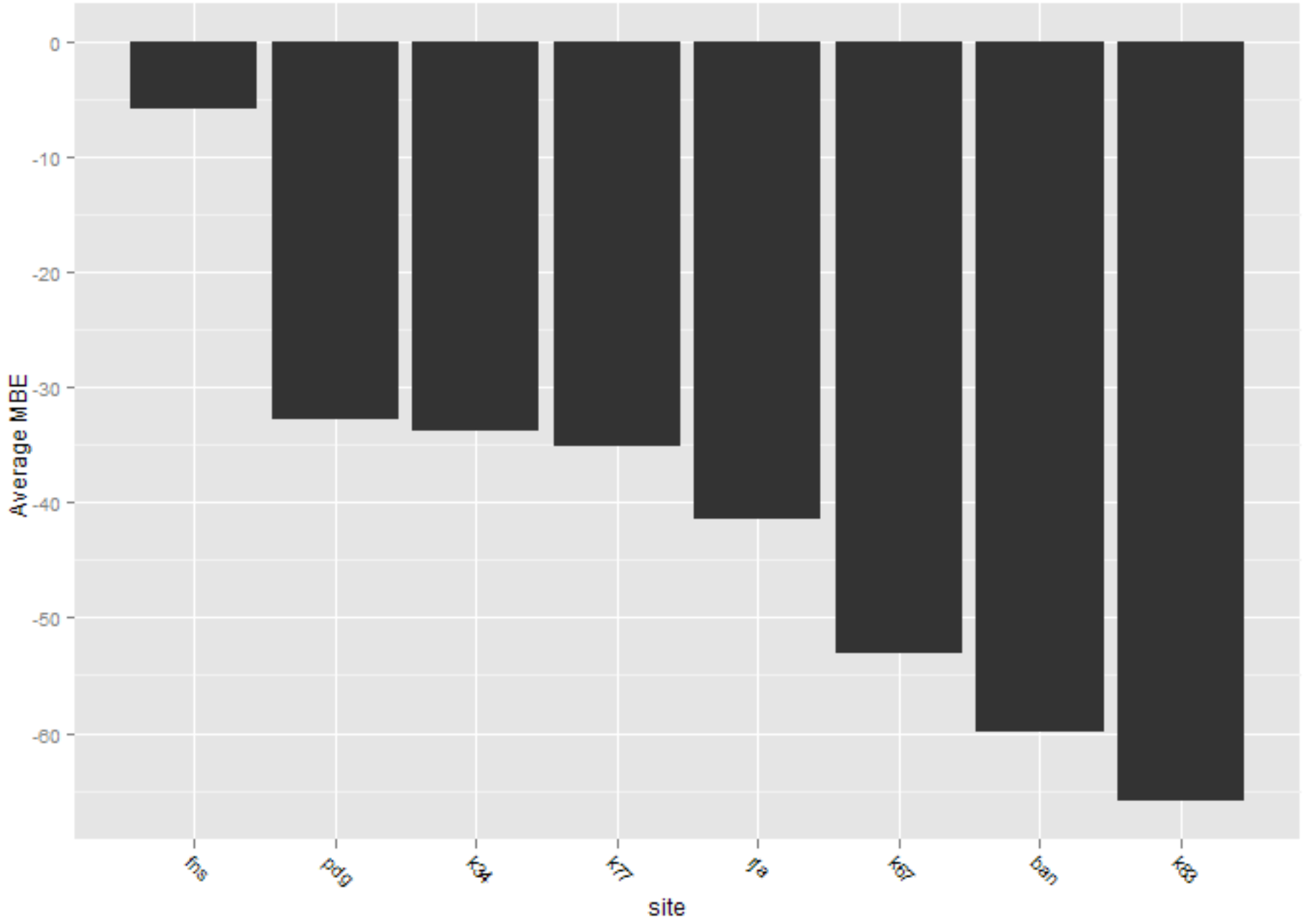
$$MBE = \frac{\sum_{i=1}^{n}(P_i - O_i)}{N}$$

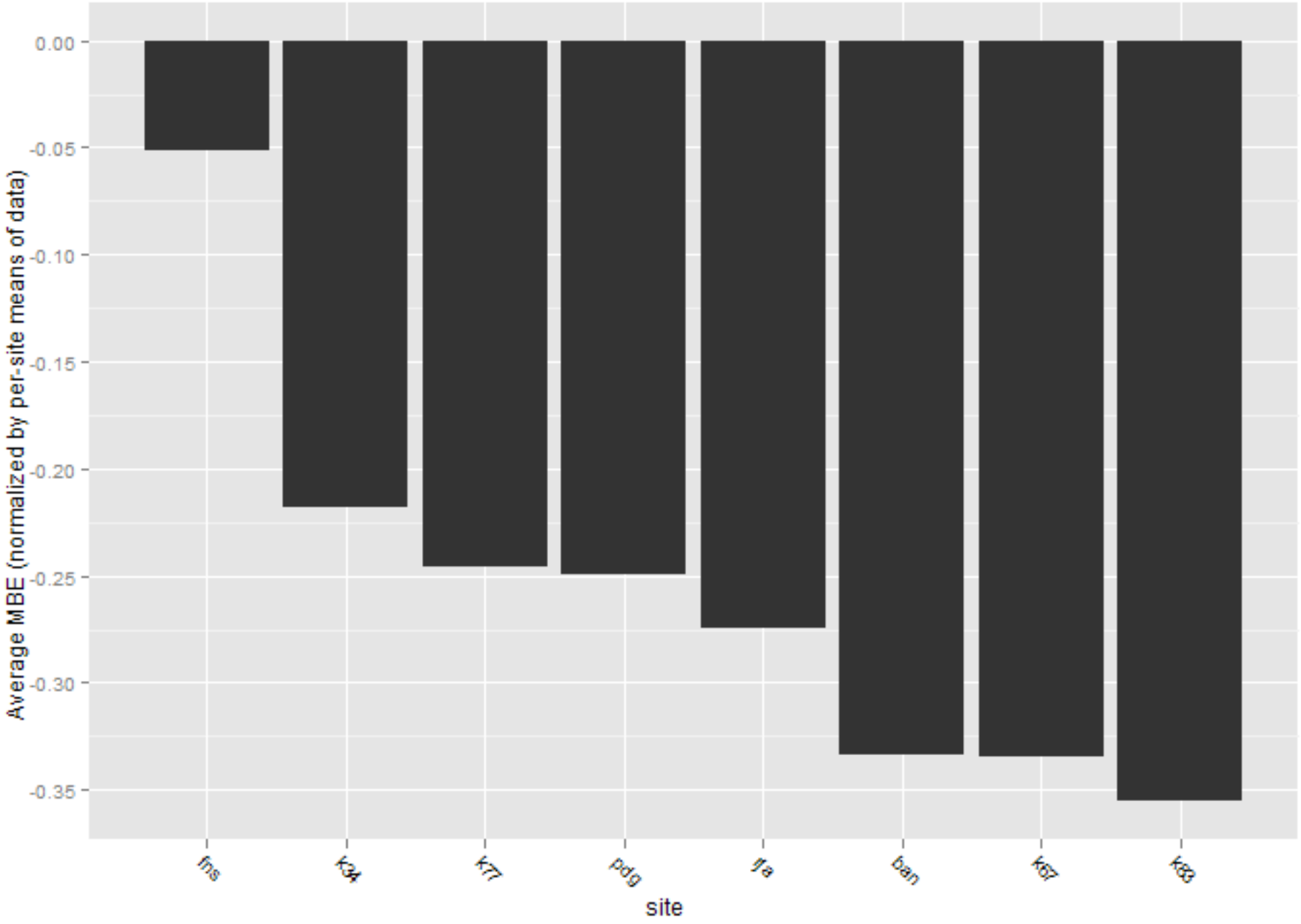
$$MBE_{norm} = \frac{\left(\frac{\sum_{i=1}^{n}(P_i - O_i)}{N}\right)}{\overline{O}}$$

