

Philippe Crochet

A Linear Model for Mapping Precipitation in Iceland

I) Introduction

Estimating the amounts of precipitation is of paramount importance for number of applications such as agricultural purposes, water management, hydrological models, climatic studies, global circulation models and for the validation of indirect estimation methods based on remote sensors such as radars or satellites and numerical prediction models.

The most common way to estimate ground precipitation is to interpolate scattered observations. There is not a unique solution and various interpolation methods are available for this purpose (see [1] to [3] for instance). Most of the time, these methods operate in 2D and rely only on the information based on the precipitation field itself. They usually work well if the precipitation network is dense enough with respect to the spatial variability of the precipitation field. In a country like Iceland, the use of such interpolation methods alone can be questionable in mountainous regions where strong 3D variability can be expected or in poorly monitored areas such as the highlands. It is thus necessary to pre-process the data and attempt to describe and remove any trend or non-stationarities before interpolating, see [4] for instance, [5] and earlier work from [6] and [7] who successfully describe relationships between precipitation amounts and topography. In this report, we propose to follow this general idea and study the influence of the topographical environment on the precipitation patterns in Iceland. It is part of the validation protocol developed for assessing the quality of the precipitation amounts produced by the MM5 numerical model in Iceland [8] and [9].

II) Relationship between precipitation and topographical environment

II-1) The predictors

The precipitation at each site is described by a linear model. The 9 selected predictors (p1) to (p9) give an information about the geographical and topographical features in the vicinity of each site.

$$P(x, y, k) = a_0 + a_1 x + a_2 y + a_3 d_{\min} + a_4 \bar{h} + a_5 \bar{a} + a_6 \bar{s} + a_7 \sigma_h + a_8 \sigma_a + a_9 w \quad (1)$$

Where:

$P(x, y, k)$ = Precipitation at location (x,y) and time k

and

(p0): a_0 = intercept term

(p1): x = x coordinate (in Lambert Conformal)

(p2): y = y coordinate (in Lambert Conformal)

- (p3): d_{\min} = shortest distance to the ocean in km
- (p4): \bar{h} = average elevation (in metre) in the vicinity of point (x,y)
- (p5): \bar{a} = average orientation of the hillslope (in degree) in the vicinity of point (x,y).
- (p6): \bar{s} = average hillslope (in %) in the vicinity of point (x,y)
- (p7): σ_h = standard-deviation of the elevation in the vicinity of point (x,y)
- (p8): σ_a = standard-deviation of the hillslope orientation in the vicinity of point (x,y)
- (p9): w = difference in metre between the highest elevation in the vicinity of point (x,y) and elevation at point (x,y).

The predictors (p1) and (p2) inform about the spatial location of the site in the plane, while predictor (p3) describes the location more dynamically, with respect to the Icelandic shoreline and accounts for the maritime influence on precipitation. Predictor (p4) describes the effect of the elevation around the site. Predictor (p5) integrates information about how exposed the site is in order to account for the dominant direction of the weather systems and possible shelter effects. Predictors (p6) to (p9) attempt to describe the complexity of the topographical environment and the degree of geographical isolation of the site being considered.

The 7 predictors (p3) to (p9) are estimated for each site once for all with the use of a digital elevation model (DEM) (figure 1). First, the original DEM (resolution of about 1 km) is resampled to a coarser resolution of about 2 km for computational purposes. From this coarser DEM, the elevation, hillslope and orientation are extracted at each grid point (figure 2). In the model presented here, the hillslope is taken positive downhill. The orientation (-180° , 180°) is positive clockwise from north (y-axis) towards east (x-axis) and negative anti-clockwise from north toward west. Then, the 7 predictors are estimated by considering a window of 5 km around each site. Further improvements in this first generation model will be considered later.

Figure 1: Digital Elevation Model

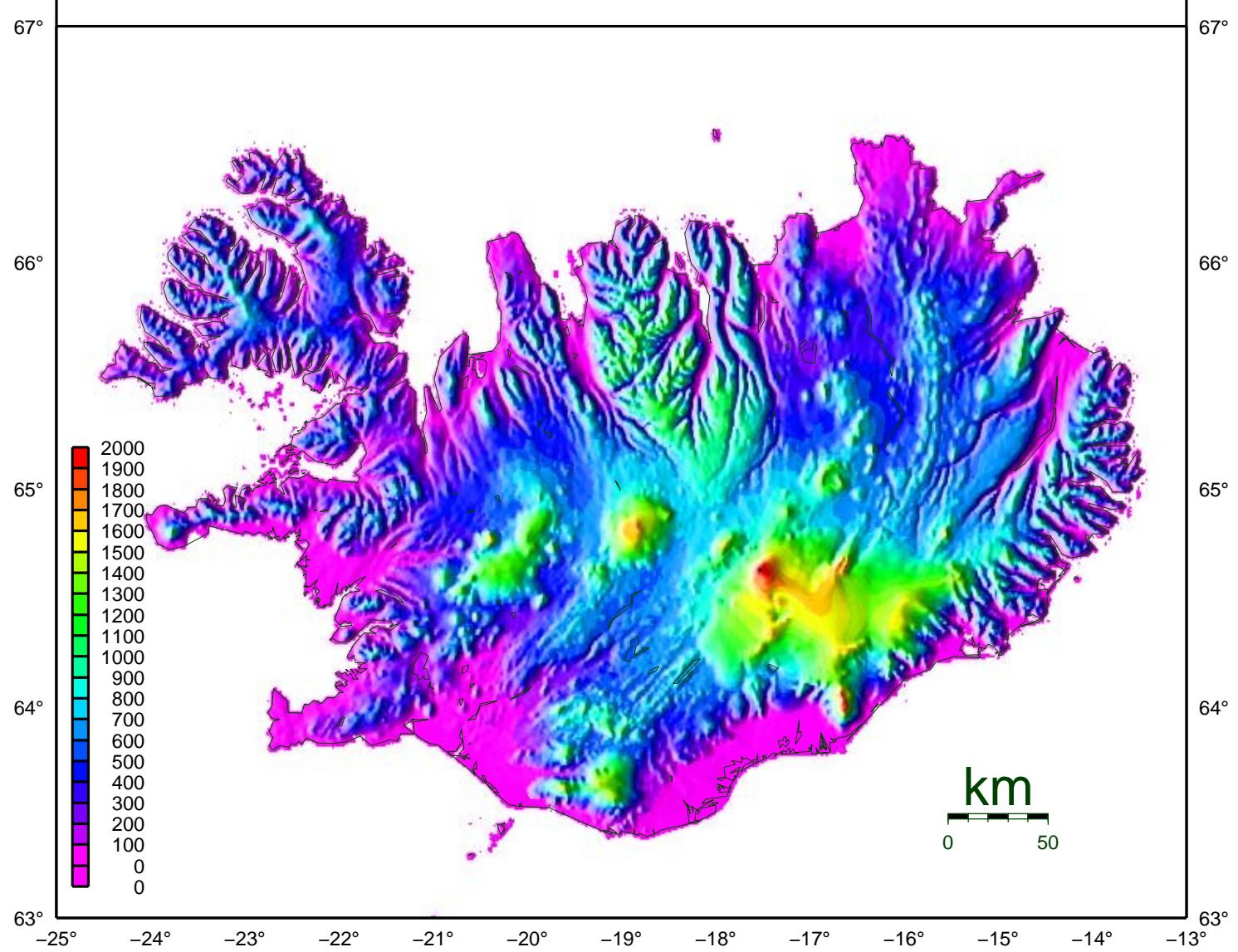
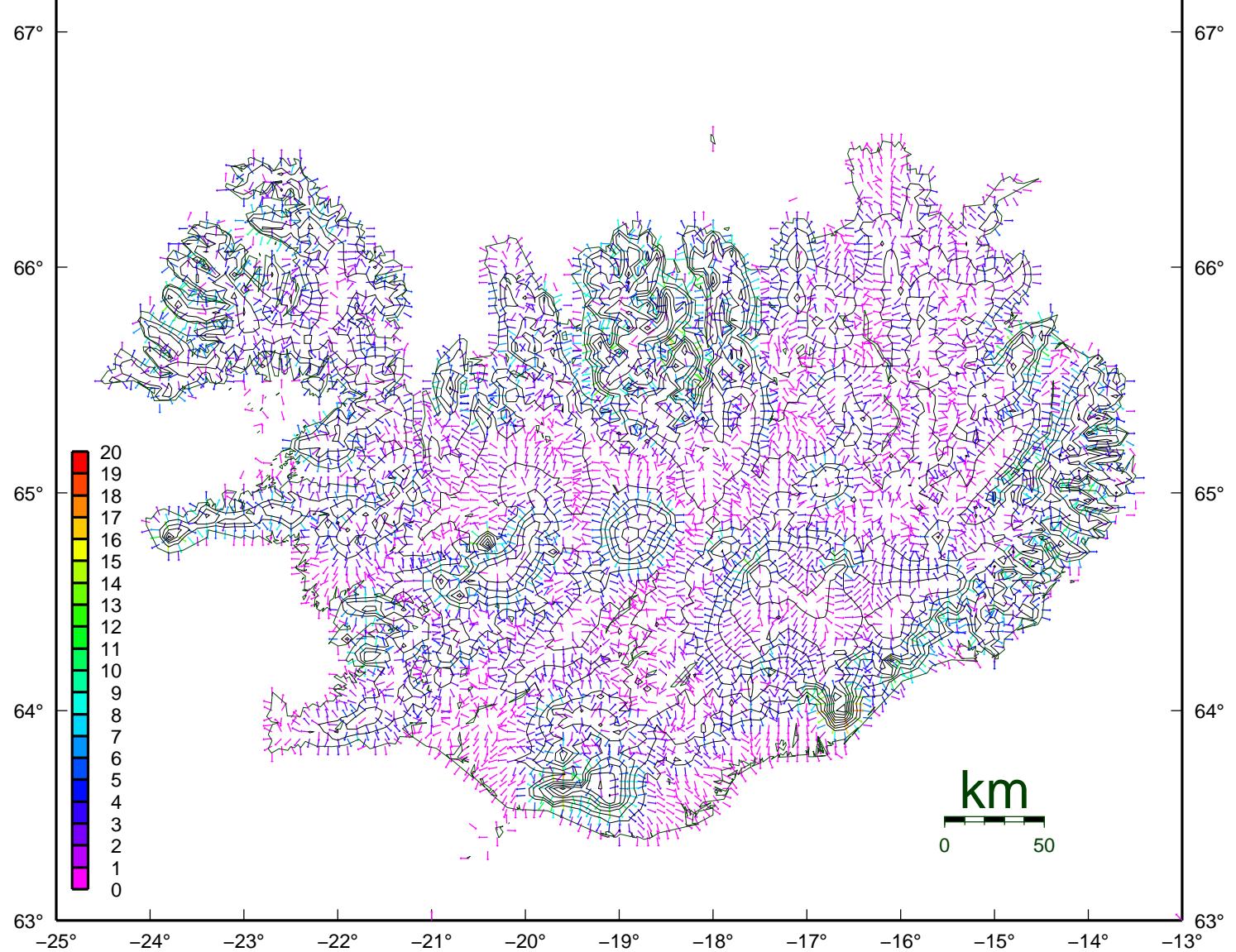


Figure 2: Directional gradient of hillslopes



II-2) Operational estimation procedure

The 9 predictors are not expected to have the same effect on the precipitation amounts all over Iceland nor from time to time. For this reason, the following procedure is defined and applied individually at each time step k:

The regression coefficients are estimated by a multiple linear regression (MLR). A set of 5 successive MLR is made over 3 sectors of 120° each, centered on the top of Hofsjökull glacier (figure 3). There is a rotation of 10° to 15° between one set to the next. There are several reasons that motivate this choice. As a matter of fact, we do not know for sure the best way to divide Iceland into homogeneous regions, and by making a set of 5 successive estimates, we expect the final estimation to be closer to the true value. The reason for dividing Iceland into 3 sectors is a compromise between having a reasonable number of stations per sector (greater than the number of predictors) and the definition of what we expect are homogeneous regions, i.e. regions under the influence of the same weather systems. Finally, in using this successive estimation method for mapping, we expect to avoid unpleasant sharp edge effects between the different sectors. This estimation procedure is not definitive and modifications might be considered later.

For each set, the method produces 3 series of regression coefficients, and the precipitation at site j is estimated by:

$$\begin{aligned} P^*(x_j, y_j, k, r) = & a_0(k, s[r, l]) + a_1(k, s[r, l])x_j + a_2(k, s[r, l])y_j + \\ & a_3(k, s[r, l])d \min_j + a_4(k, s[r, l])\bar{h}_j + a_5(k, s[r, l])\bar{a}_j + a_6(k, s[r, l])\bar{s}_j + \\ & a_7(k, s[r, l])\sigma_{h_j} + a_8(k, s[r, l])\sigma_{a_j} + a_9(k, s[r, l])w_j \end{aligned} \quad (2)$$

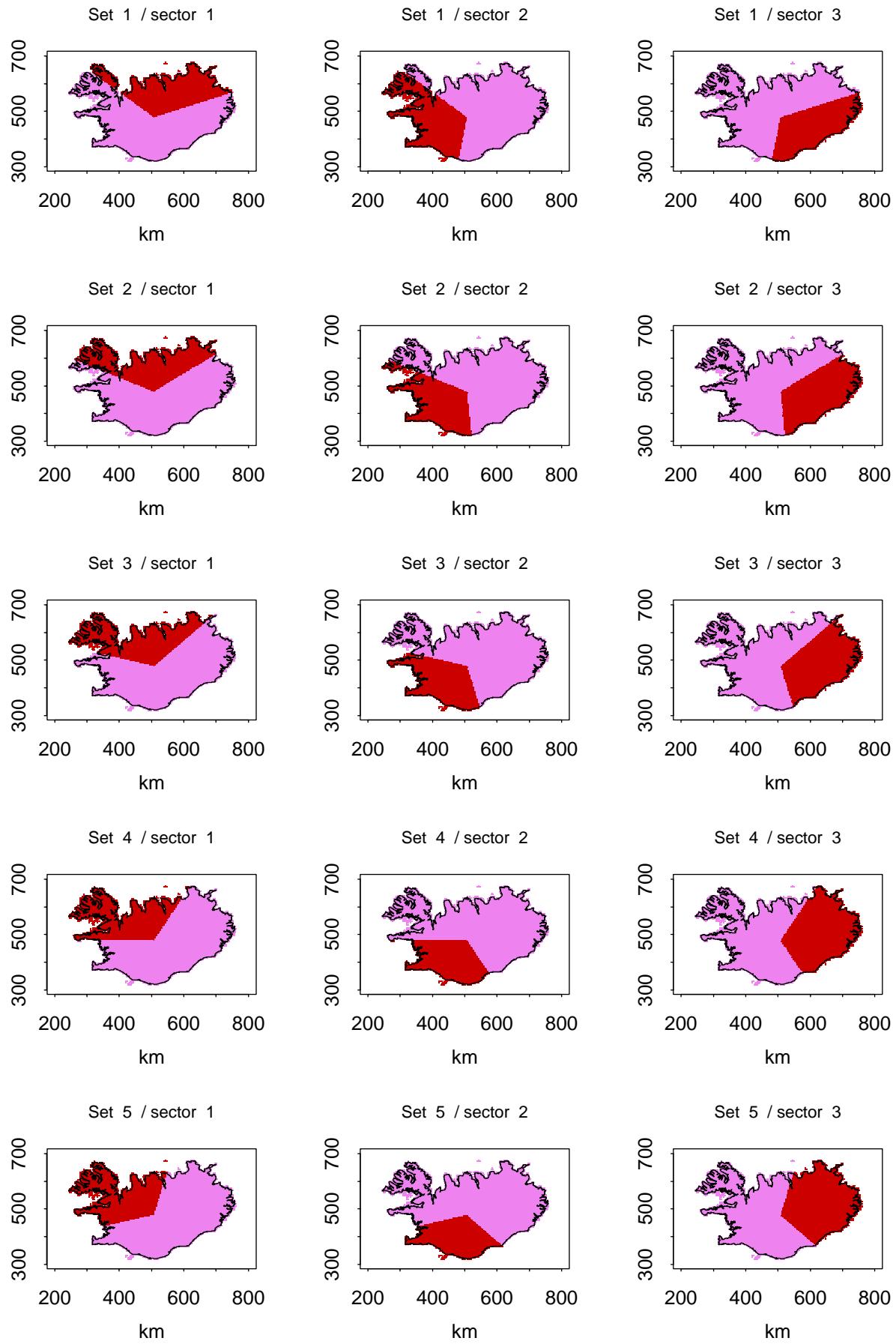
Where

$s[r, l]$ is the l^{th} sector $[1 \leq l \leq 3]$ of the r^{th} set $[1 \leq r \leq 5]$.

A given station belongs only to one sector at the time, and the final estimation is given by the mean of the 5 successive estimates:

$$P^*(x_j, y_j, k) = \frac{1}{5} \sum_{r=1}^5 P^*(x_j, y_j, k, r) \quad (3)$$

Figure 3: The different sectors



III) Evaluation of the linear model in Iceland

III-1) The data

The data used in this study are daily precipitation measured at 122 sites from 1980 to 2001 (figure 4). These precipitation stations are manual. No correction is applied to account for any wind loss for instance. The monthly precipitation is computed if the station has been in operation for at least 25 days, and the annual precipitation is computed if the station has been in operation for at least 350 days. The number of stations used simultaneously in a given month is ranging from 60 to 122, with a mean of 84.

III-2) Annual precipitation

The model was evaluated for each year between 1980 to 2001. First the regression coefficients were estimated with the entire network and the model performance was judged by plotting the relationship between the estimated values $P^*(x, y, k)$ and observed values $P(x, y, k)$. Figure 5 presents the scatter plots between $P^*(x, y, k)$ and $P(x, y, k)$ for the year 2001. Appendix 1 gives the scatter plots for the period 1980-2000. The linearity of the relationship is good. The annual estimates are unbiased in average for each year (figure 6). The correlation coefficients are ranging between 0.903 and 0.976 with an average value of 0.947. However, systematic errors can occur punctually for some stations where the estimate can be systematically too high or too low (see appendix 2).

Figure 7 presents the histograms of the different regression coefficients. Large variations in both sign and magnitude are observed. This supports the previous assumption that there is not a unique relationship between the predictors and precipitation in both space and time. A numerical example is given with the annual precipitation in 1999 and 2000 (tables 1 and 2). The stepwise selection procedure (tables 3 and 4) gives the rank of the variable that gives the largest reduction of the residual sum of squares. It is obvious that a given predictor does not influence the estimation in the same manner, both in space and time. The results tend to show that the topography "captures" long term accumulated precipitation features but the complexity of the relationships cannot be easily described by one single model.

Figure 4: precipitation network

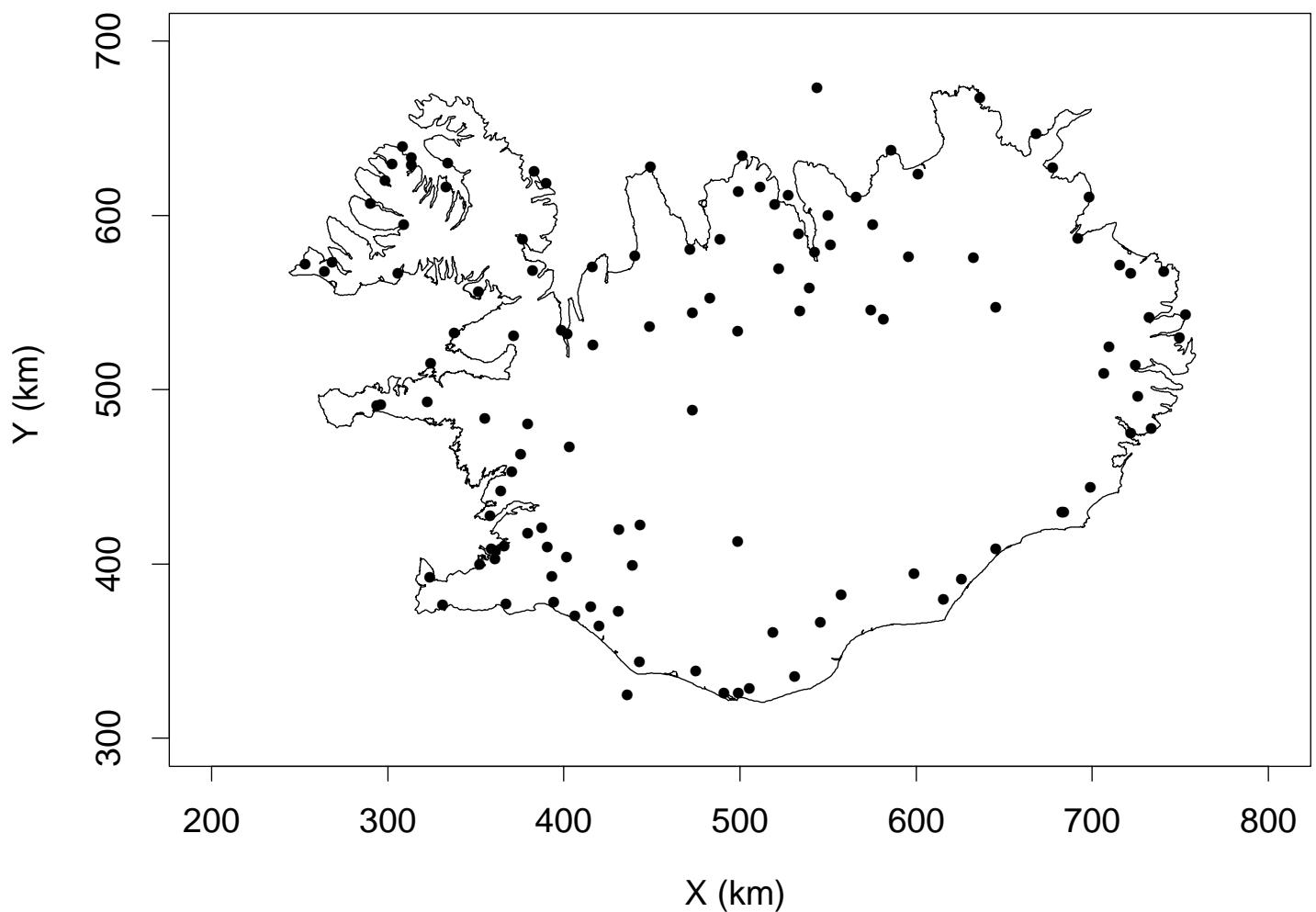


Figure 5: Annual precipitation in 2001

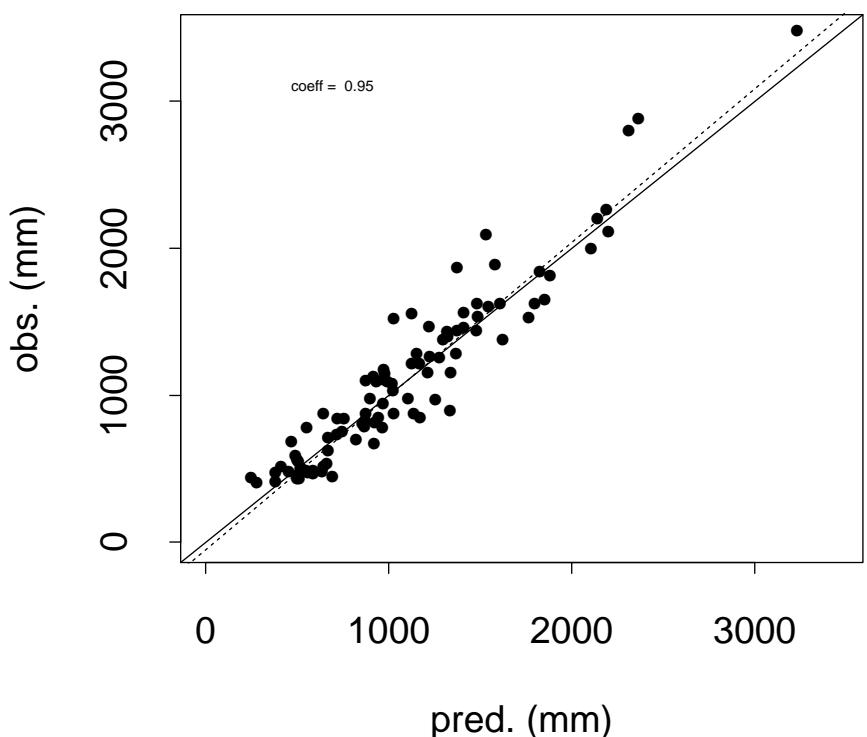


Figure 6: Annual precipitation : mean error

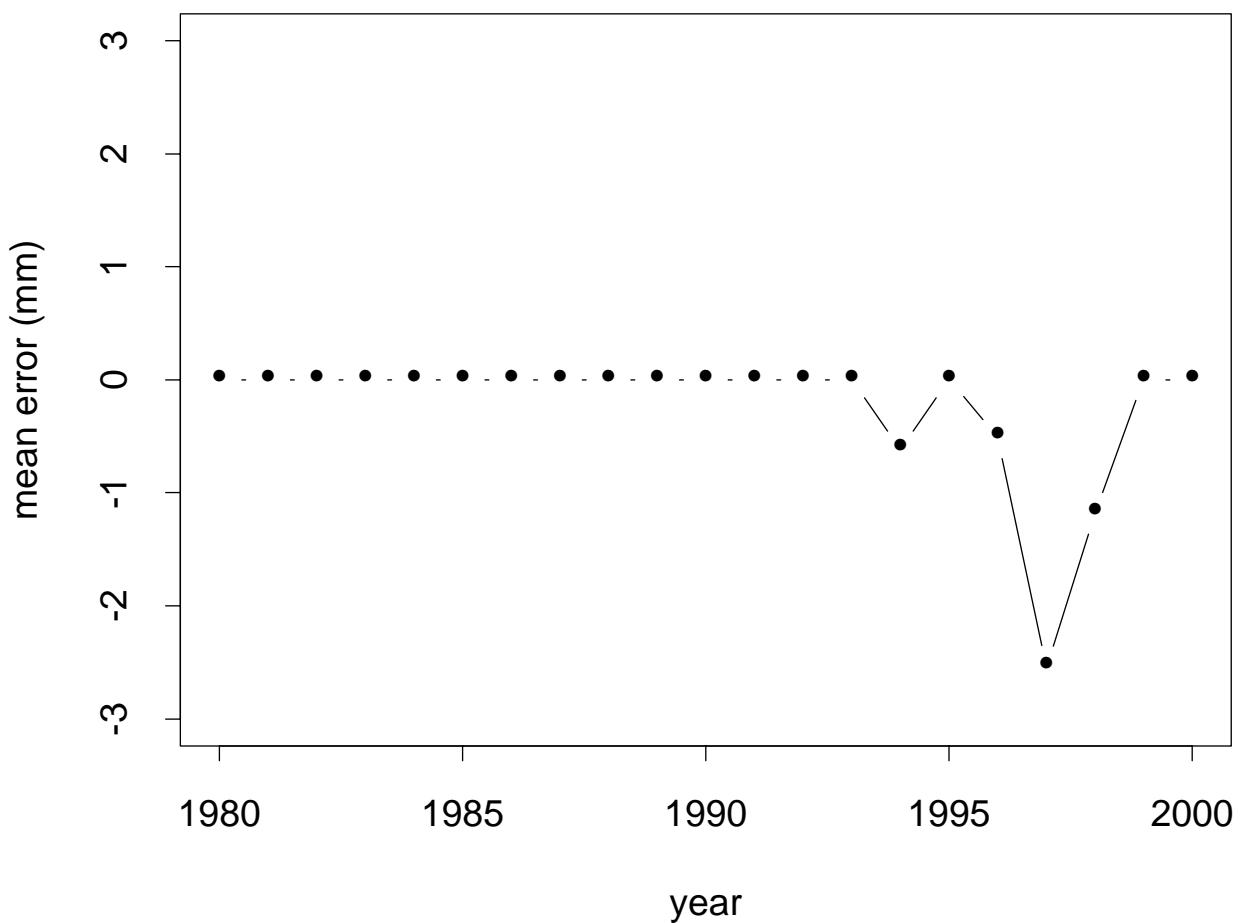
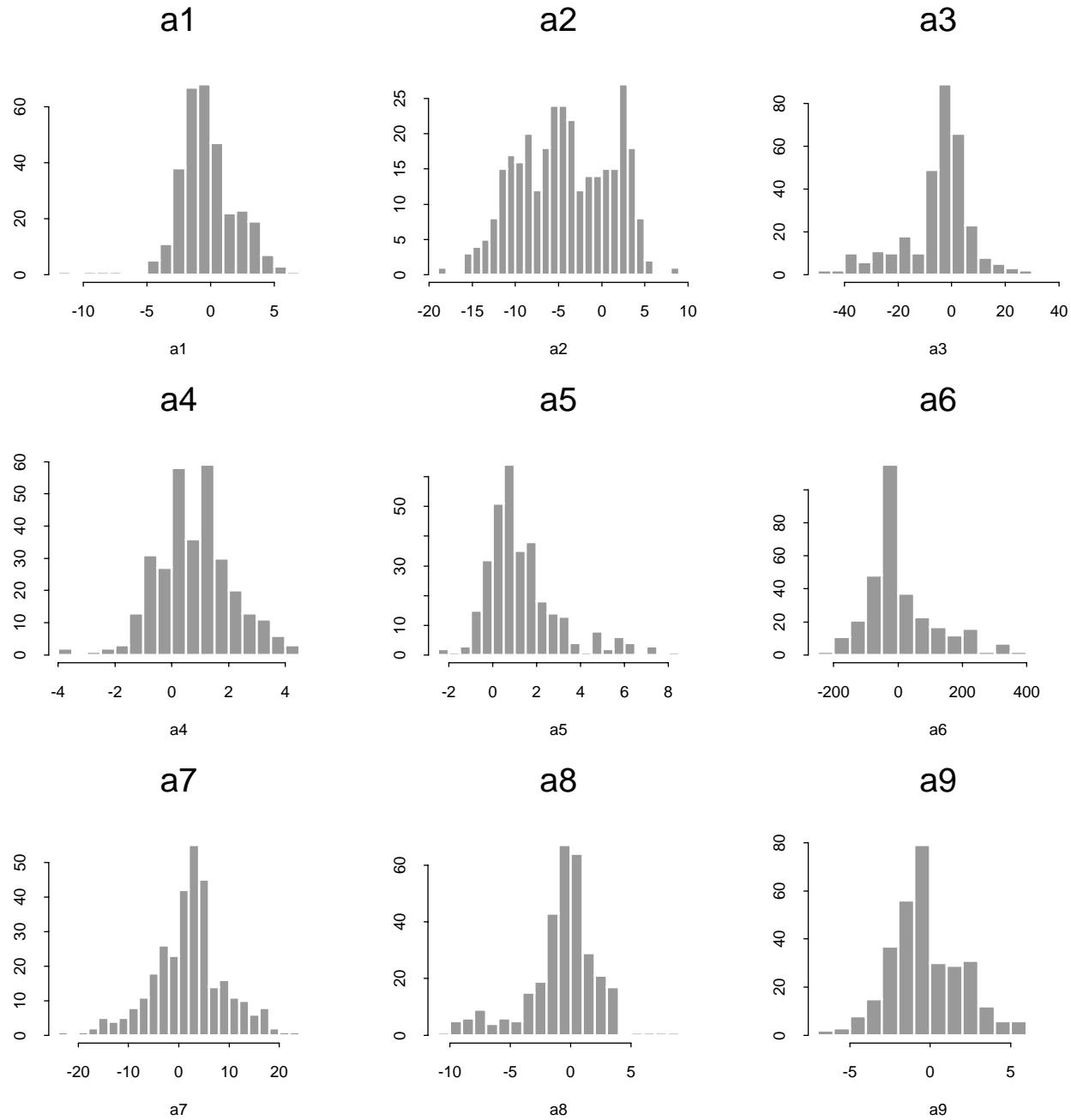


Figure 7: Annual precipitation (1980-2000): Histograms of the regression coefficients



In order to assess the quality of the model in a more consistent manner, a validation procedure was considered by defining two independent samples (figure 8):

- A calibration sample (95 stations) was used to estimate the model coefficients.
- A reference sample (27 stations independent from the calibration sample) was used to estimate precipitation with the model defined with the calibration sample.