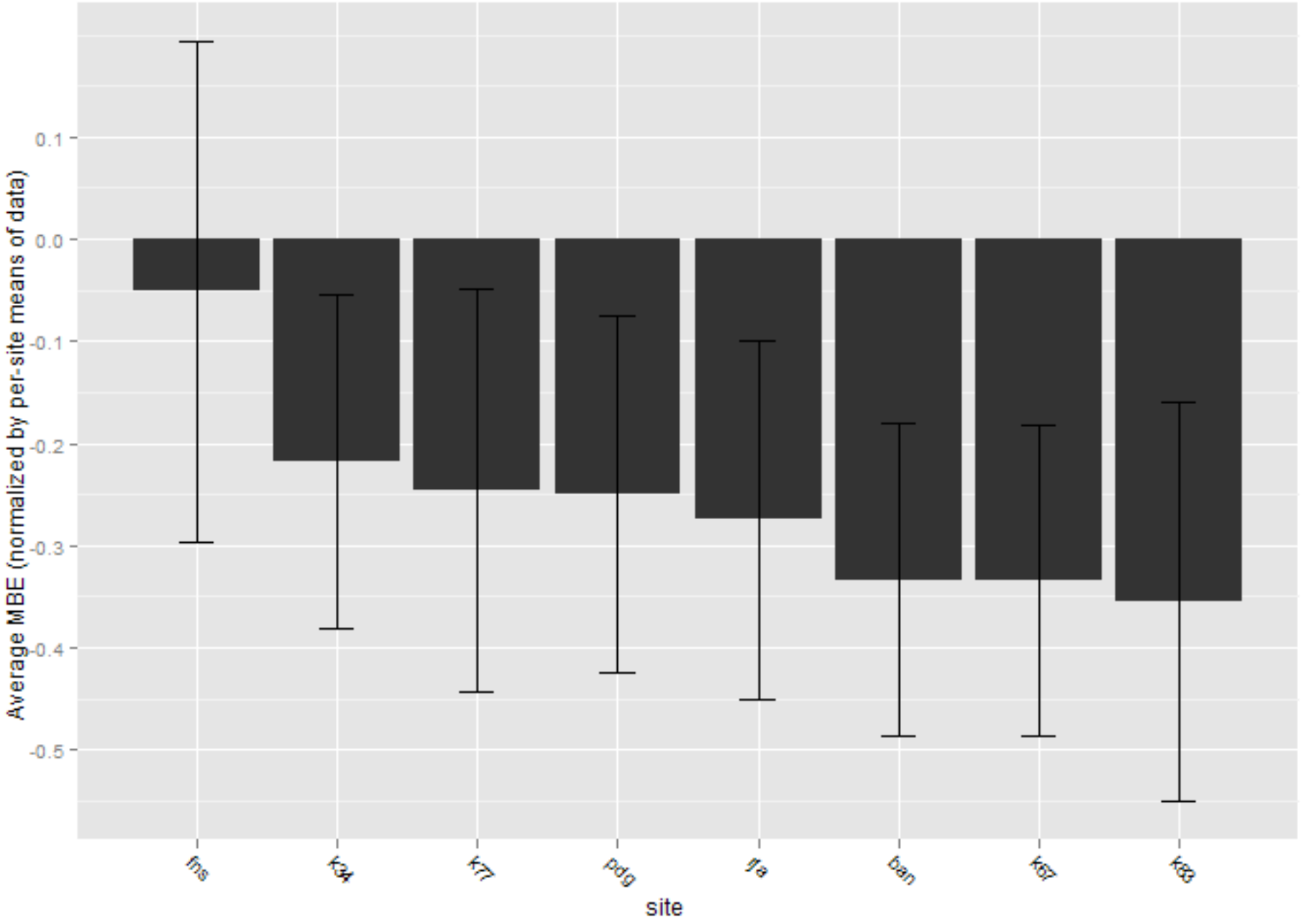
$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{N}}$$

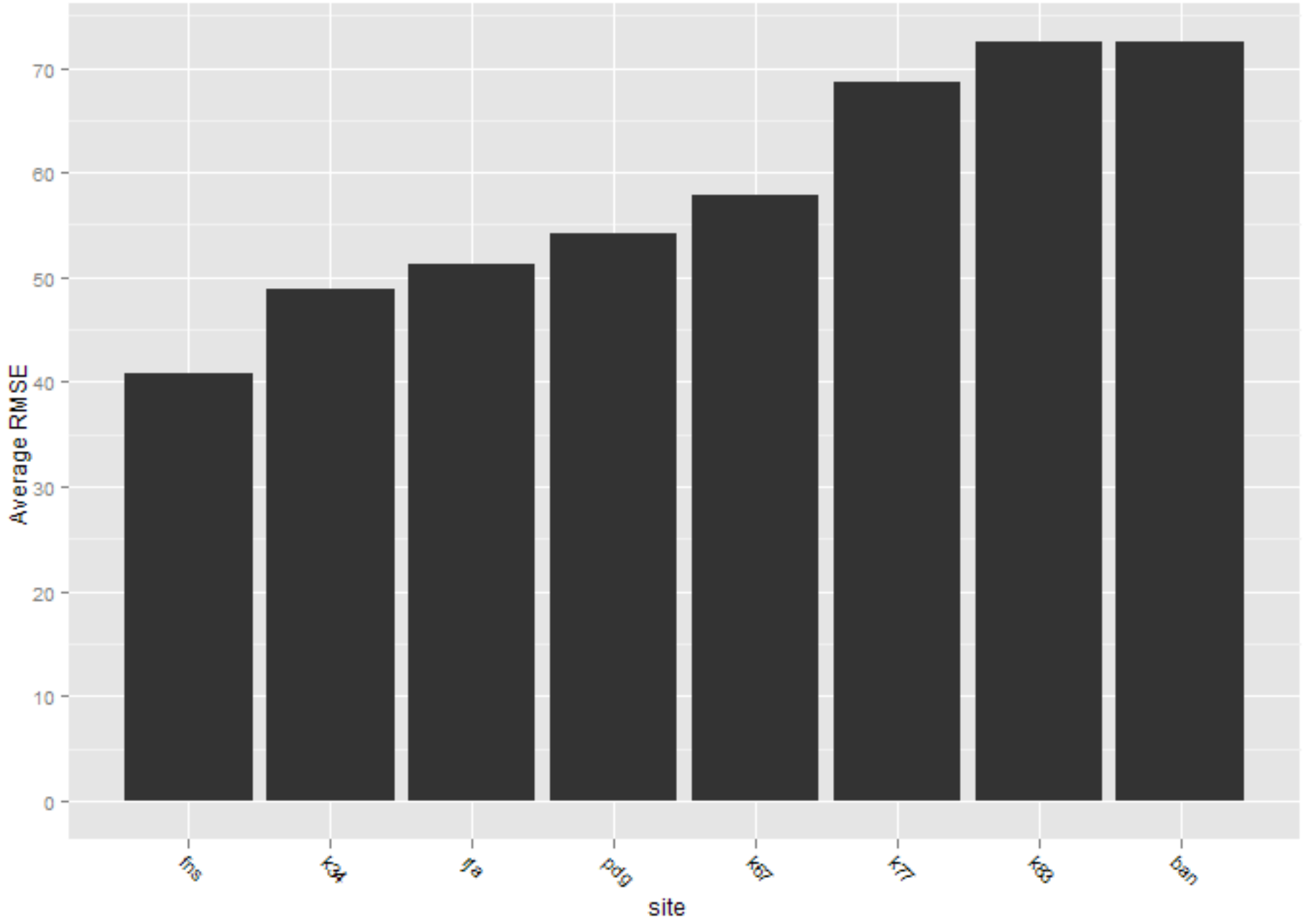
$$RMSE_{norm} = \frac{\left(\sqrt{\frac{\sum_{i=1}^{n}(P_i - O_i)^2}{N}}\right)}{SE_{Obs}}$$