The scatter plots between $P^*(x_{ref}, y_{ref}, k)$ and $P(x_{ref}, y_{ref}, k)$ for the years 1980 to 2000 are given in appendix 3. As expected, the correlation coefficients are not as high as those estimated in the first place, i.e. when the calibration and validation are made with the entire network. These results represent somehow the lower limit of what can be expected in reality when using the model for mapping purposes in poorly monitored areas. The entire network hardly reaches 122 stations in total and any reduction in this number will most likely have a dramatic consequence on the estimation of the model coefficients and the robustness of the method. When using the method for mapping purposes, all the available information will be used and any additional information such as ice core data for instance will be of great value.

Figure 8 : Validation network (red) and calibration network (black)

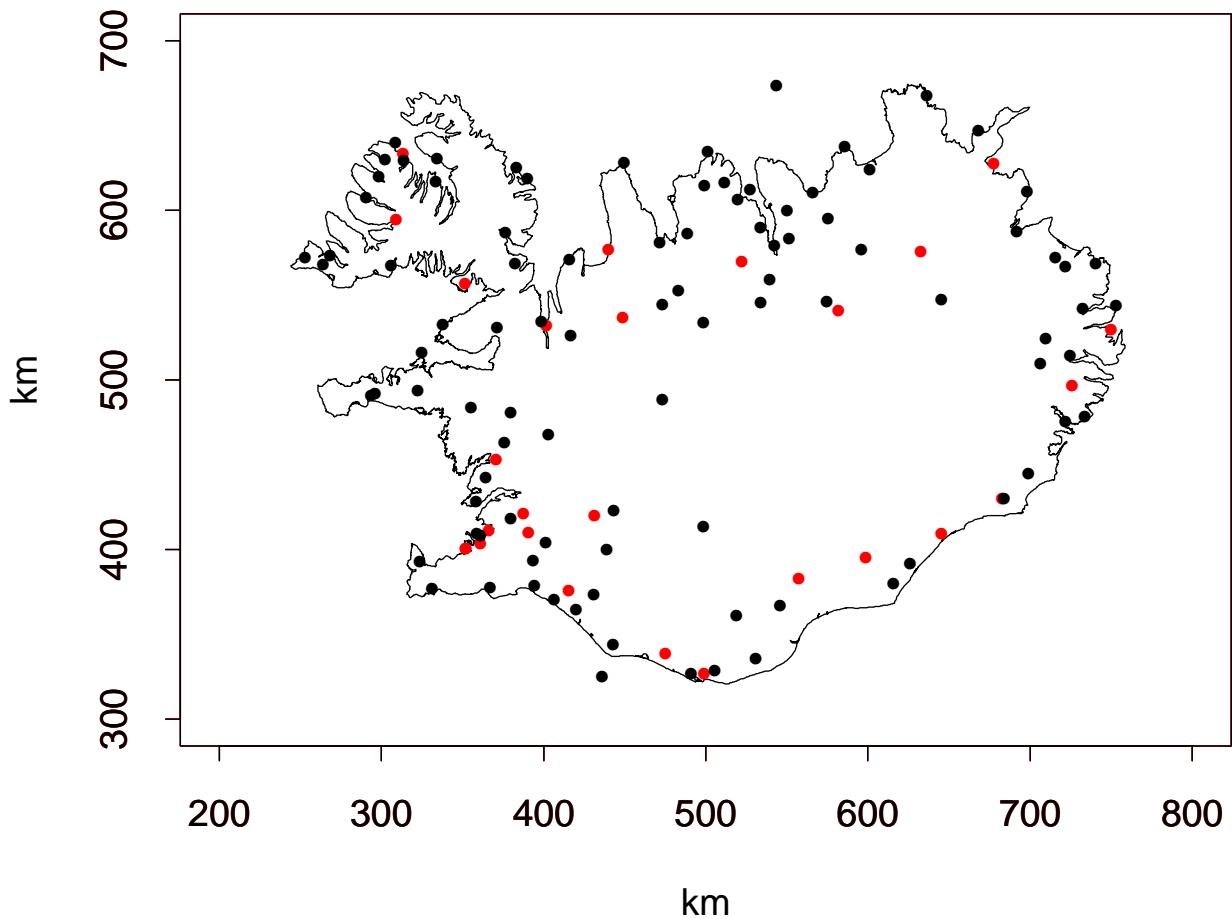


Table 1: Year 1999: regression coefficients for the 3rd set

	a_0	$a1$	$a2$	$a3$	$a4$	$a5$	a_6	$a7$	$a8$	$a9$
S1	-108	-1.24	2.36	-8.68	0.84	0.115	-6.11	-0.648	0.304	-0.02
S2	2173	2.47	-5.12	-2.79	0.16	1.57	51.47	4.47	-0.002	-0.695
S3	3759	-0.143	-5.39	-19.6	1.91	1.77	-3.95	9	-1.26	-3.48

Table 2: Year 2000: regression coefficients for the 3rd set

	a_0	$a1$	$a2$	$a3$	$a4$	$a5$	a_6	$a7$	$a8$	$a9$
S1	-230	-0.802	2.09	-11.41	1.08	0.155	7.84	-0.003	0.598	-0.448
S2	1410	3.61	-4.47	-8.83	0.135	1.566	94.18	2.27	0.259	-0.019
S3	1199	2.15	-3.28	-15.16	1.483	3.2	48.38	5.88	-1.23	-2.83

Table 3: Year 1999: rank of predictors in the stepwise selection for the 3rd set

	1	2	3	4	5	6	7	8	9
S1	x	y	\bar{h}	$d \min$	σ_h	σ_a	\bar{s}	\bar{a}	w
S2	x	σ_h	y	\bar{a}	\bar{s}	$d \min$	w	\bar{h}	σ_a
S3	y	$d \min$	\bar{h}	\bar{a}	\bar{s}	w	σ_h	σ_a	x

Table 4: Year 2000: rank of predictors in the stepwise selection for the 3rd set

	1	2	3	4	5	6	7	8	9
S1	$d \min$	\bar{h}	x	y	w	σ_a	\bar{a}	\bar{s}	σ_h
S2	y	w	\bar{a}	\bar{s}	$d \min$	x	σ_h	σ_a	\bar{h}
S3	$d \min$	\bar{a}	\bar{s}	y	x	w	σ_h	\bar{h}	σ_a

III-3) Application to monthly precipitation

The model was evaluated for each month from 1980 to 2001. First, the entire network was used both to derive the regression coefficients and to check the model performances. Figure 9 presents the scatter plots between $P^*(x, y, k)$ and $P(x, y, k)$ for each month of the year 2001, and appendix 4 the scatter plots for the period 1980-2000. The linearity of the relationships between $P^*(x, y, k)$ and $P(x, y, k)$ is usually high. The correlation coefficients are ranging between 0.736 and 0.976 with an average value of 0.915 (figure 10). In average, the proposed model is able to produce unbiased estimates of monthly precipitation (figure 11), but systematic errors can occur punctually for some stations where the estimate can be systematically too high or too low (see appendix 5). Here too, the different regression coefficients display large spatio-temporal variations in both sign and magnitude (figure 12).

Then, the same validation procedure defined for the annual precipitation in section III-2 was considered. The scatter plots are presented in appendix 6. The relationships are usually rather linear but the performances are lower, as expected.

In conclusion, the developed model is quite suitable for estimating monthly precipitation in Iceland, but the quality of the estimation will depend on the number of available sites for calibrating the model parameters.