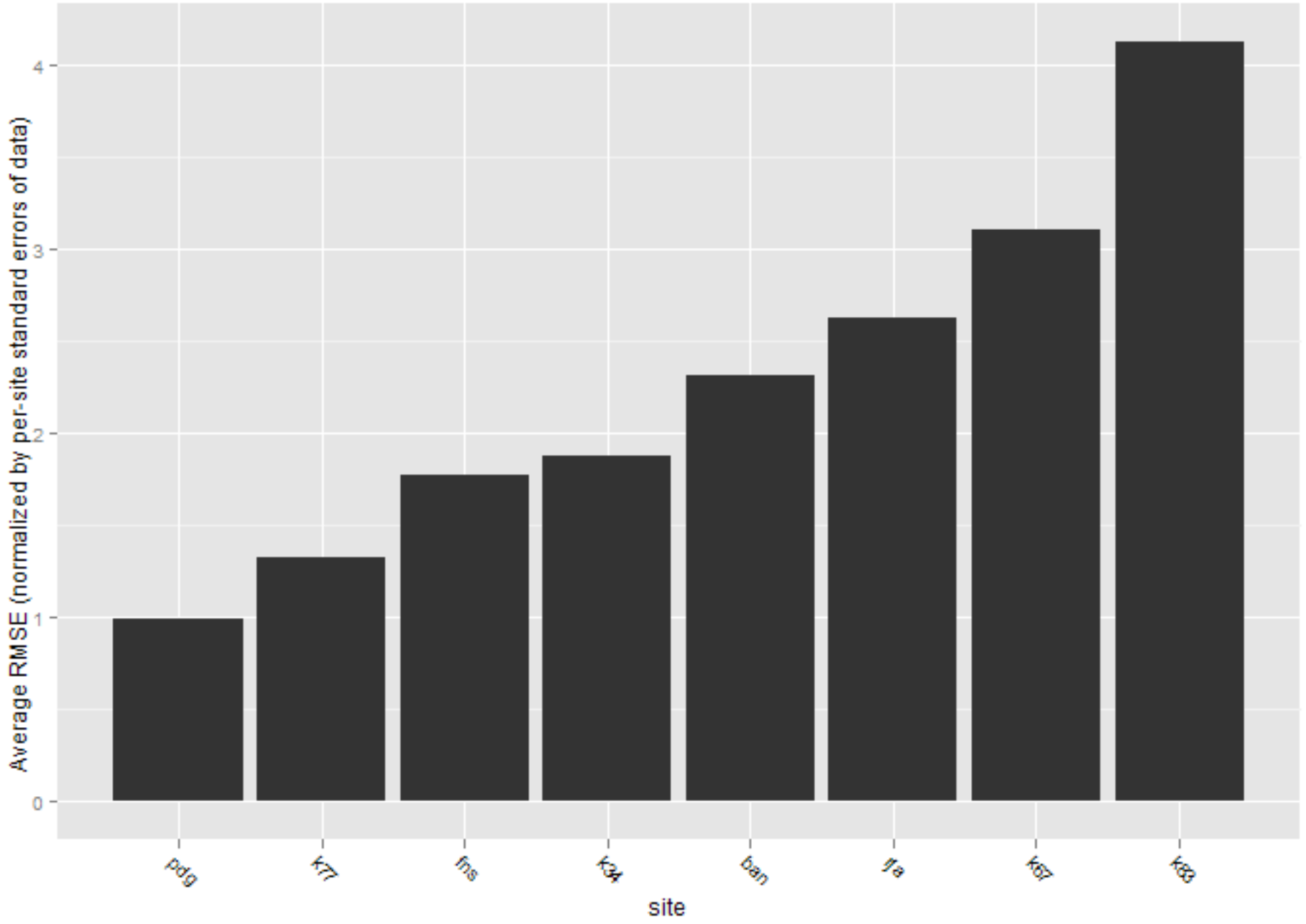
$$R^2 = \frac{SS_{err}}{SS_{tot}}$$

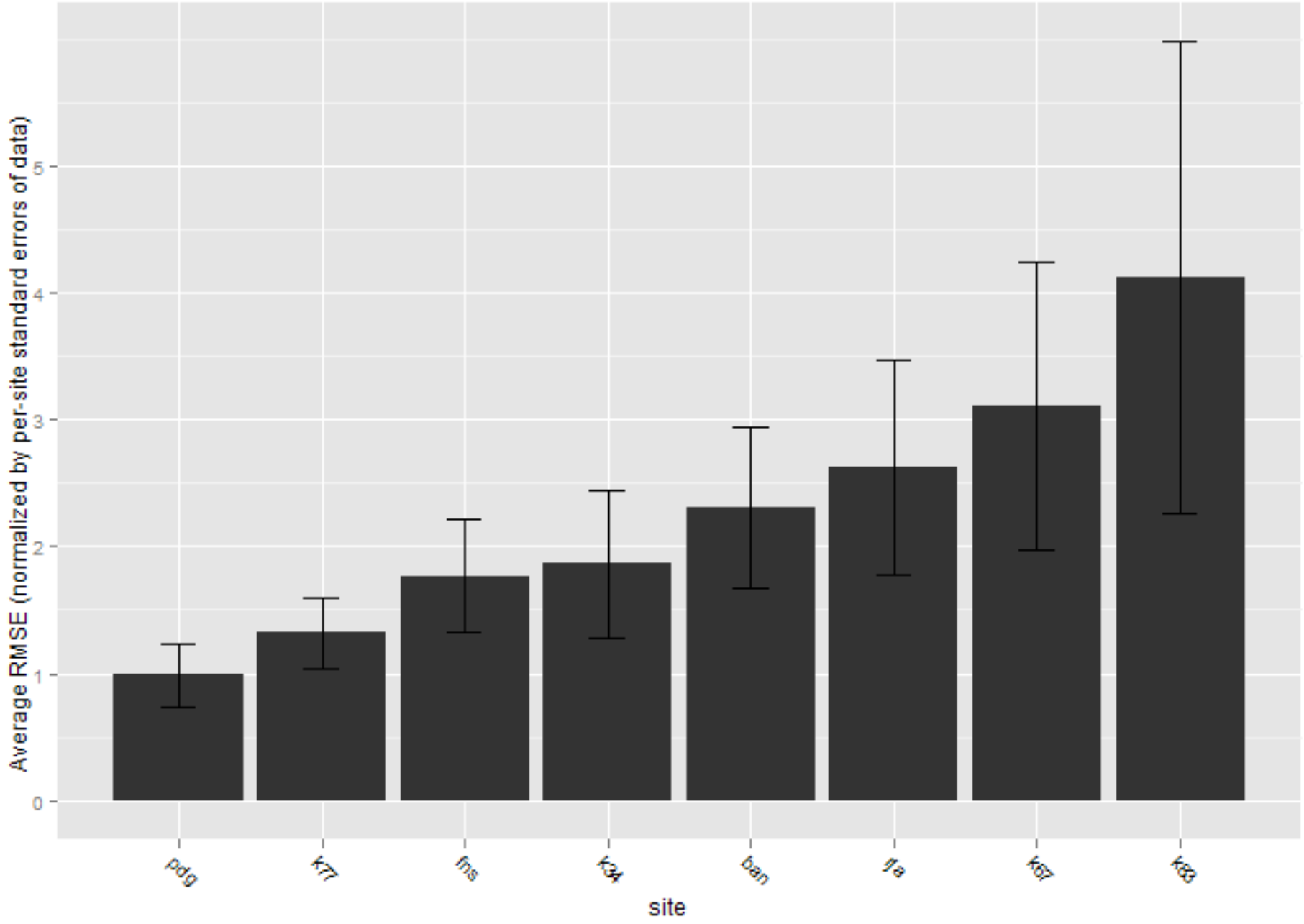

0
-10
-20
-30
-40
-50
-60
Average MBE
fns
pdg
k34
k77
fja
k67
ban
k83
site

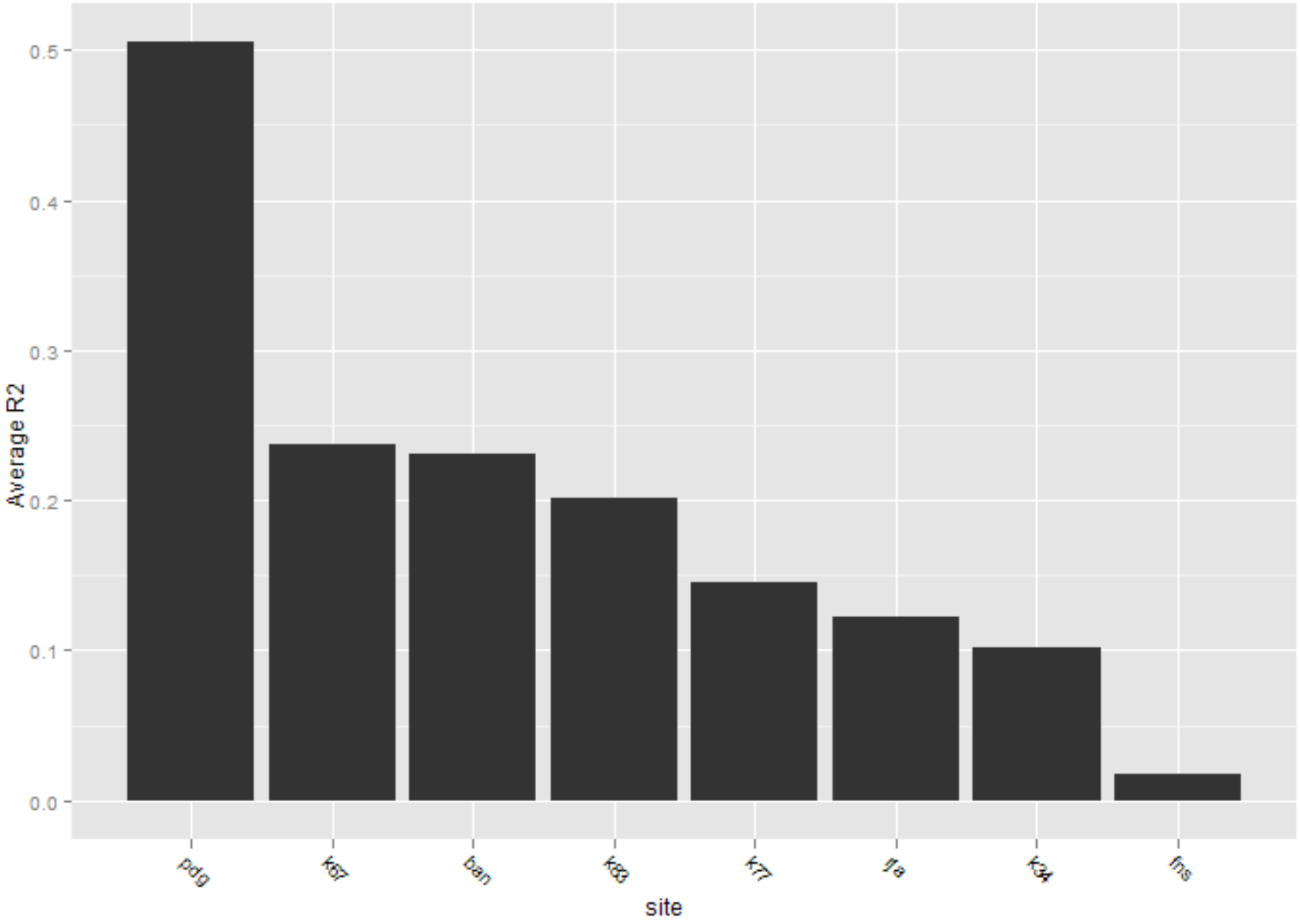

0.00
-0.05
-0.10
-0.15
-0.20
-0.25
-0.30
-0.35
Average MBE (normalized by per-site means of data)
fns
k34
k77
pdg
ffa
ban
k67
k83
site