Figure 9: Monthly precipitation in 2001

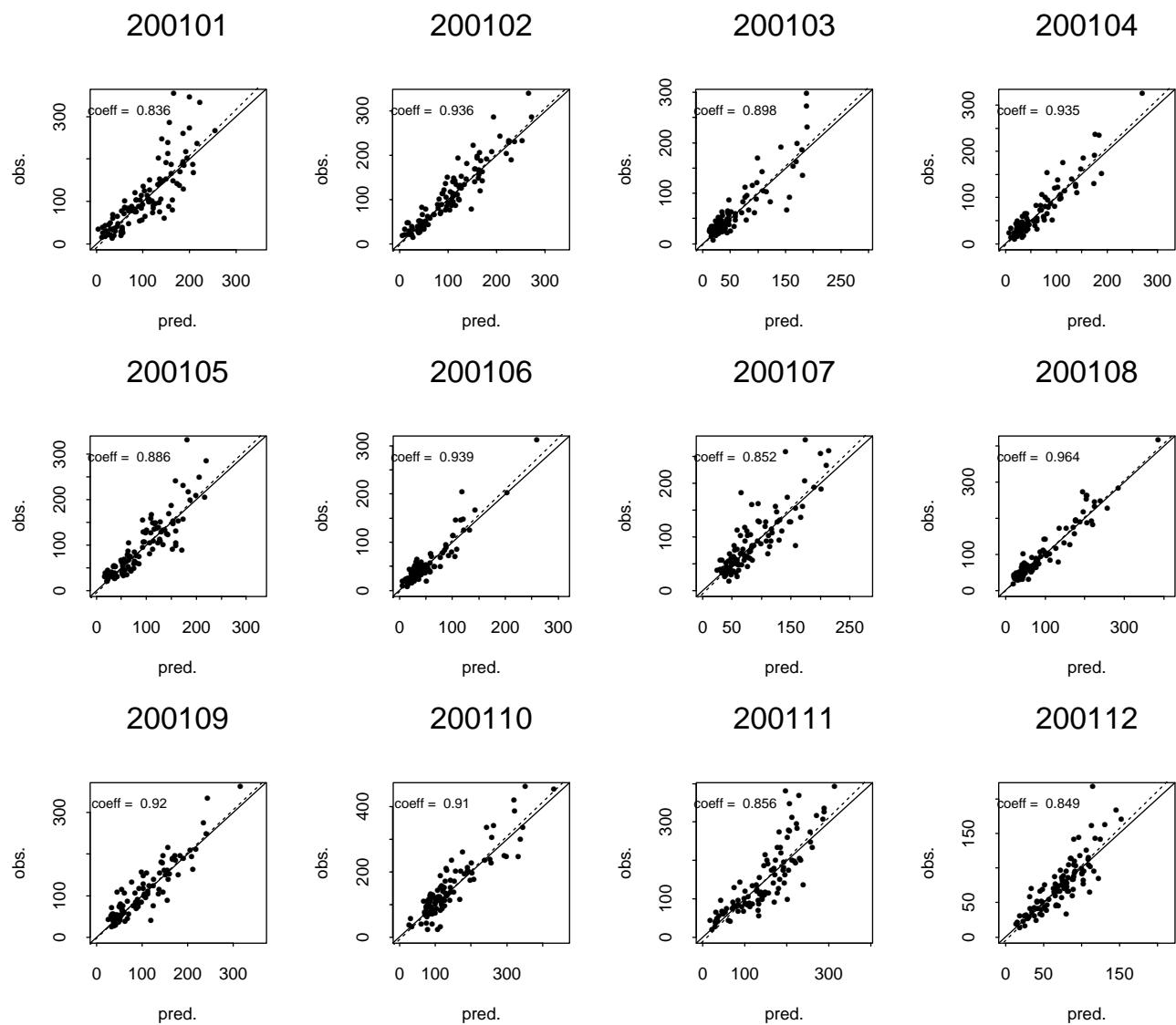


Figure 10: Monthly precipitation : correlation coefficients

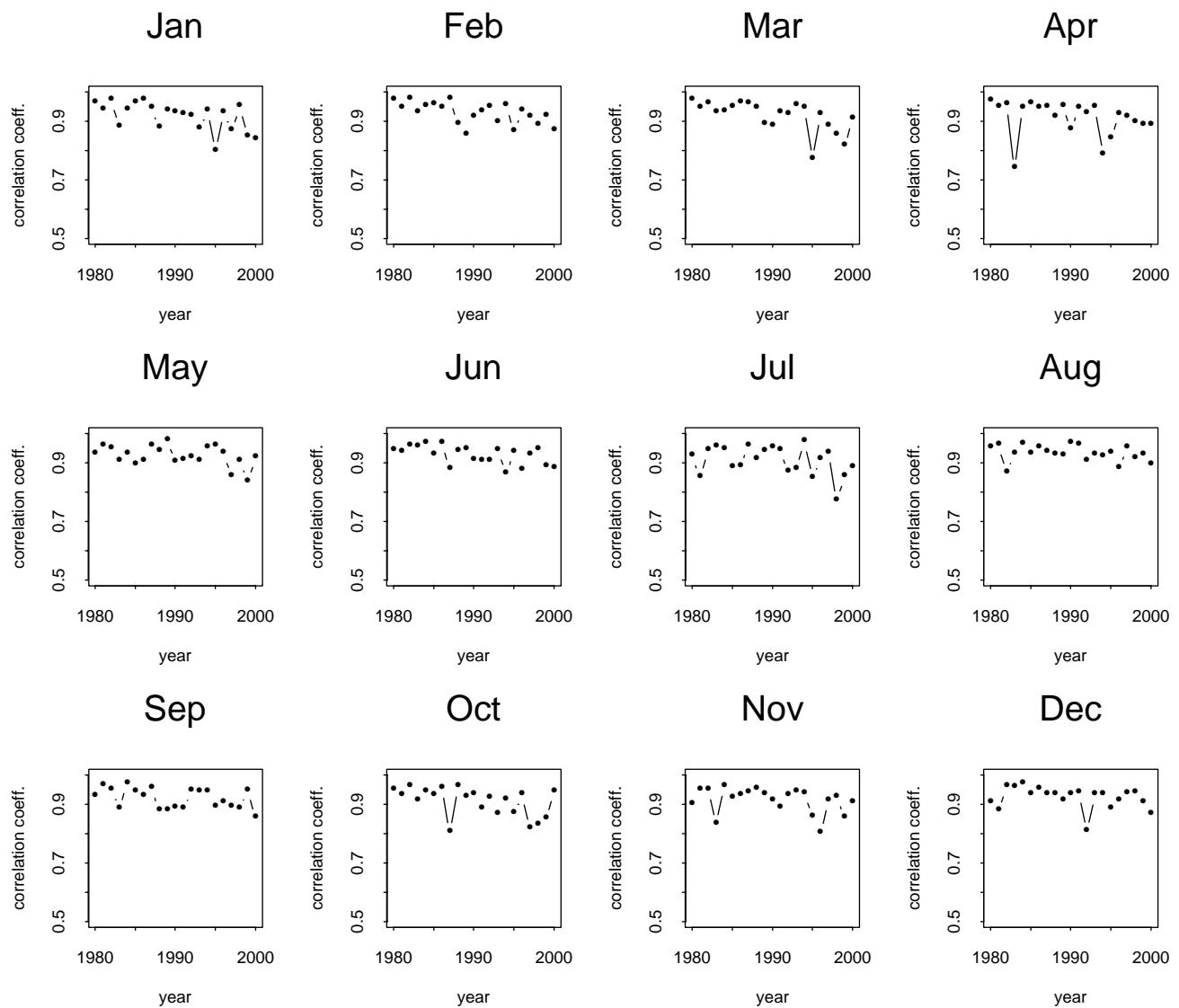


Figure 11: Monthly precipitation : mean error

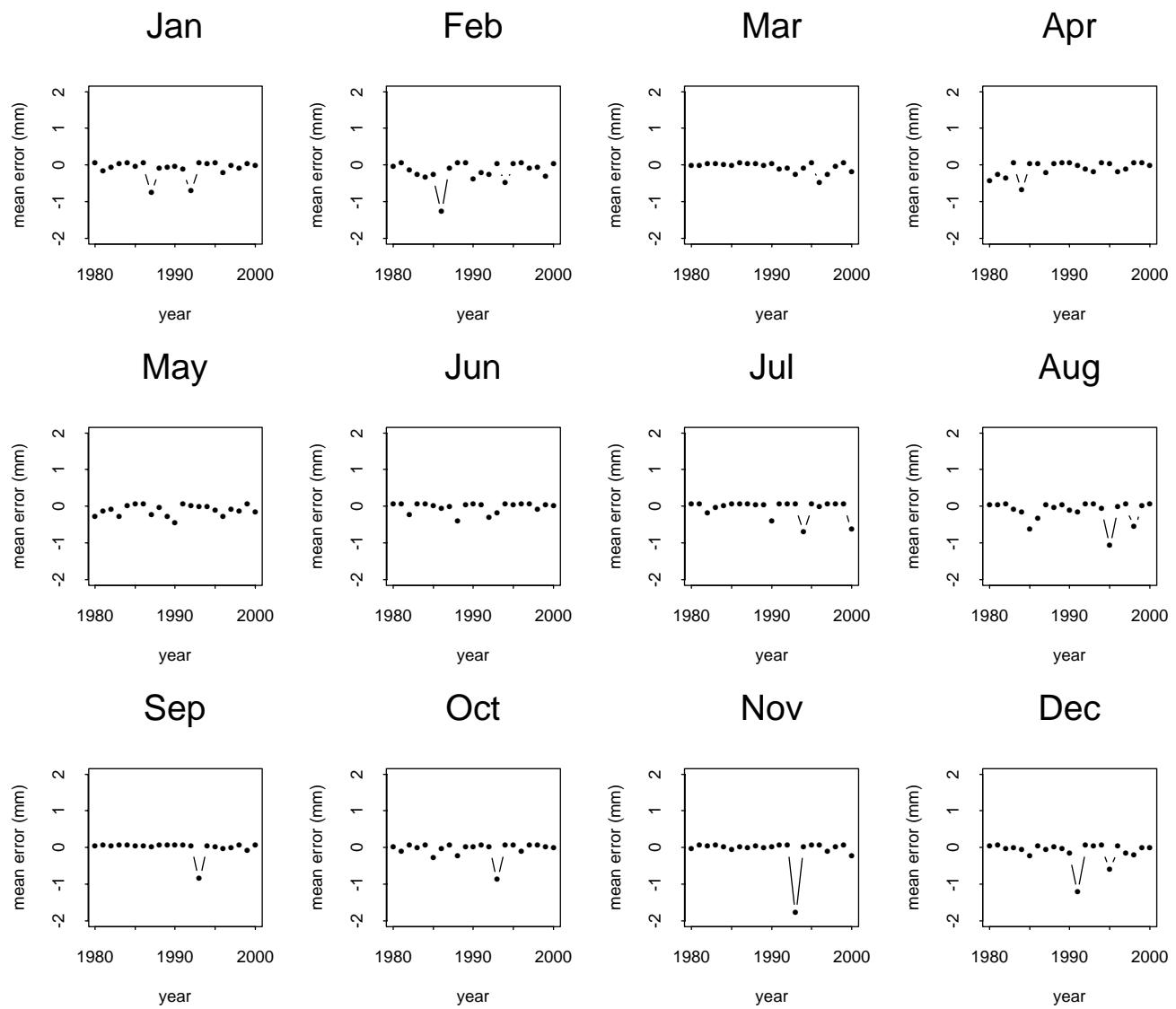
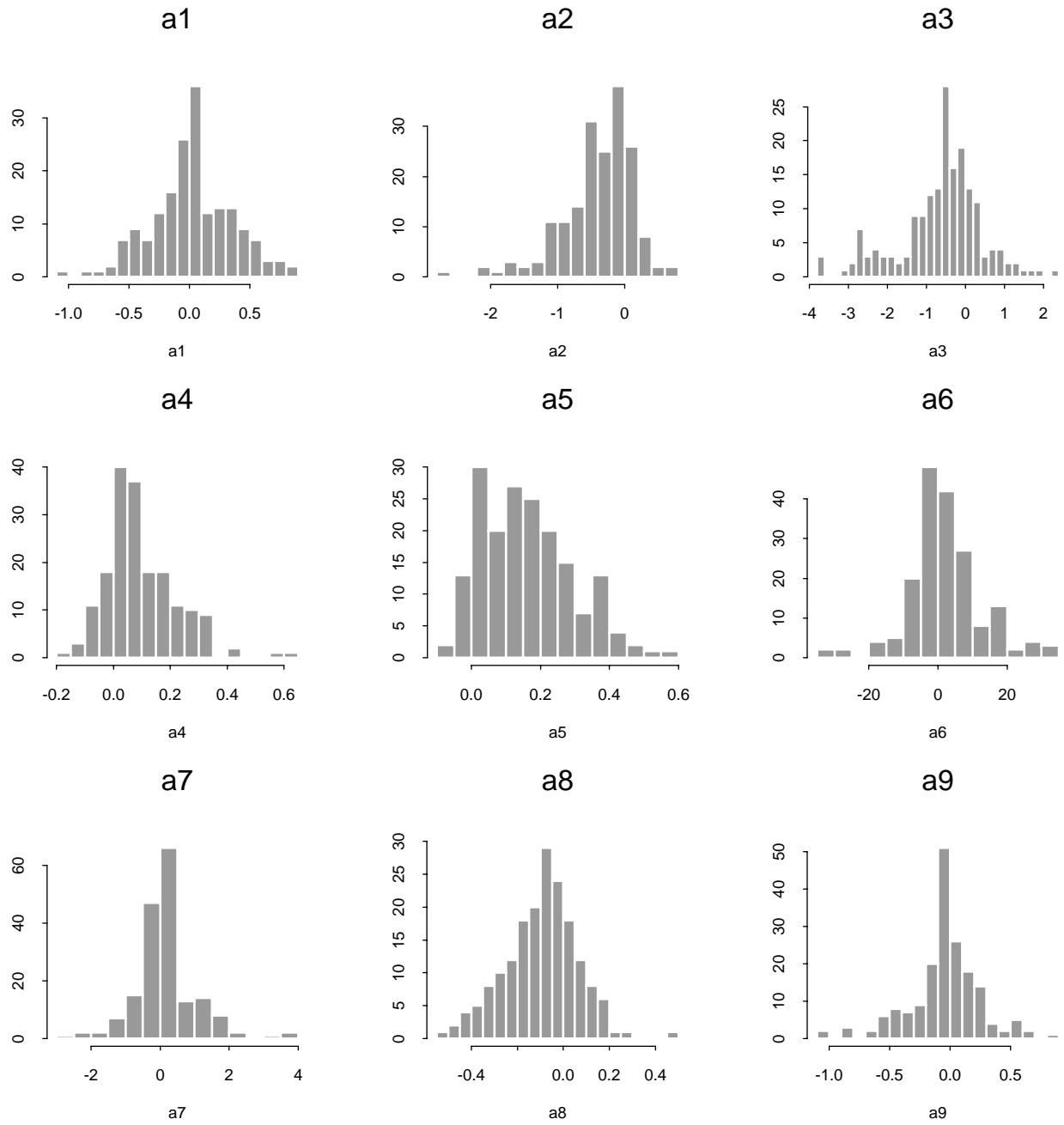


Figure 12: Monthly precipitation in 2001: Histogram of the regression coefficients



III-4) Spatial structure of residuals

The residuals of the regression are defined as follows:

$$e(x, y, k) = P(x, y, k) - P^*(x, y, k) \quad (4)$$

If the residuals exhibit any spatial structure, then an interpolation procedure can be considered. The spatial structure of the residuals is defined by the climatological semi-variogram, see [10] for instance.

First a scaled residual is defined for each field k:

$$e_s(X, k) = \frac{e(x, y, k)}{\sigma_k[e(x, y, k)]} \quad (5)$$

where

X denotes the spatial location (x,y)

and

$\sigma_k[e(x, y, k)]$ denotes the spatial standard-deviation of $e(x, y, k)$ for the field k

The experimental climatological semi-variogram is computed as

$$\gamma(h) = \frac{1}{2m(h)} \sum_{k=1}^n \sum_{i=1}^{m(h,k)} (e_s(X_i, k) - e_s(X_i + h, k))^2 \quad (6)$$

where h is the separation distance between two stations i and j:

$$h = \|X_i - X_j\| \quad (7)$$

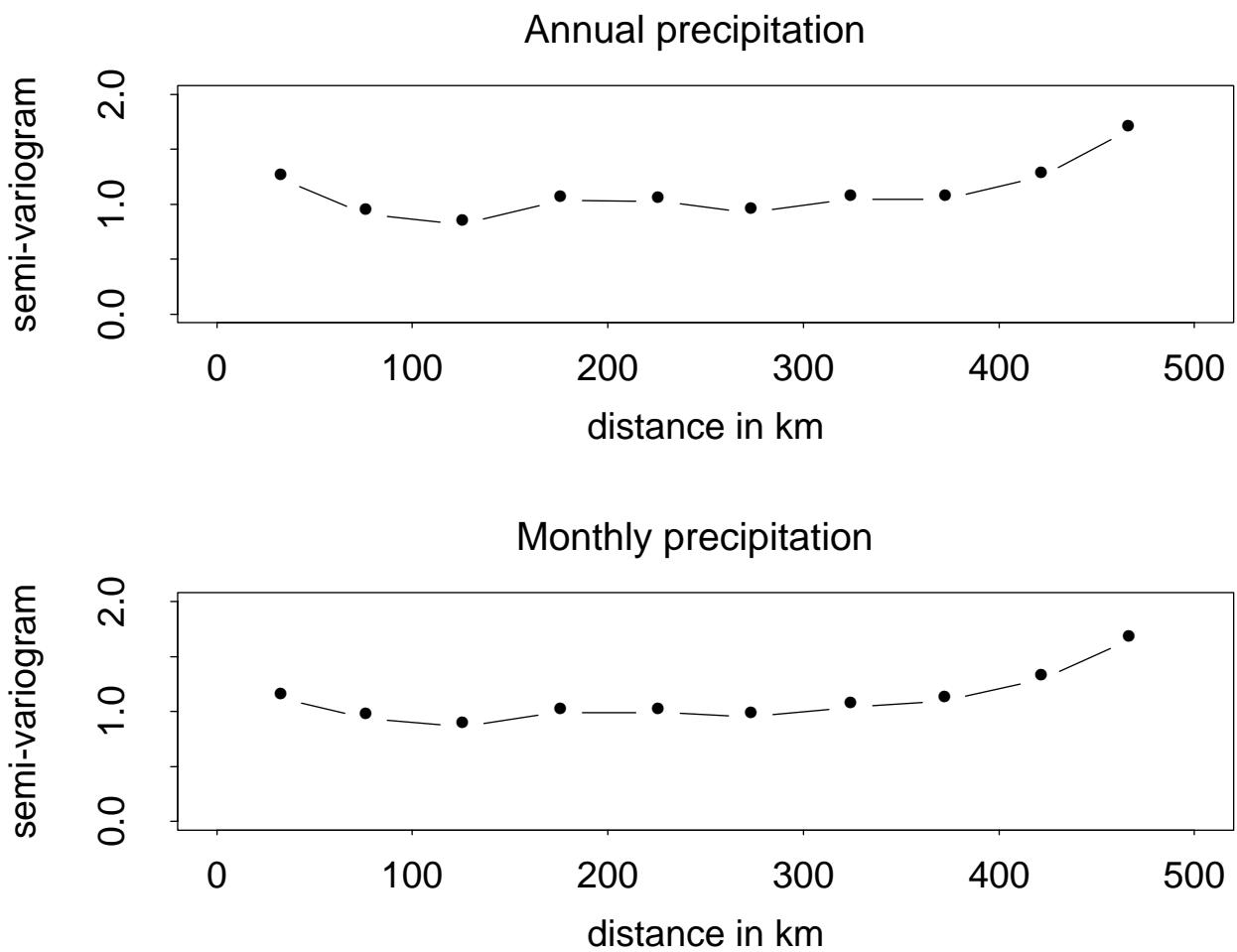
$m(h, k)$ is the number of pairs of stations having an inter-distance of $h \pm \Delta h$ for the field k, $m(h)$ is the total number of pairs of stations having an inter-distance of $h \pm \Delta h$, and n is the number of fields.

As figure 13 suggests, the monthly and annual residuals do not present any spatial structure and can be considered as decorrelated. Thus, the attempt to define an estimation procedure based on the kriging interpolation of the residuals:

$$P^k(x, y, k) = e^k(x, y, k) + P^*(x, y, k) \quad (8)$$

will not be considered in this study.

Figure 13: Climatological semi-variogram of residuals



III-5) Mapping precipitation in Iceland with the linear model

The linear model was used to map monthly and annual precipitation in Iceland. The maps have a spatial resolution of 2 km or 4 km. The predictors are defined at each grid point with the DEM (figures 14 and 15).

Figure 16 presents the annual precipitation map for 2001 (2 km resolution) and figure 17 the sum of the 12 monthly precipitation maps from January to December 2001. The two maps display a rather similar pattern, with some discrepancies over the glaciers where the second estimation produces larger values over Vatnajökull but lower values over Hofsjökull. The annual precipitation maps from 1980 to 2000 (4 km resolution) are given in appendix 7 and the monthly precipitation maps for 2001 (2 km resolution) in appendix 8.

The method is able to produce more detailed information than any classical interpolation procedure, especially in the regions poorly or not monitored such as the highlands, the glaciers and the west and east fjords. It is also quite interesting to see that the average ratio between the maximum estimated precipitation and the maximum observed precipitation is 1.7 (min = 1, max = 2.45) for the annual precipitation (period 1980-2000) and 1.95 (min = 0.9, max = 4.6) for monthly precipitation (period 1990-2000). These results contribute to show how unrepresentative the network probably is, and a simple interpolation procedure will not be able to capture this detailed pattern and will give estimates ranging only between the minimum and maximum observed values.