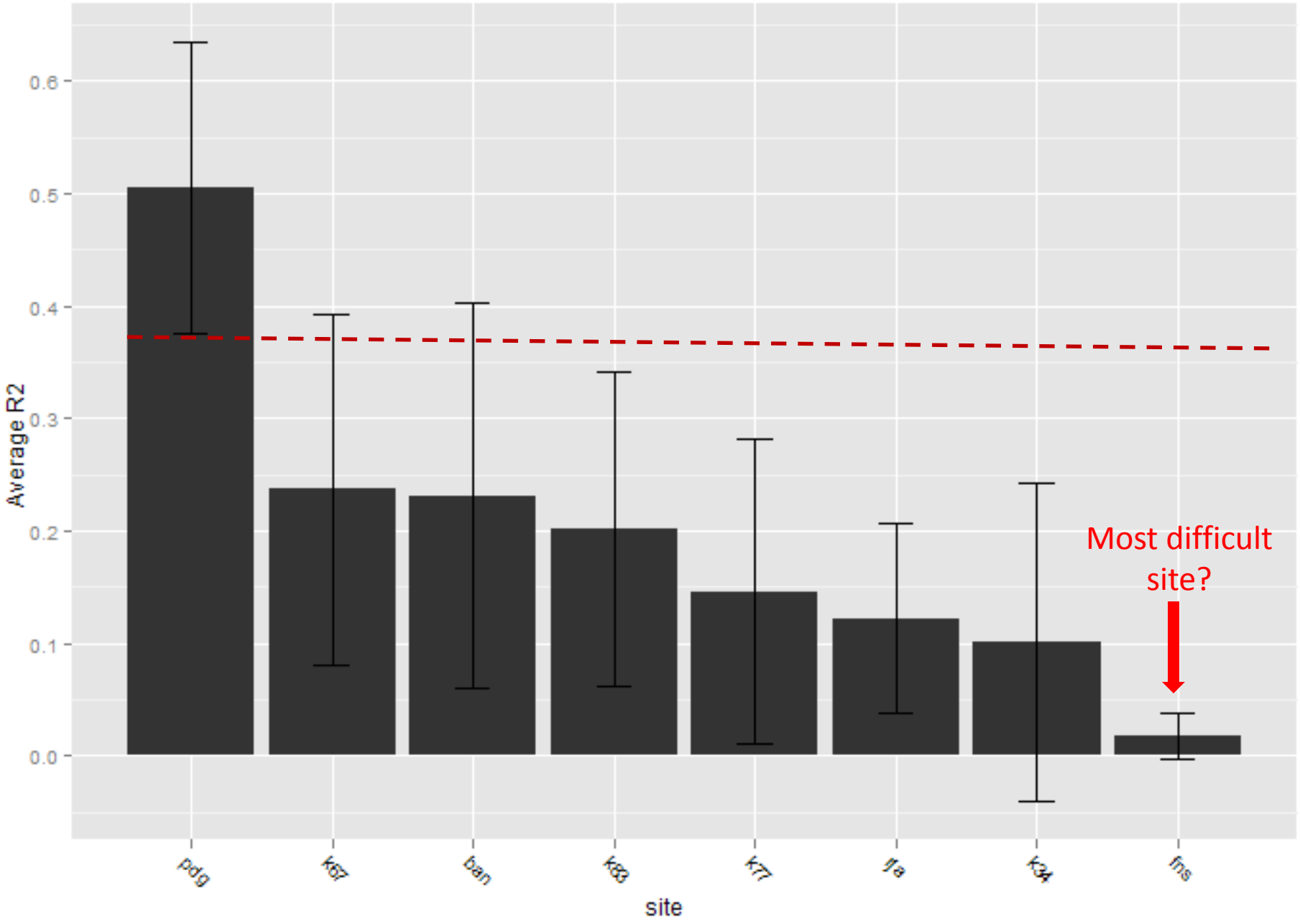

Average MBE (normalized by per-site means of data)
0.1
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
fns
k34
k77
pdg
fja
ban
k67
k83
site

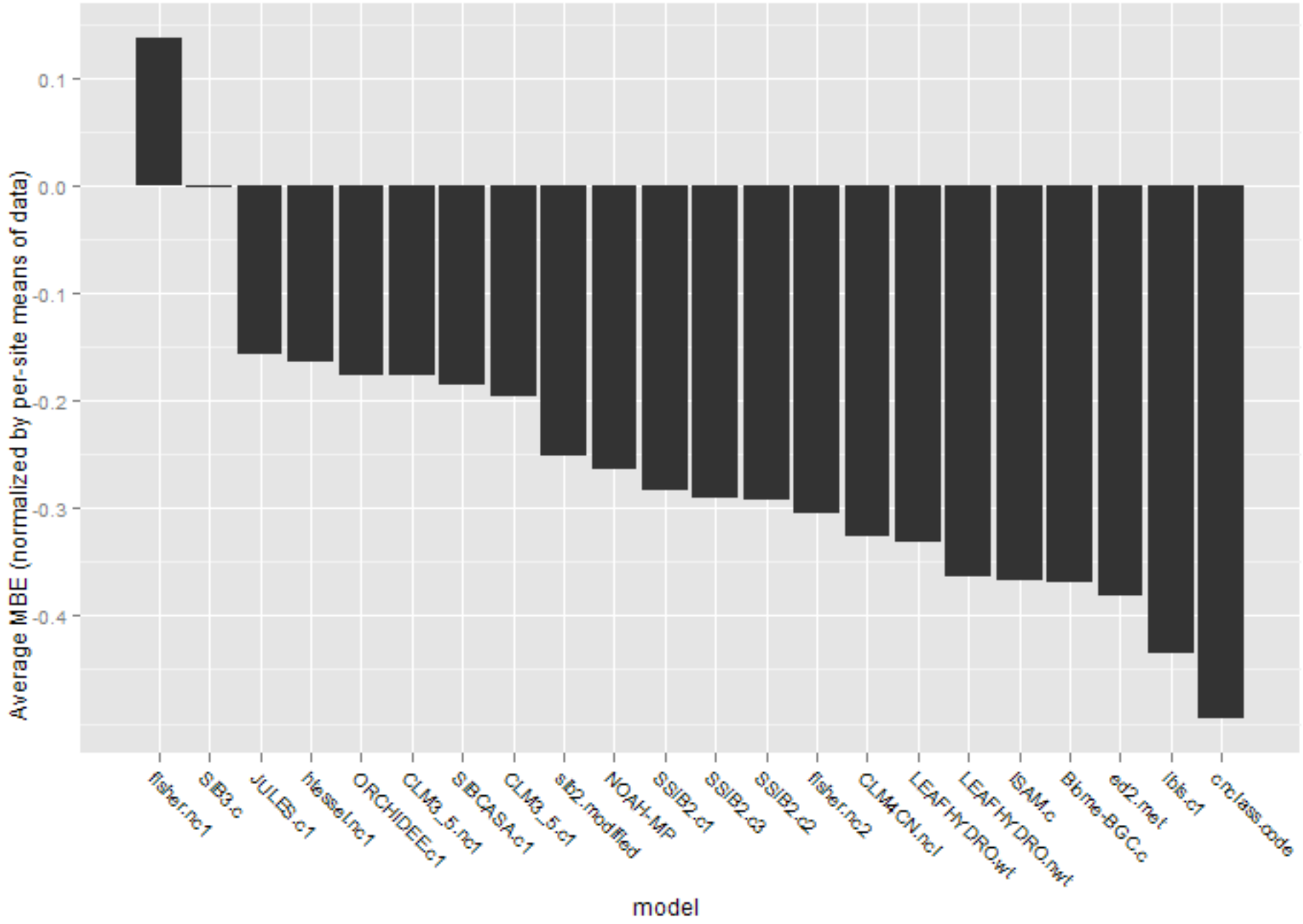

Average RMSE
0
10
20
30
40
50
60
70
fns
k34
ifa
pdg
k67
k77
k83
ban
site

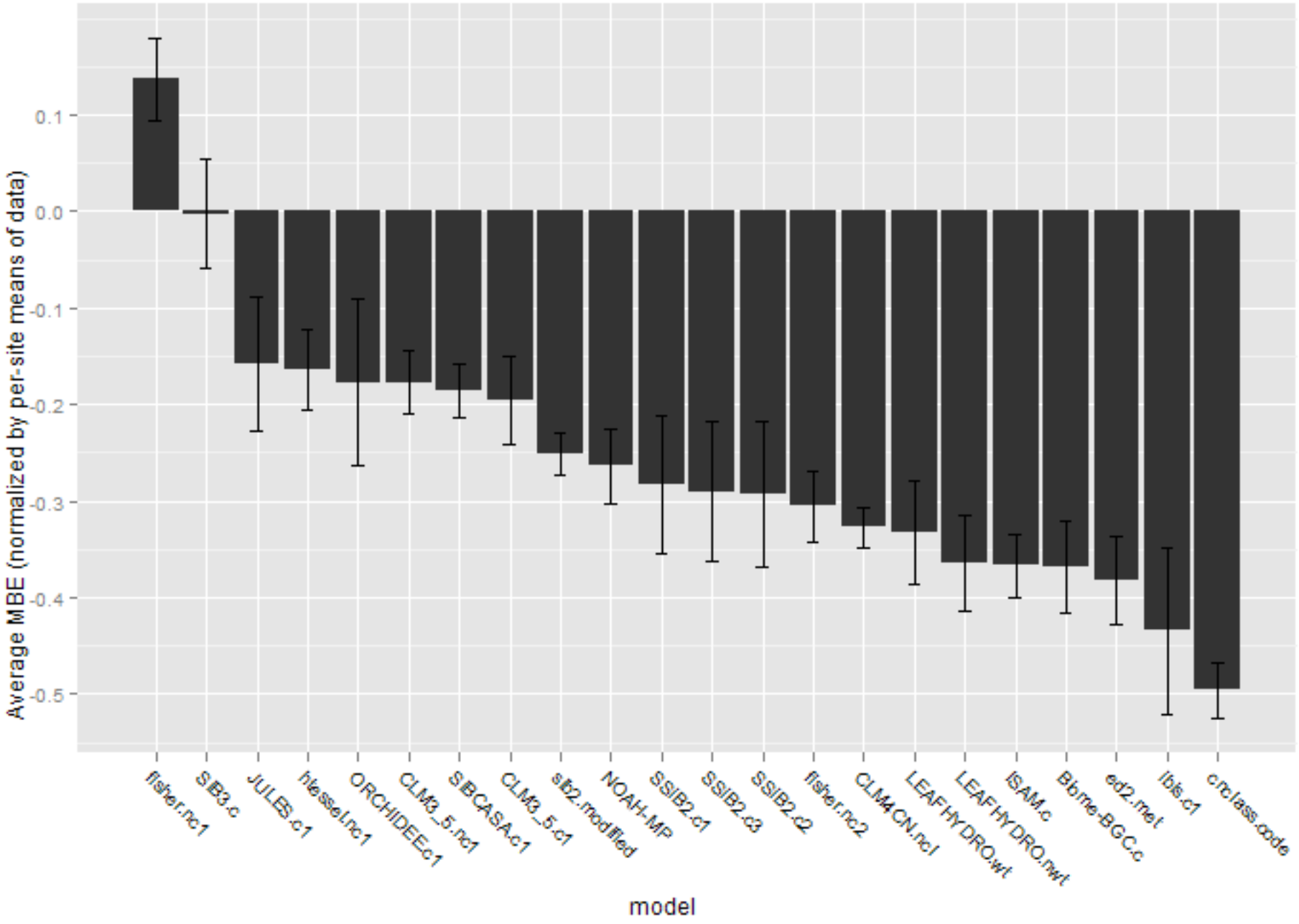

Average RMSE (normalized by per-site standard errors of data)