Figure 14: grid 2 km resolution : predictors (p3) to (p6)

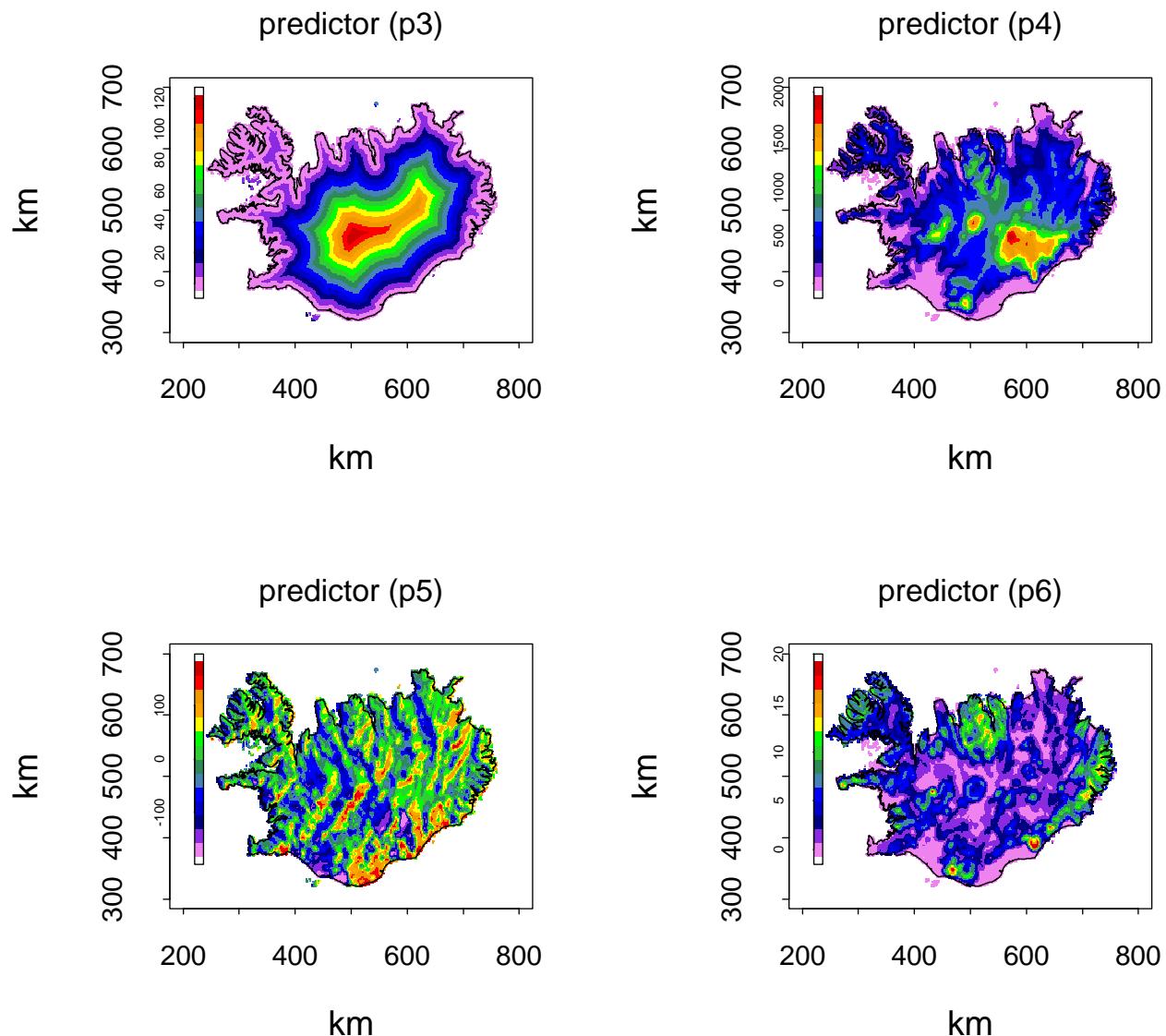


Figure 15: grid 2 km resolution : predictors (p7) to (p9)

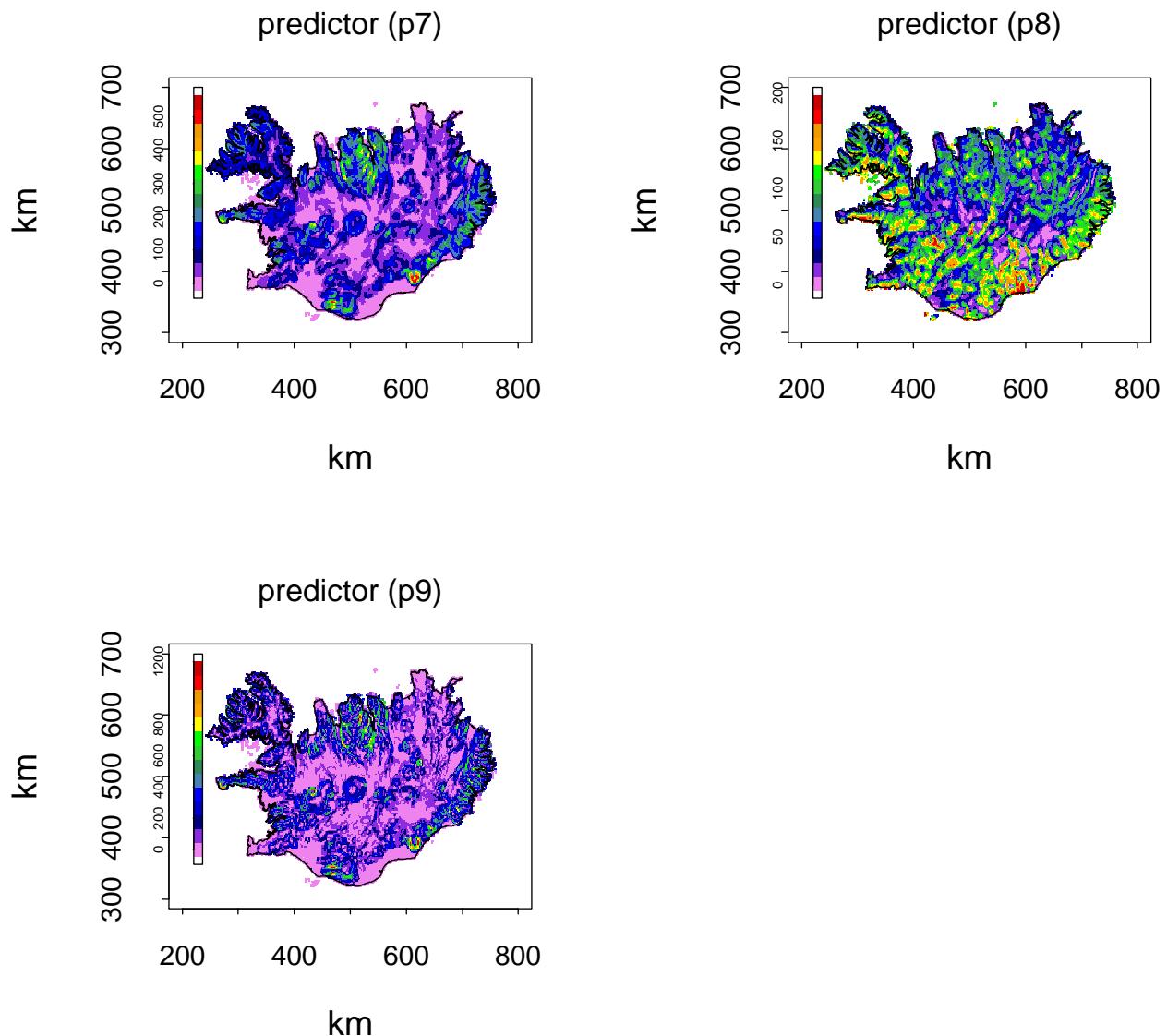


Figure 16: Annual precipitation in 2001

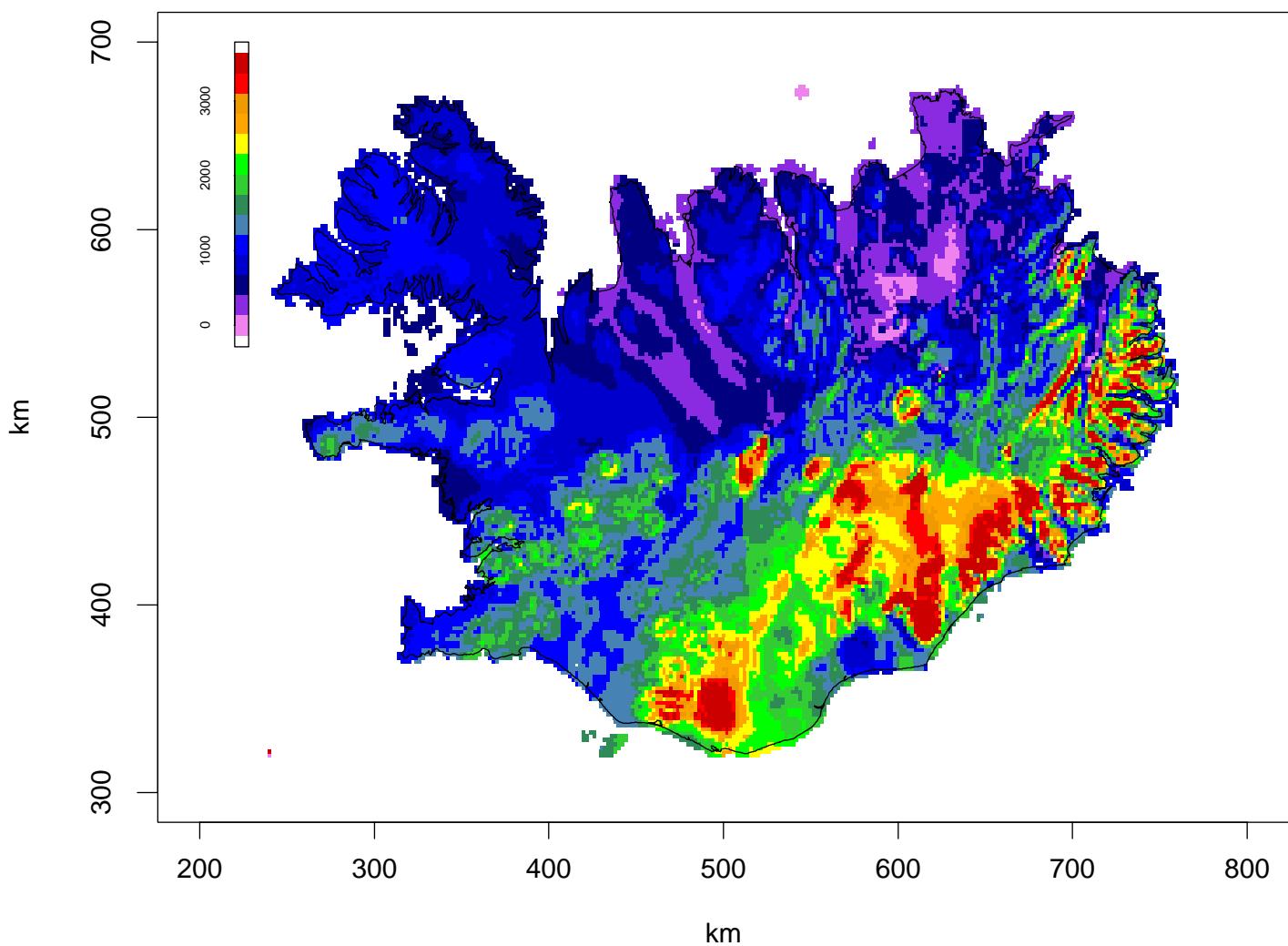
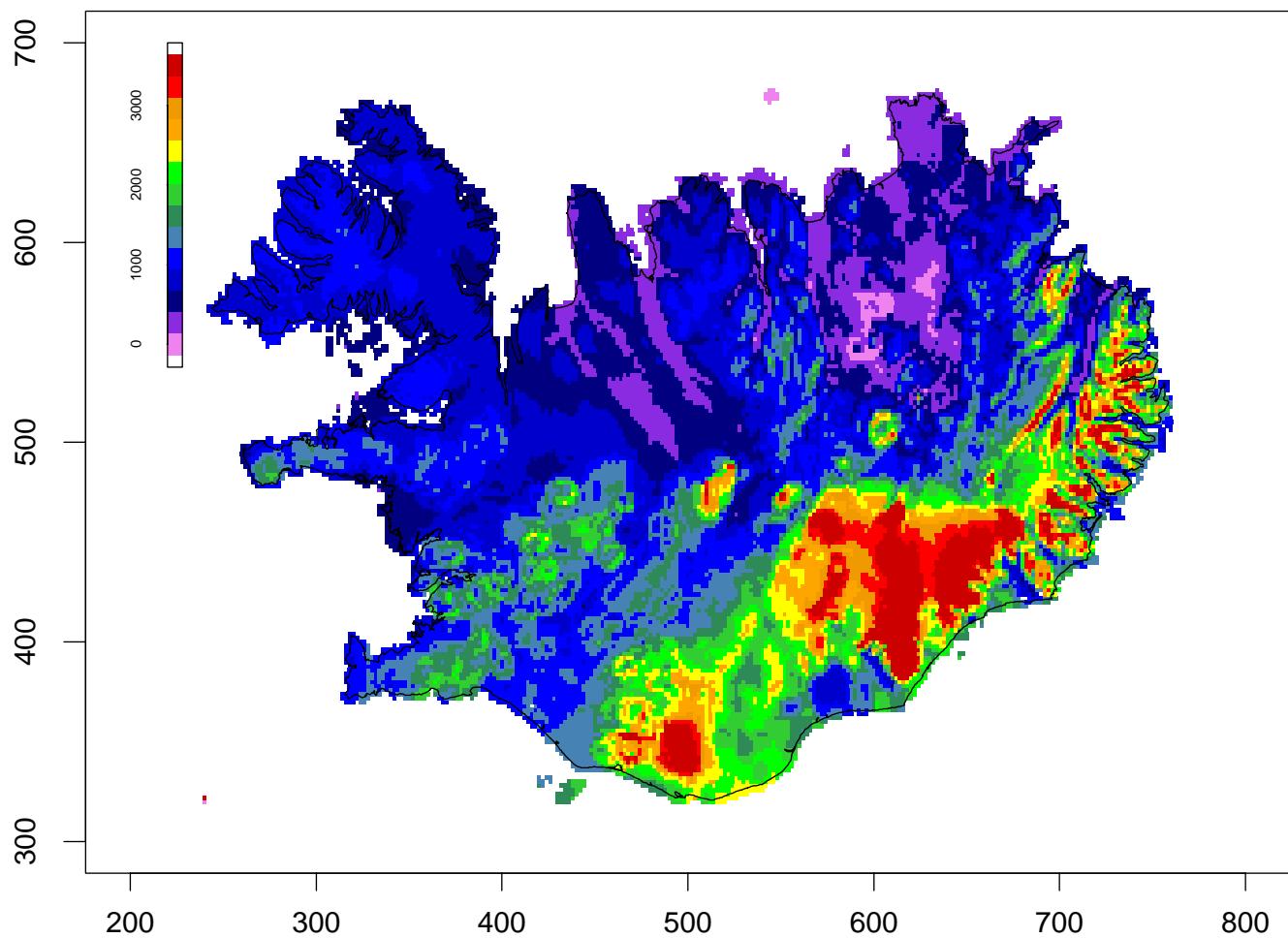


Figure 17: Annual precipitation in 2001 (sum of 12 months)



IV) Conclusion

In this work, a model has been developed for studying the influence of geographical and topographical factors on precipitation. According to the model characteristics, it is observed that a large part of the spatial variability of monthly and annual precipitation in Iceland is explained by the geographical and topographical environment in the vicinity of the considered locations. The results suggest that the predictors do not contribute to precipitation in the same manner in space and time. The systematic errors observed at some locations show that a more detailed division of Iceland will most likely improve the quality of the estimation. In the future, the model will hopefully gain in accuracy when information over the glaciers such as ice core data will be available. The main range of application of this model is to produce high resolution precipitation maps. For the near future, it is intended to be used in order to assess the quality of the precipitation amounts generated by the MM5 numerical prediction model. Further work in both the selection of the model parameters and the estimation procedure will be considered later, as well as the extension of the method to shorter time steps and other variables such as statistical parameters and other weather parameters such as the temperature.

Acknowledgements

This work was supported by the Icelandic Research Council (RANNÍS) and the Nordic Climate Water and Energy (CWE) project. It is part of the validation protocol developed for assessing the quality of the precipitation amounts produced by the MM5 model in Iceland. I wish to thank Thomas Jóhannesson and Haraldur Ólafson for involving me in this research project, and Trausti Jónsson for the interesting discussions concerning the work presented in the present report.