0
1
2
3
4
pdg
k77
fns
k34
ban
fla
k67
k83
site

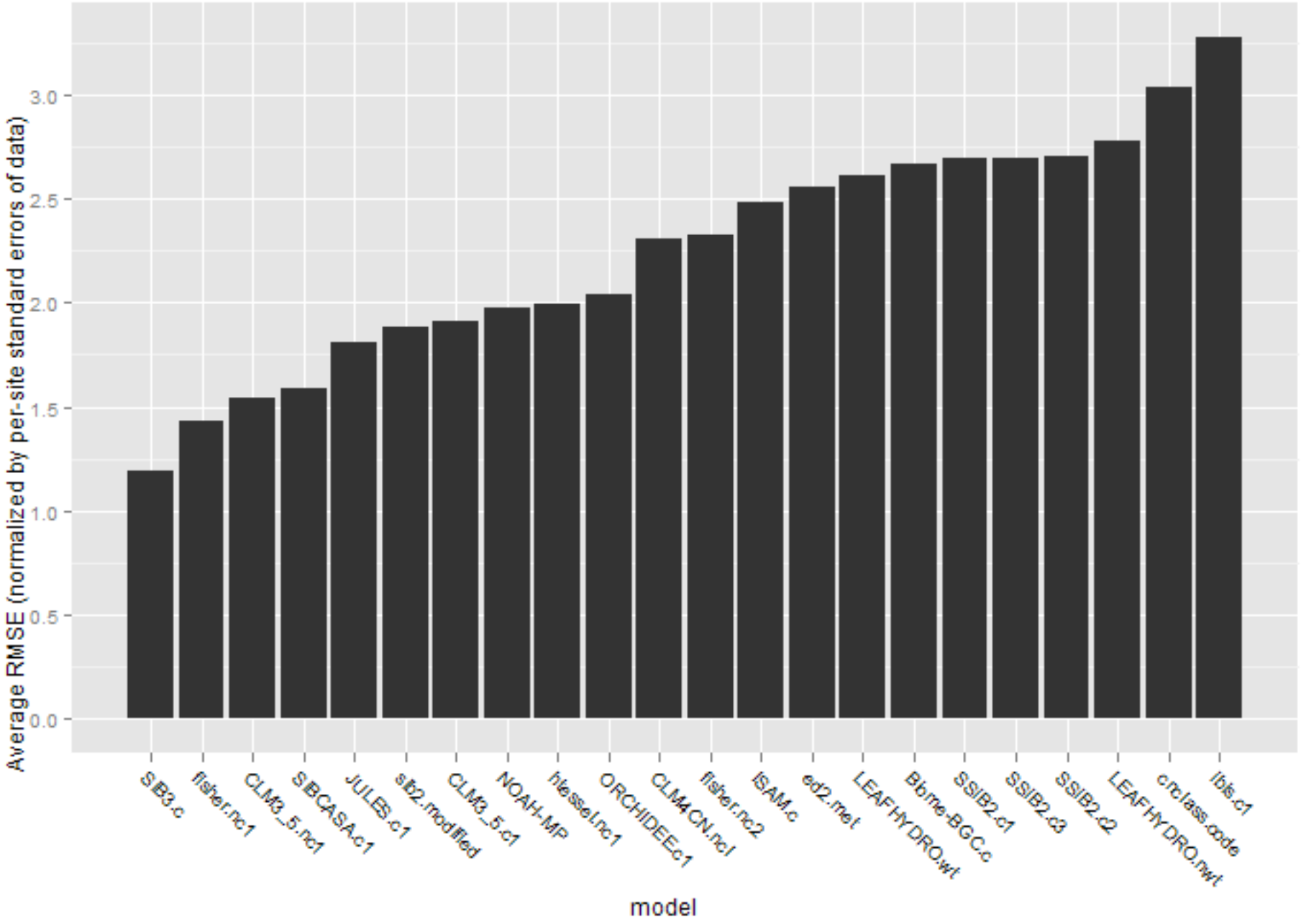

Average RMSE (normalized by per-site standard errors of data)
0
1
2
3
4
5
pdg
k77
fns
k34
ban
fla
k67
k83
site

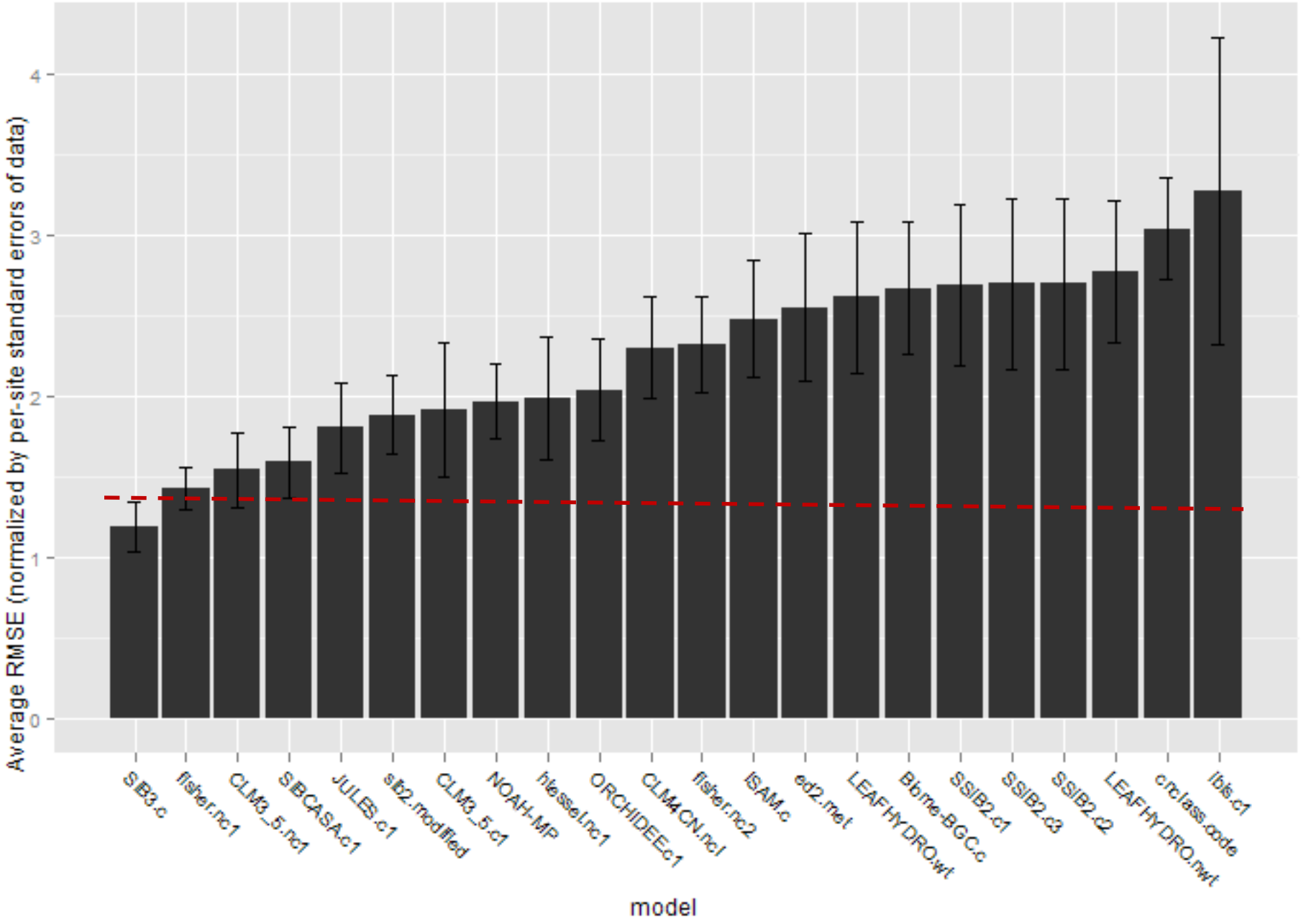

0.5
0.4
0.3
0.2
0.1
0.0
Average R2
pdg
k67
ban
k68
k77
fja
k34
fns
site

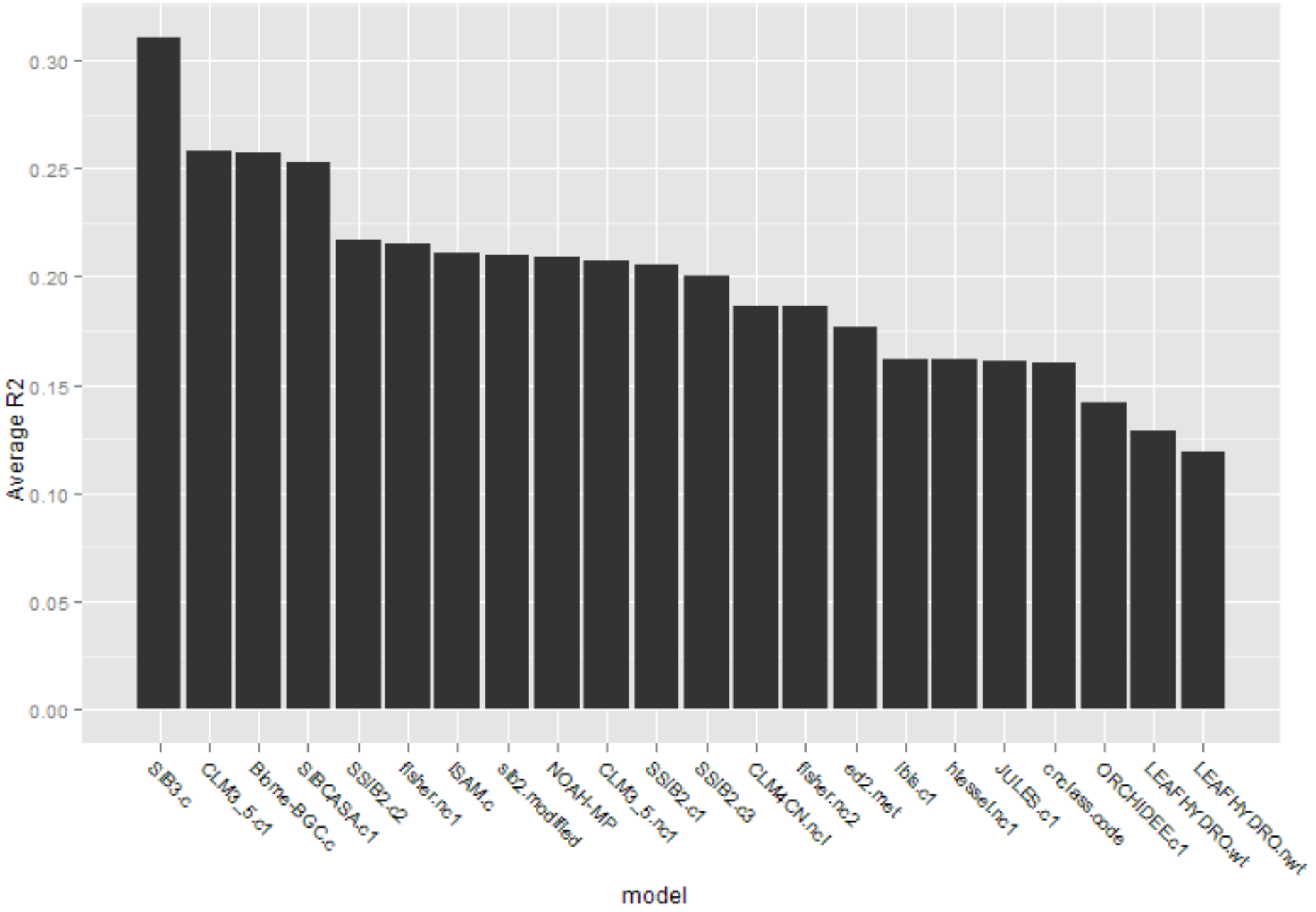

0.6
0.5
0.4
0.3
0.2
0.1
0.0
Average R2
pdg
k67
ban
k83
k77
fja
k34
fns
site
Most difficult site?