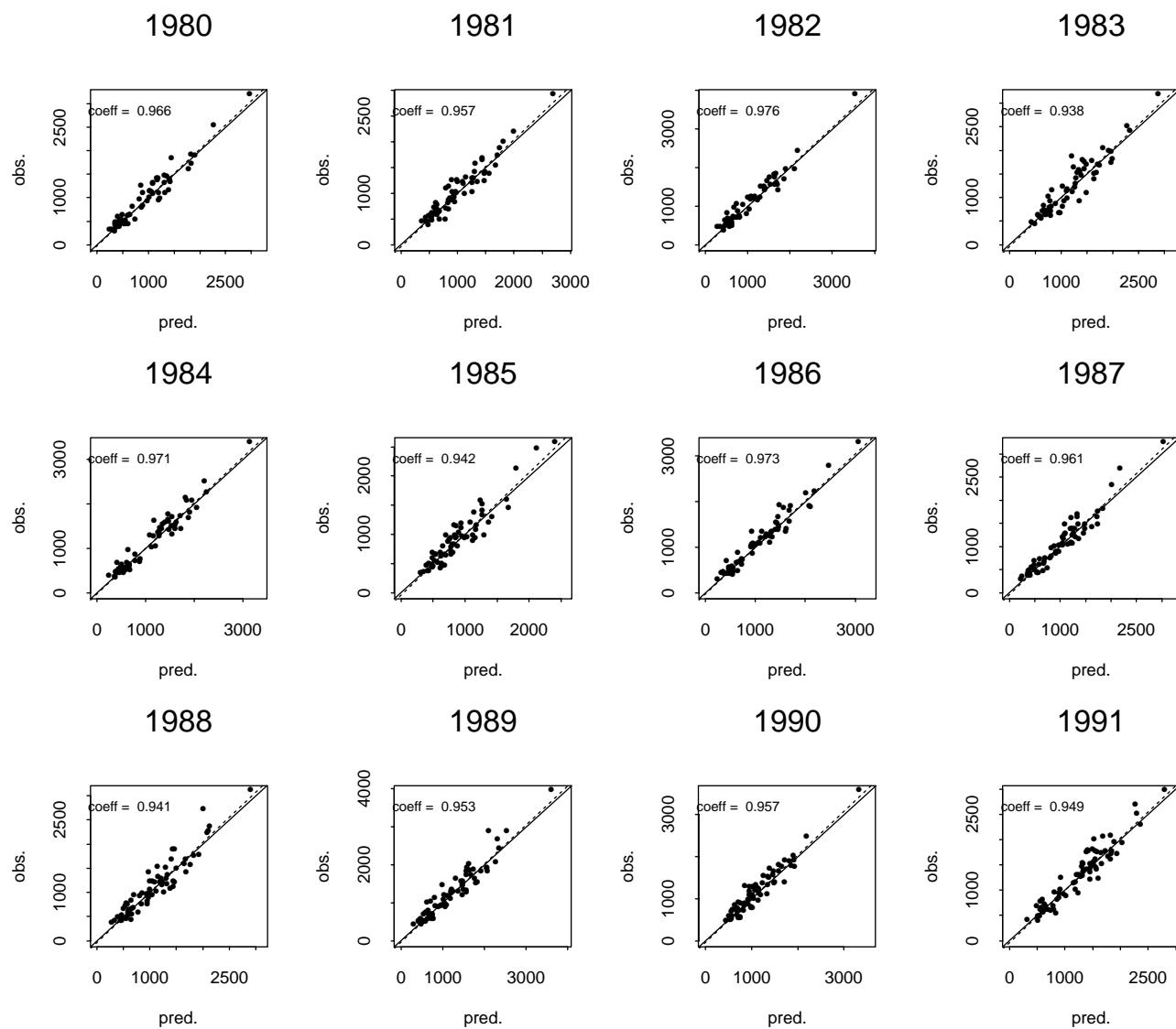
References

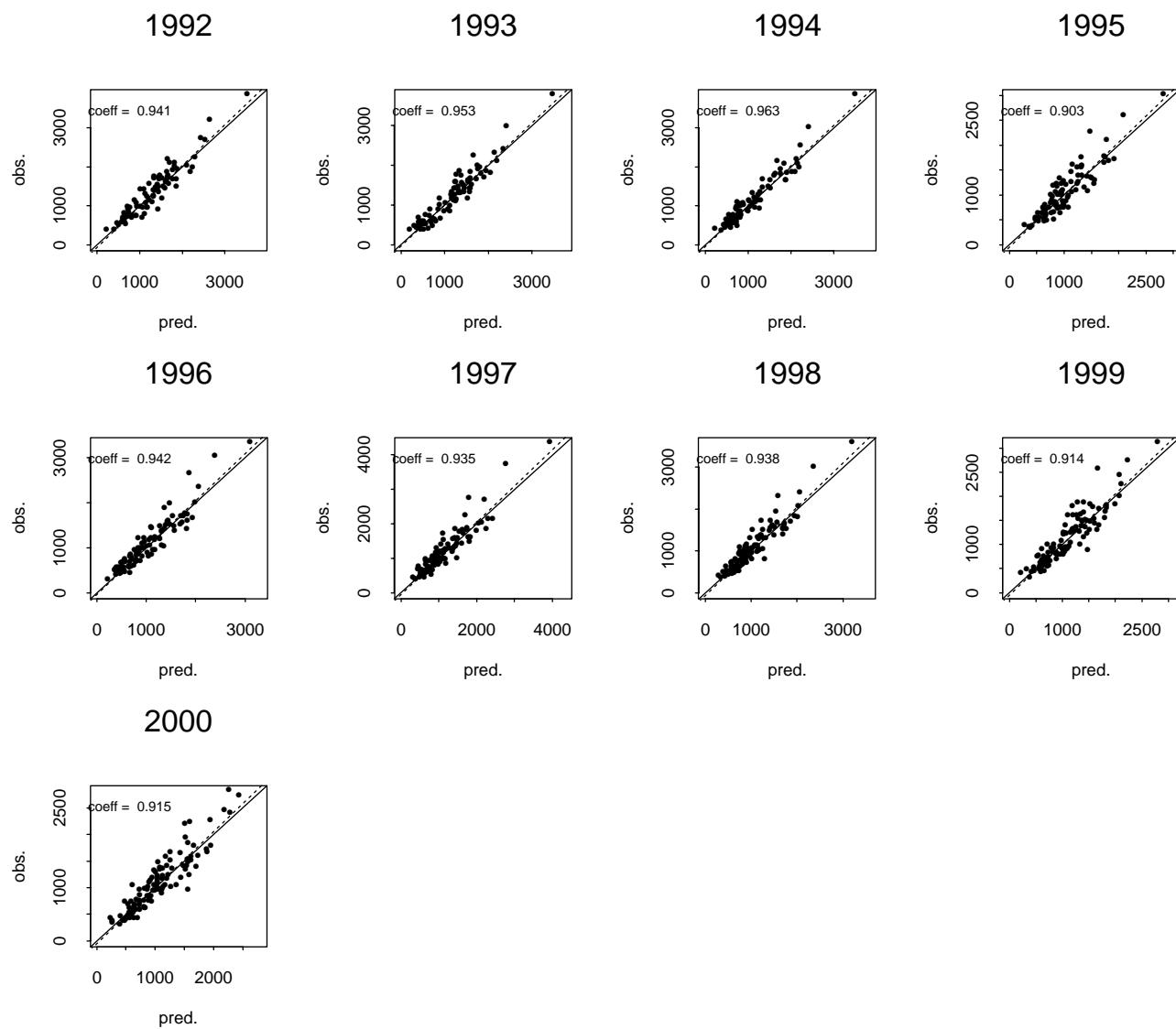
- [1] **Creutin J.D. and C. Obled, 1982:** Objective analyses and mapping techniques for rainfall fields: an objective comparison. WRR, Vol. 18 no 2, 413-431.
- [2] **Delhomme J.P., 1978:** Kriging in the hydrosciences. Advances in water resources, vol 1, no 5, 251-266.
- [3] **Kitanidis P.K., 1997:** Introduction to geostatistics, application in hydrogeology. Cambridge University Press, 249 pp.
- [4] **Barancourt C., J.D. Creutin and J. Rivoirard, 1990:** A method for delineating and estimating rainfall fields. WRR, Vol. 28 no. 4, 1133-1144.
- [5] **Benichou, P. and O. Le Breton, 1987:** Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques. La Météorologie, 7e série - no 19.
- [6] **Whitemore J.S., 1972:** The variation of mean annual rainfall with altitude and locality in south Africa, as determined by multiple curvilinear regression analysis. Geilo symposium WMO/OMM no 326.
- [7] **Storr, D. and H.L. Ferguson, 1972:** The distribution of precipitation in some mountainous canadian watersheds. Geilo symposium WMO/OMM no 326.
- [8] **Rögnvaldsson, Ó and H. Ólafsson, 2001:** Validation of high-resolution simulations with the MM5 system. Proc. NCAR MM5 workshop, Boulder, USA, June 2002.
- [9] **Ólafsson, H., Á. J. Elíasson og Egill Þorsteins, 2002:** Orographic influence on wet snow icing conditions, Part I: Upstream of mountains. Proc. Intern. Workshop on Atmos. Icing on Structures, Prag, Czech Republic, June 2002.
- [10] **Lebel T., G. Bastin, C. Obled and J.D. Creutin, 1987:** On the accuracy of areal rainfall estimation: a case study. WRR, Vol. 23 no 11, 2123-2134.