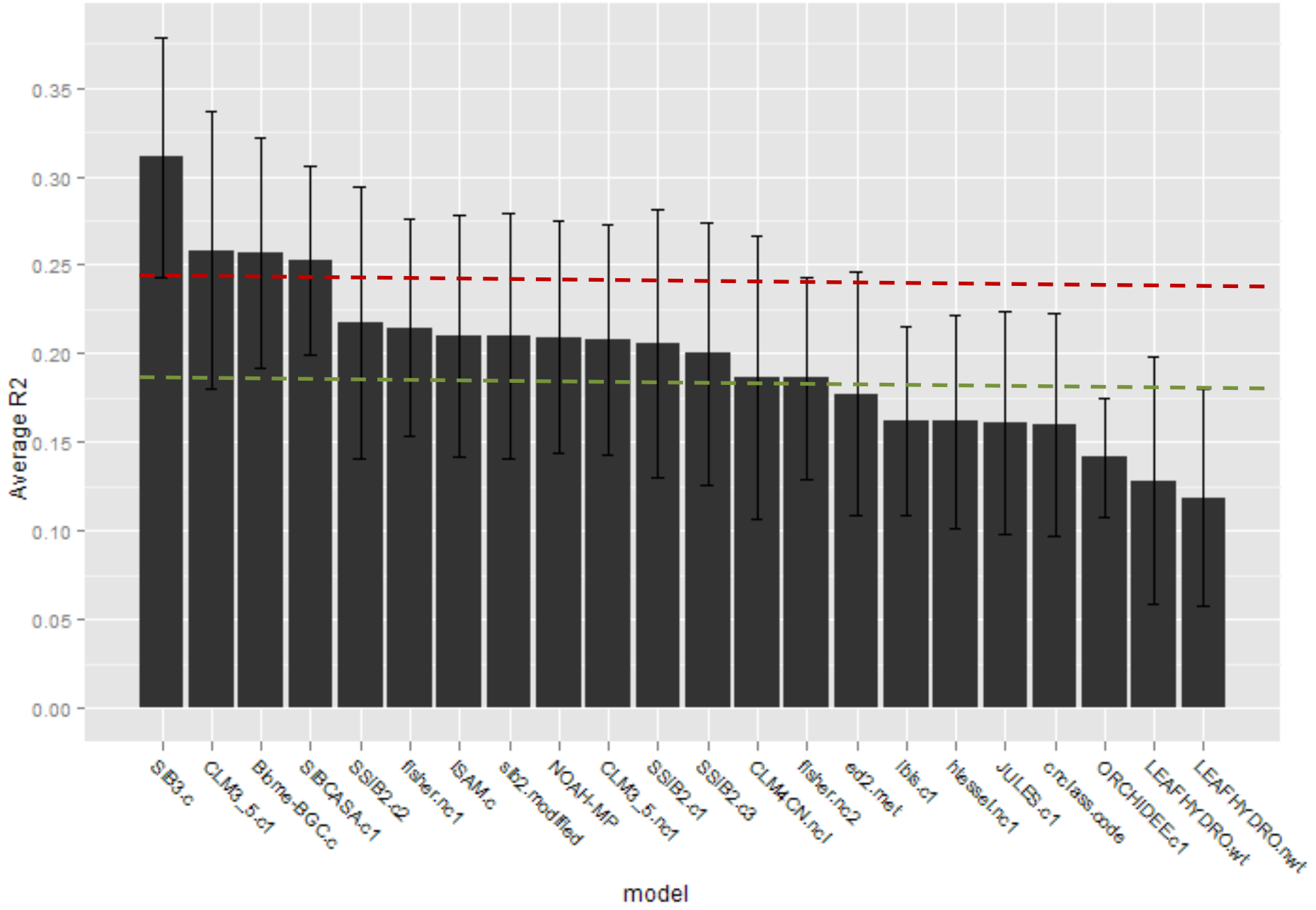

Average MBE (normalized by per-site means of data)
0.1
0.0
-0.1
-0.2
-0.3
-0.4
fisher.rc1
SIB3.c
JULES.c1
hlessel.nc1
ORCHIDEE.c1
CLM3_5.nc1
SIBCASA.c1
CLM3_5.c1
s62.modified
NOAH-MP
SSIB2.c1
SSIB2.c3
SSIB2.c2
fisher.rc2
CLM4CN.nc1
LEAFHYDRO.wt
LEAFHYDRO.nwt
ISAM.c
Biome-BGC.c
ed2.met
lbls.c1
crc.lass.code
model


Average MBE (normalized by per-site means of data)
0.1
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
fisher.rc1
SiB3.c
JULES.c1
hlesselnc1
ORCHIDEE.c1
CLM3_5.nc1
SiBCASA.c1
CLM3_5.c1
sib2.modified
NOAH-MP
SSiB2.c1
SSiB2.c3
SSiB2.c2
fisher.rc2
CLM4CN.nc1
LEAFHYDRO.wt
LEAFHYDRO.nwt
ISAM.c
Biome-BGC.c
ed2.met
lbls.c1
crc.lass.code
model


Average RMSE (normalized by per-site standard errors of data)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
SiB3.c
fisher.nc1
CLM3_5.nc1
SiBCASA.c1
JULES.c1
sib2.modified
CLM3_5.c1
NOAH-MP
hlesselnc1
ORCHIDEE.c1
CLM4CN.nc1
fisher.nc2
ISAM.c
ed2.met
LEAFHYDRO.wt
Biome-BGC.c
SSiB2.c1
SSiB2.c3
SSiB2.c2
LEAFHYDRO.nwt
crtclass.code
lbls.c1
model


Average RMSE (normalized by per-site standard errors of data)
0
1
2
3
4
SiB3.c
fisher.nc1
CLM3_5.nc1
SiBCASA.c1
JULES.c1
sib2.modified
CLM3_5.c1
NOAH-MP
htessel.nc1
ORCHIDEE.c1
CLM4CN.nc1
fisher.nc2
ISAM.c
ed2.met
LEAFHYDRO.wt
Biome-BGC.c
SSiB2.c1
SSiB2.c3
SSiB2.c2
LEAFHYDRO.nwt
crtclass.code
ibis.c1
model


Average R2
0.30
0.25
0.20
0.15
0.10
0.05
0.00
SiB3.c
CLM3_5.c1
Biome-BGC.c
SiBCASA.c1
SSiB2.c2
fisher.nc1
ISAM.c
sib2.modified
NOAH-MP
CLM3_5.nc1
SSiB2.c1
SSiB2.c3
CLM4CN.nc1
fisher.nc2
ed2.met
lpjs.c1
hesse.nc1
JULES.c1
crc.lass.code
ORCHIDEE.c1
LEAFHYDRO.wt
LEAFHYDRO.nwt
model


Average R2
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
SiB3.c
CLM3_5.c1
Biome-BGC.c
SiBCASA.c1
SSiB2.c2
fisher.nc1
ISAM.c
sib2.modified
NOAH-MP
CLM3_5.nc1
SSiB2.c1
SSiB2.c3
CLM4CN.nc1
fisher.nc2
ed2.met
lbls.c1
hesse1.nc1
JULES.c1
crc1ass.code
ORCHIDEE.c1
LEAFHYDRO.wt
LEAFHYDRO.nwt
model

Discussion

- SiB3 emerges as overall ‘strongest performer,’ though with a large degree of variation.
- Models were in the right range at the FNS site, but lacked agreement with the observed data as measured by R^2.
- PDG emerges at the best overall site, but with a tendency toward model under—prediction.
- Seasonality dominates performance metrics much more at some sites than others.
- Observed data is very stochastic for some sites—how confident are we that these values are ‘true’ as opposed to noisy?
- No model or site is the clear `winner’—at this point standard errors are too large.
- What is observed data uncertainty? May need to think about whether predictions fall within error range of observations.

Next steps…

- Protocol for error attribution—a mix of summary stats, OLS, and machine learning algorithms

Thanks!
Questions?
Suggestions?