Appendix 1

Annual precipitation

Scatter plots for the period 1980-2000

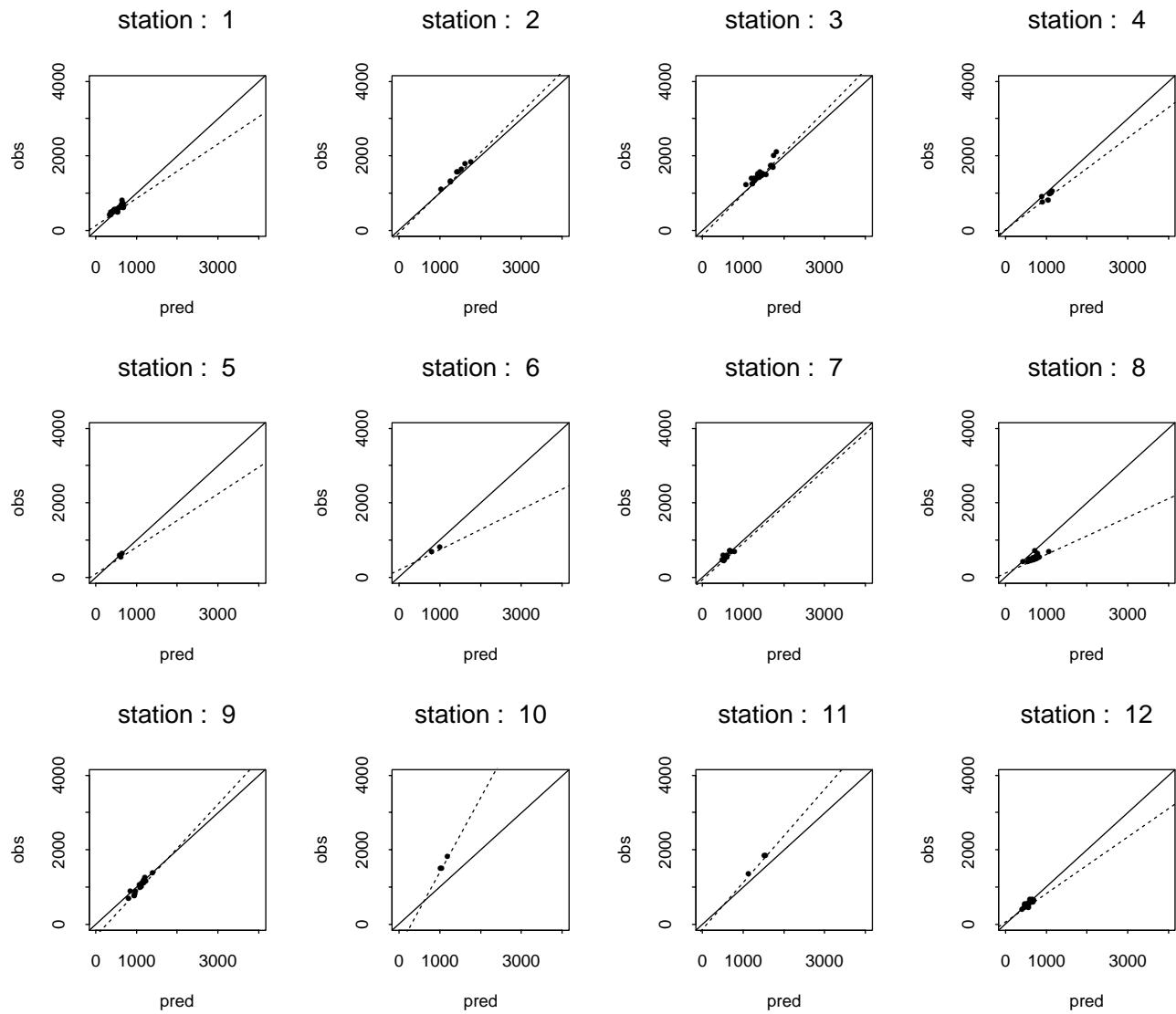


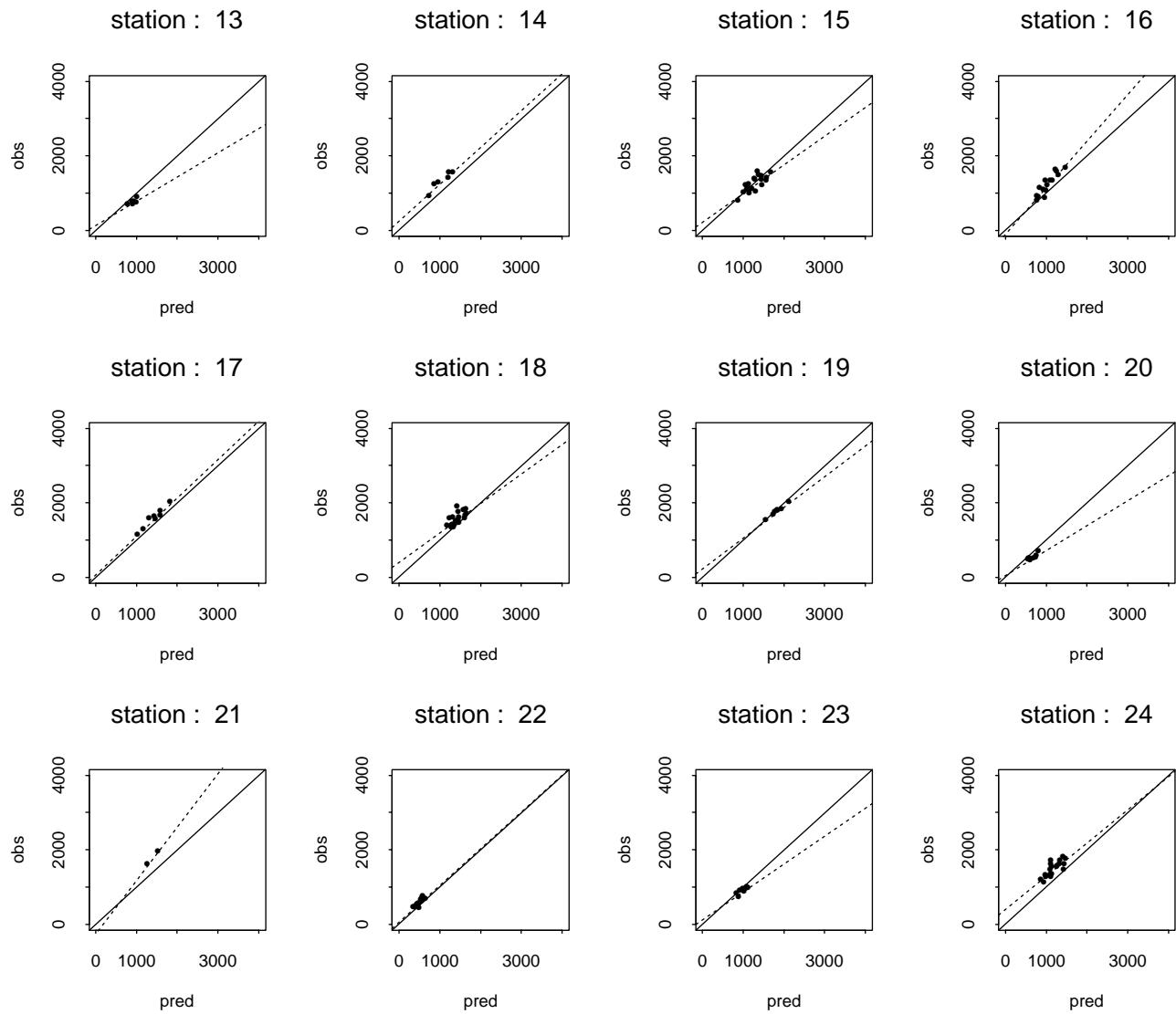


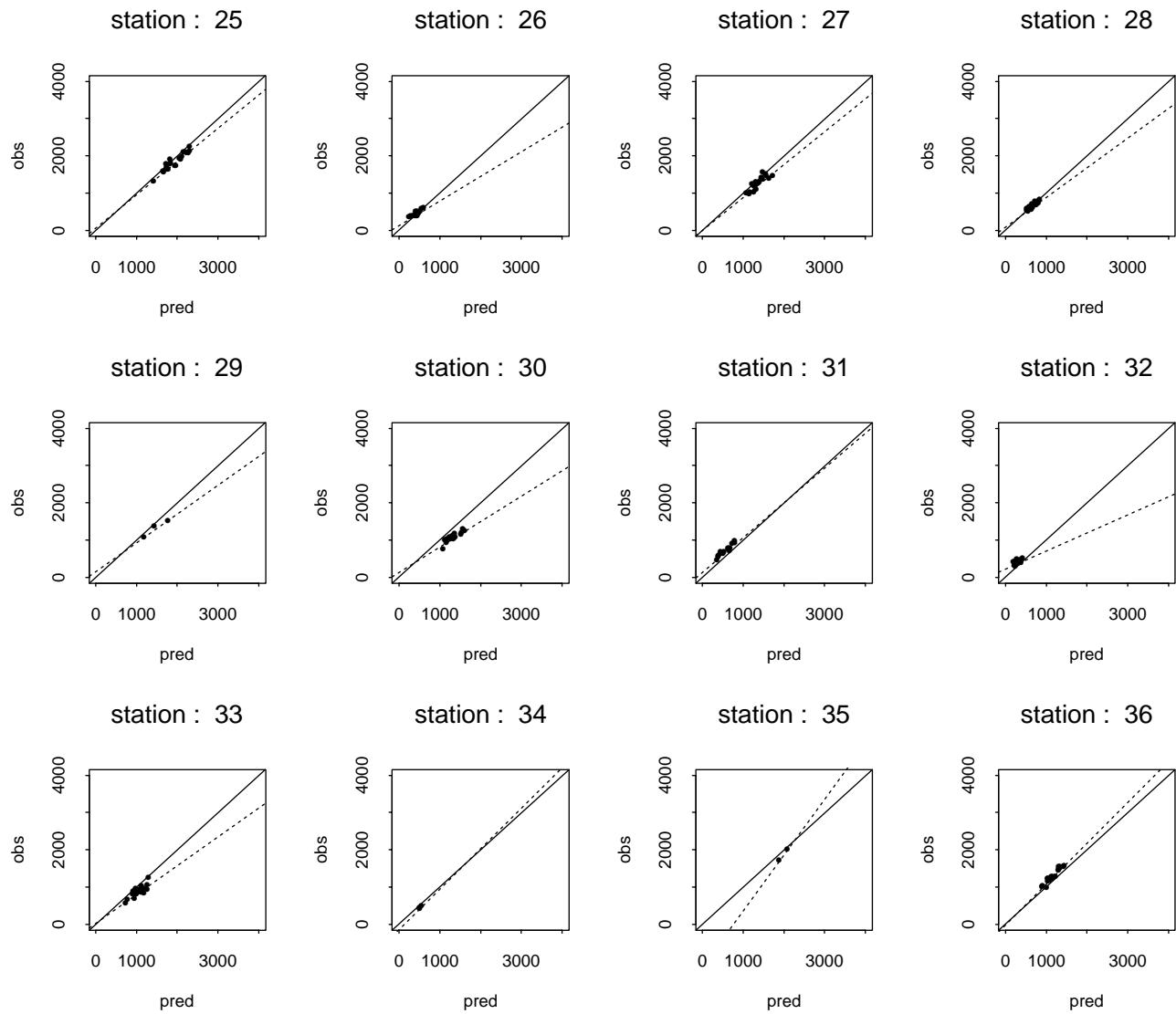
Appendix 2

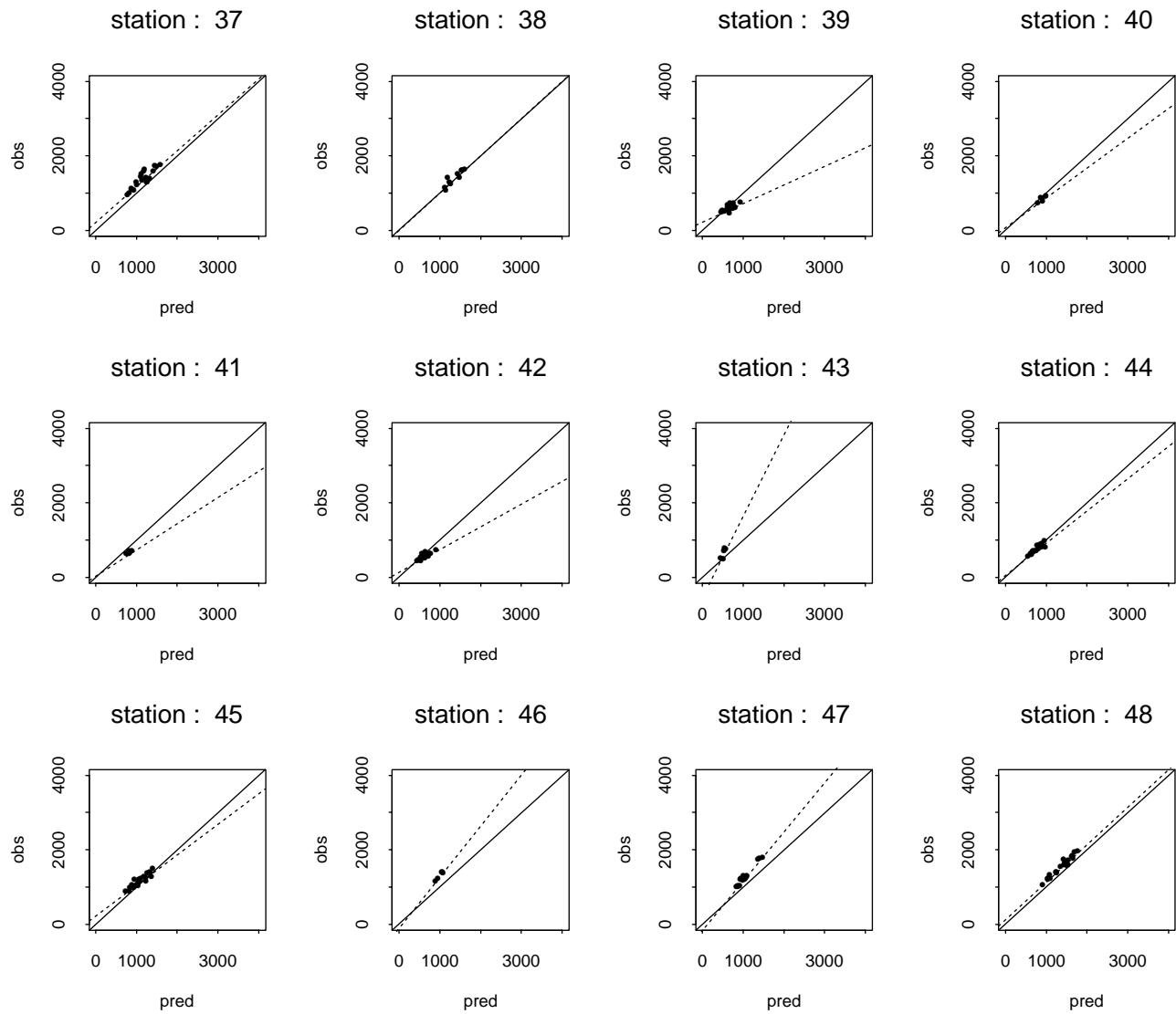
Annual precipitation

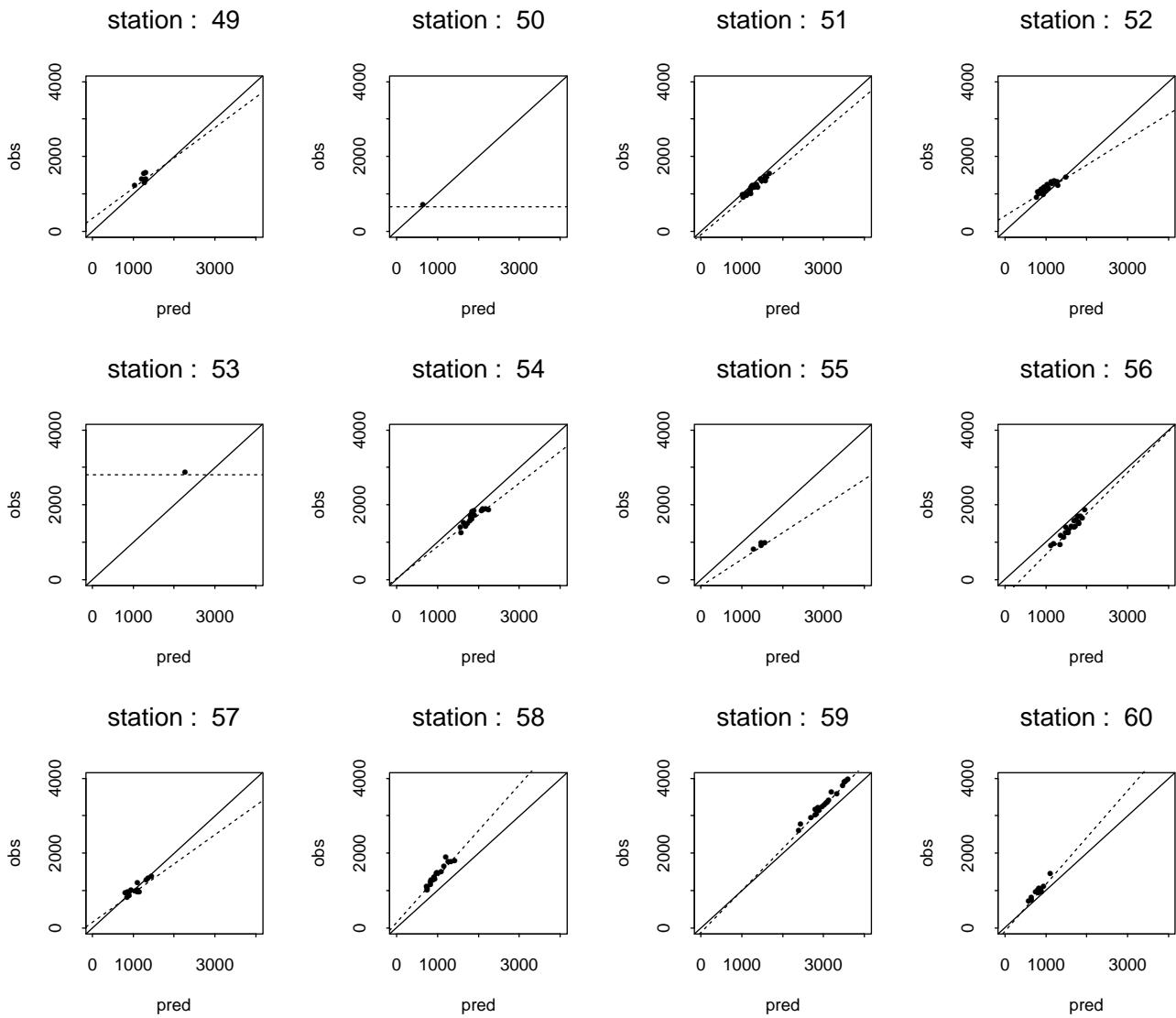
Scatter plots for each station (period 1980-2000)

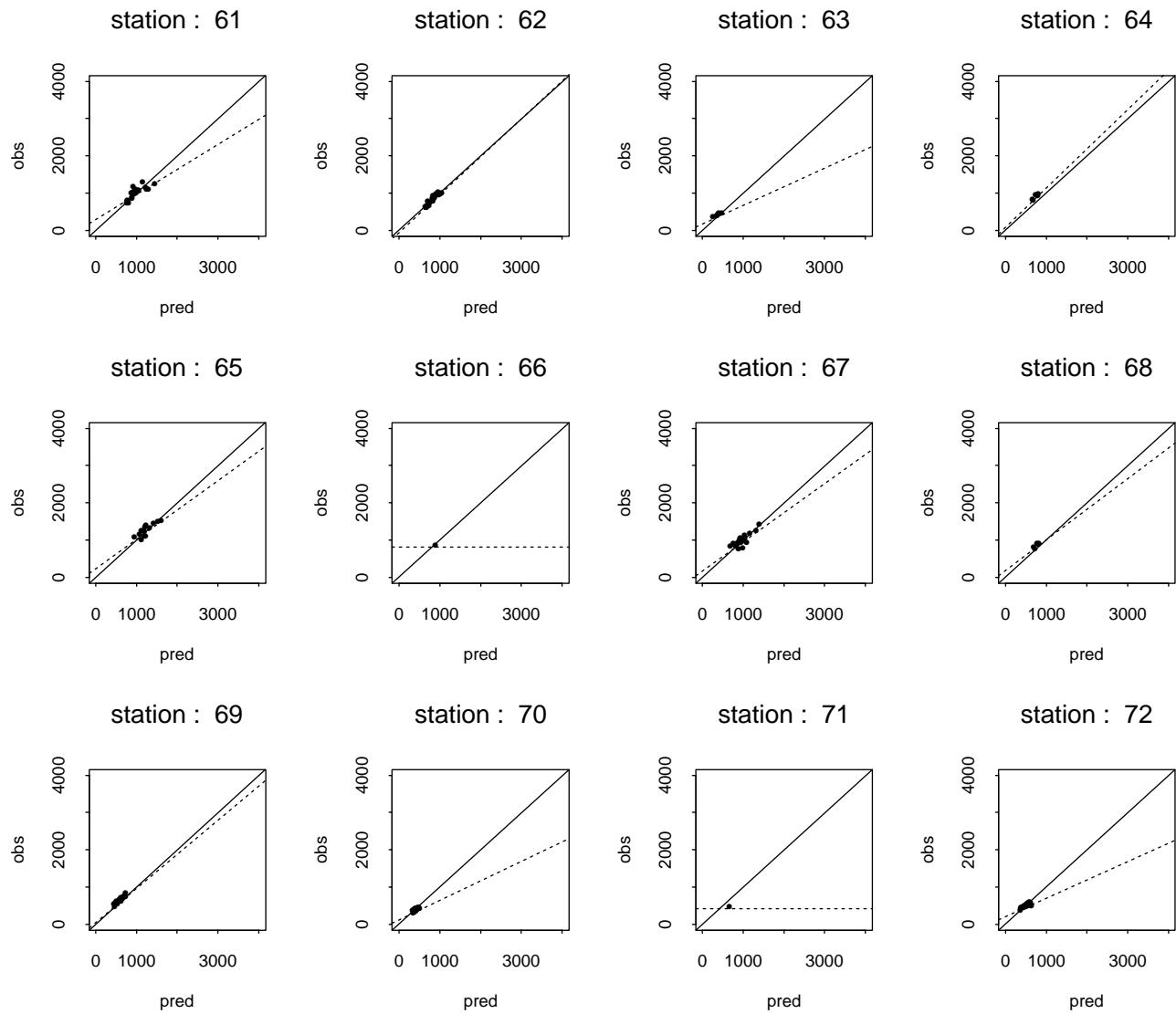


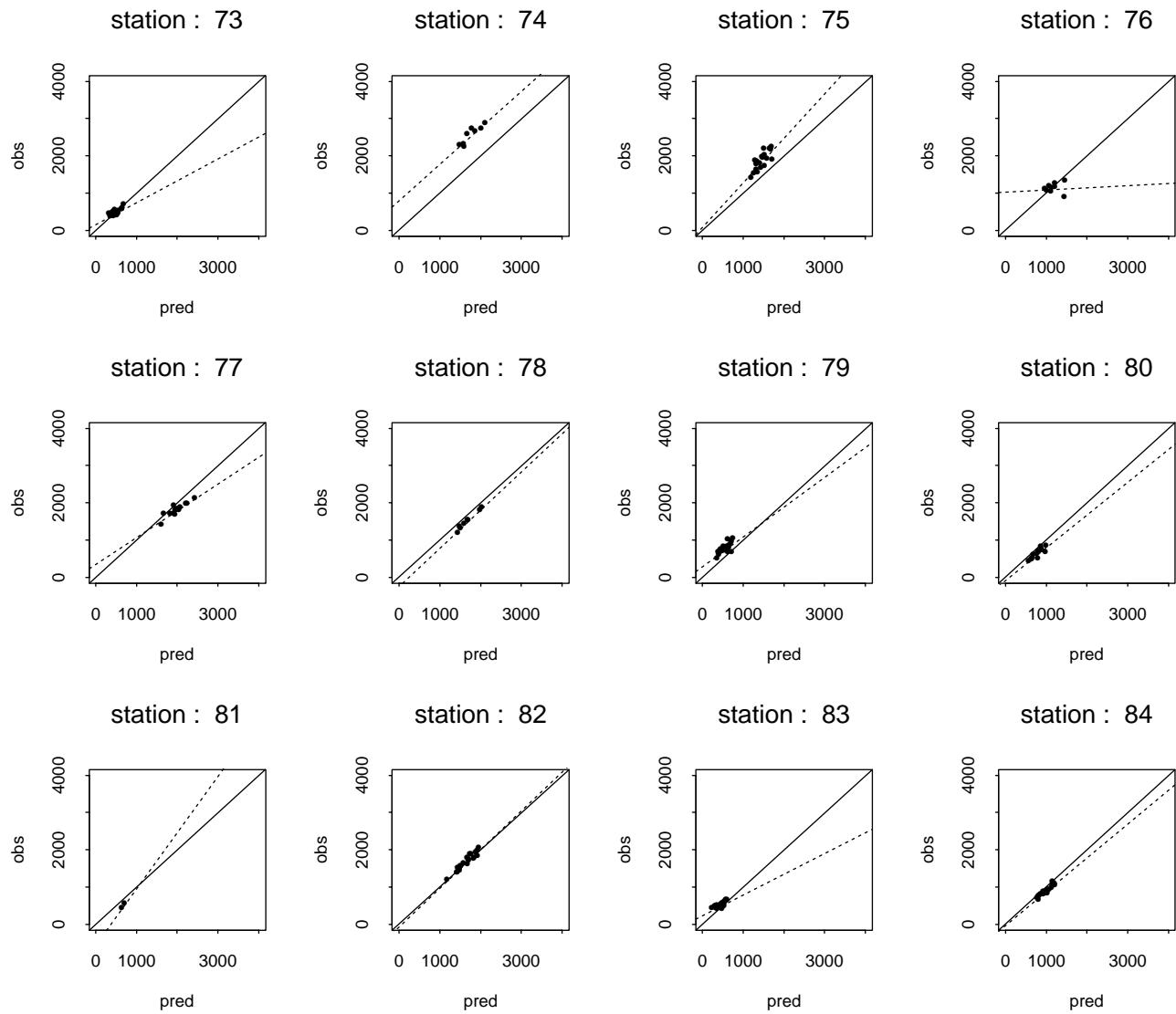


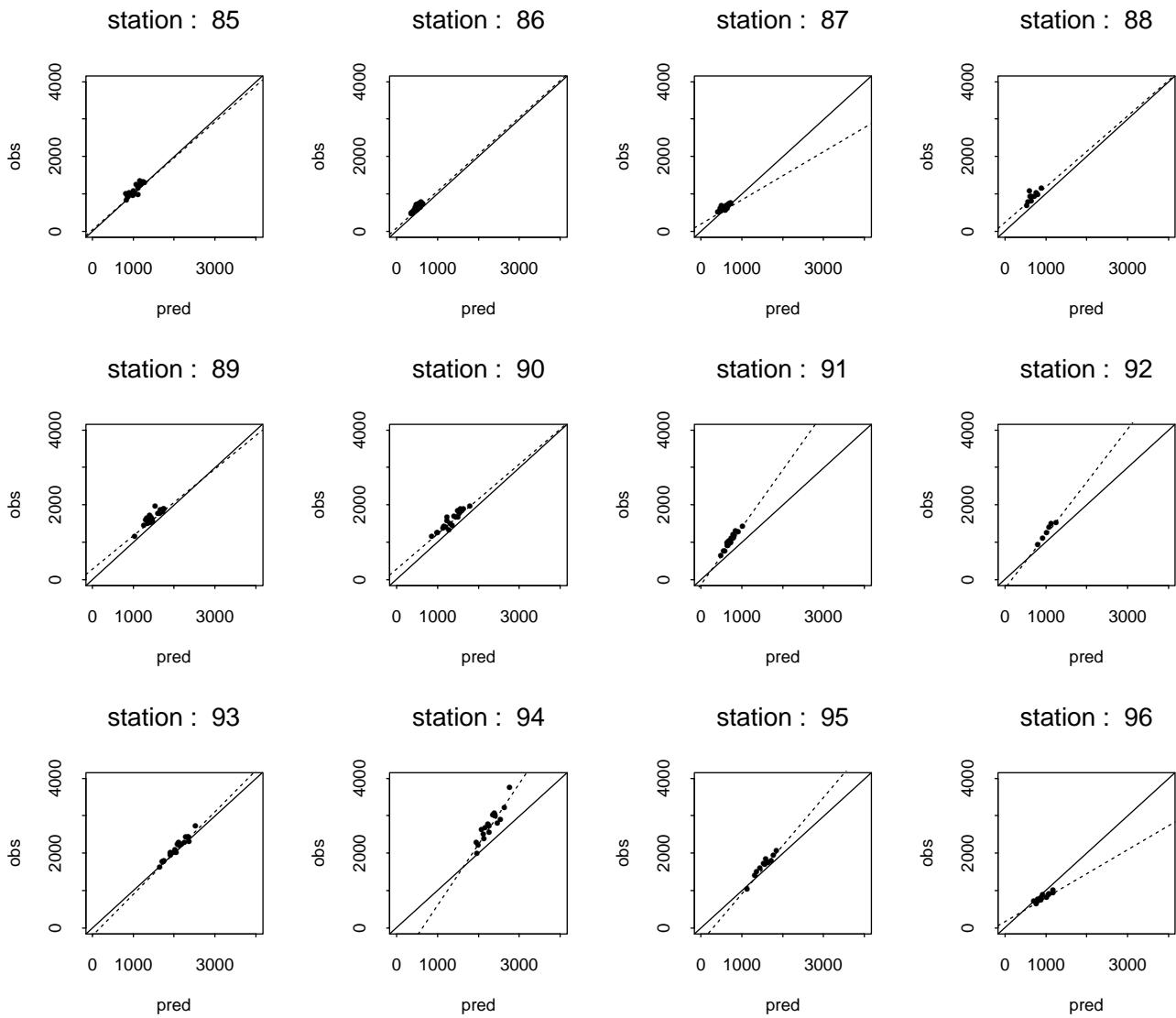


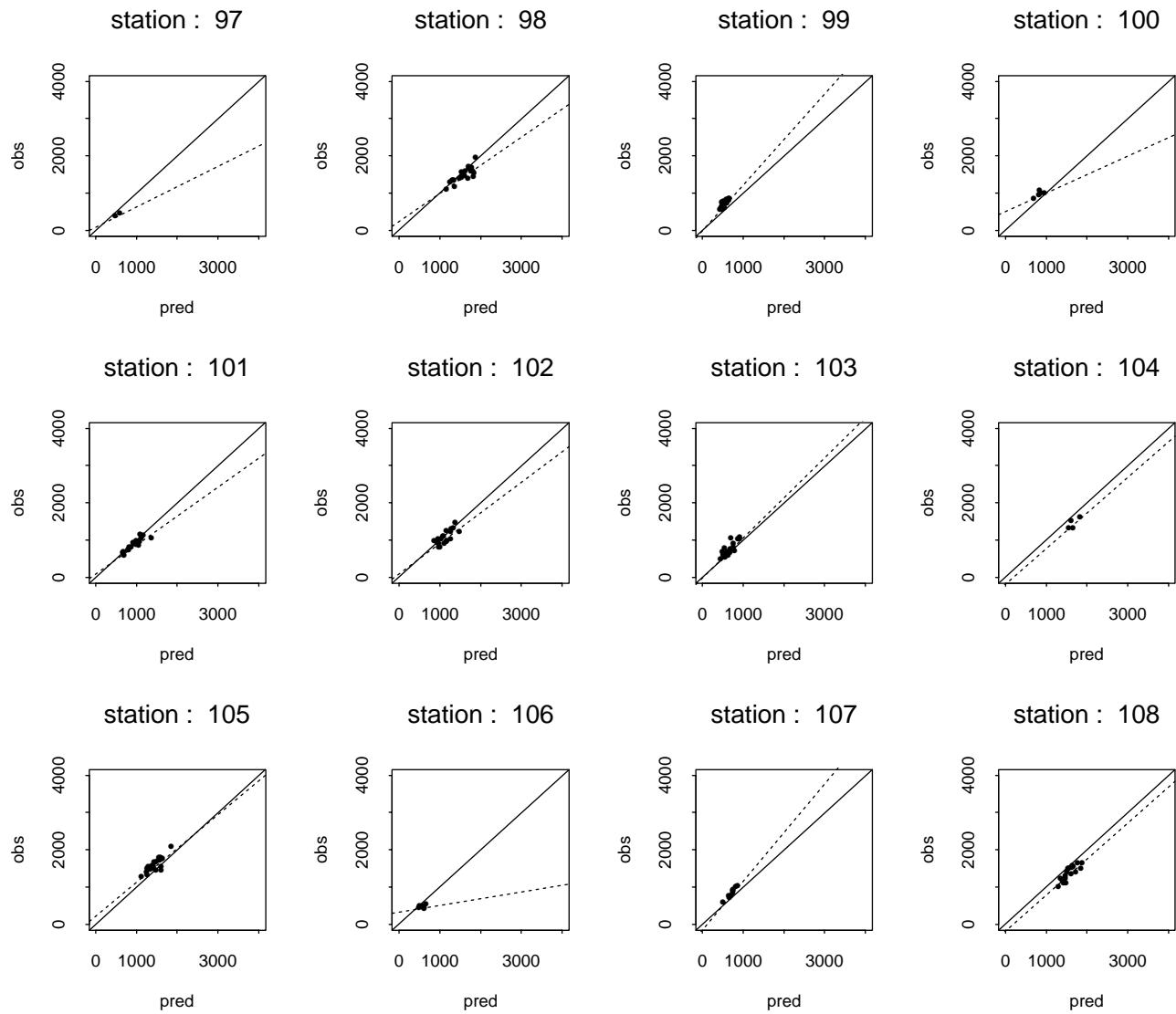


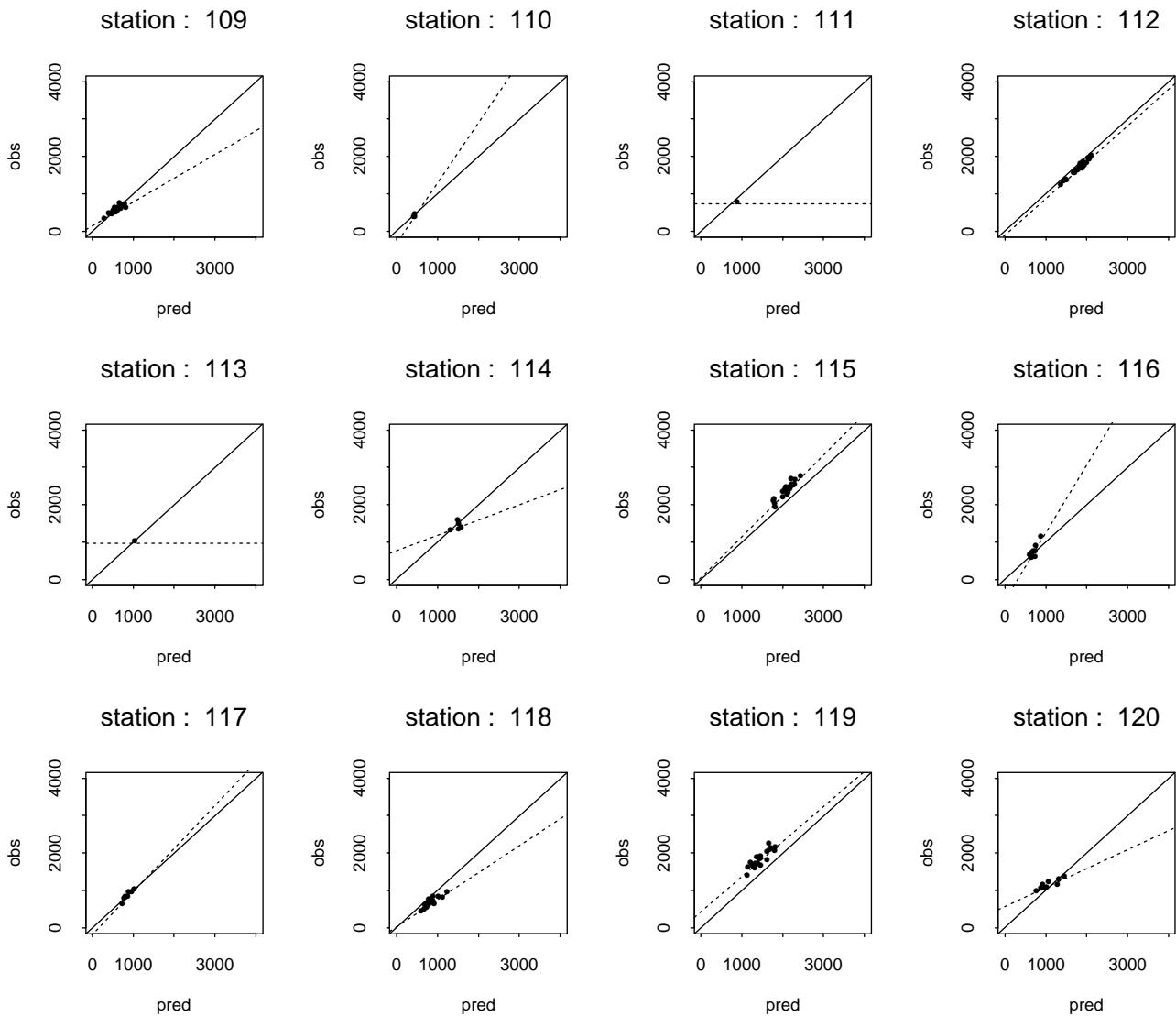


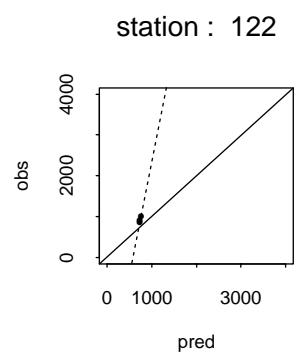
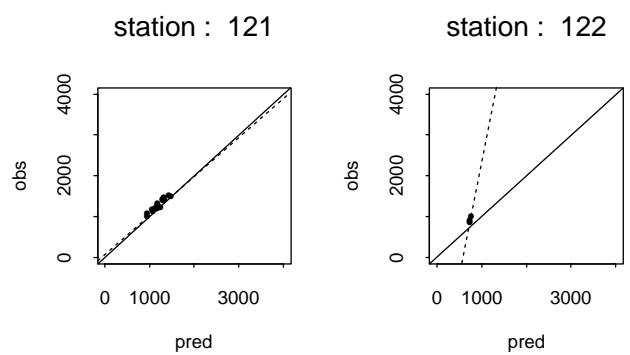










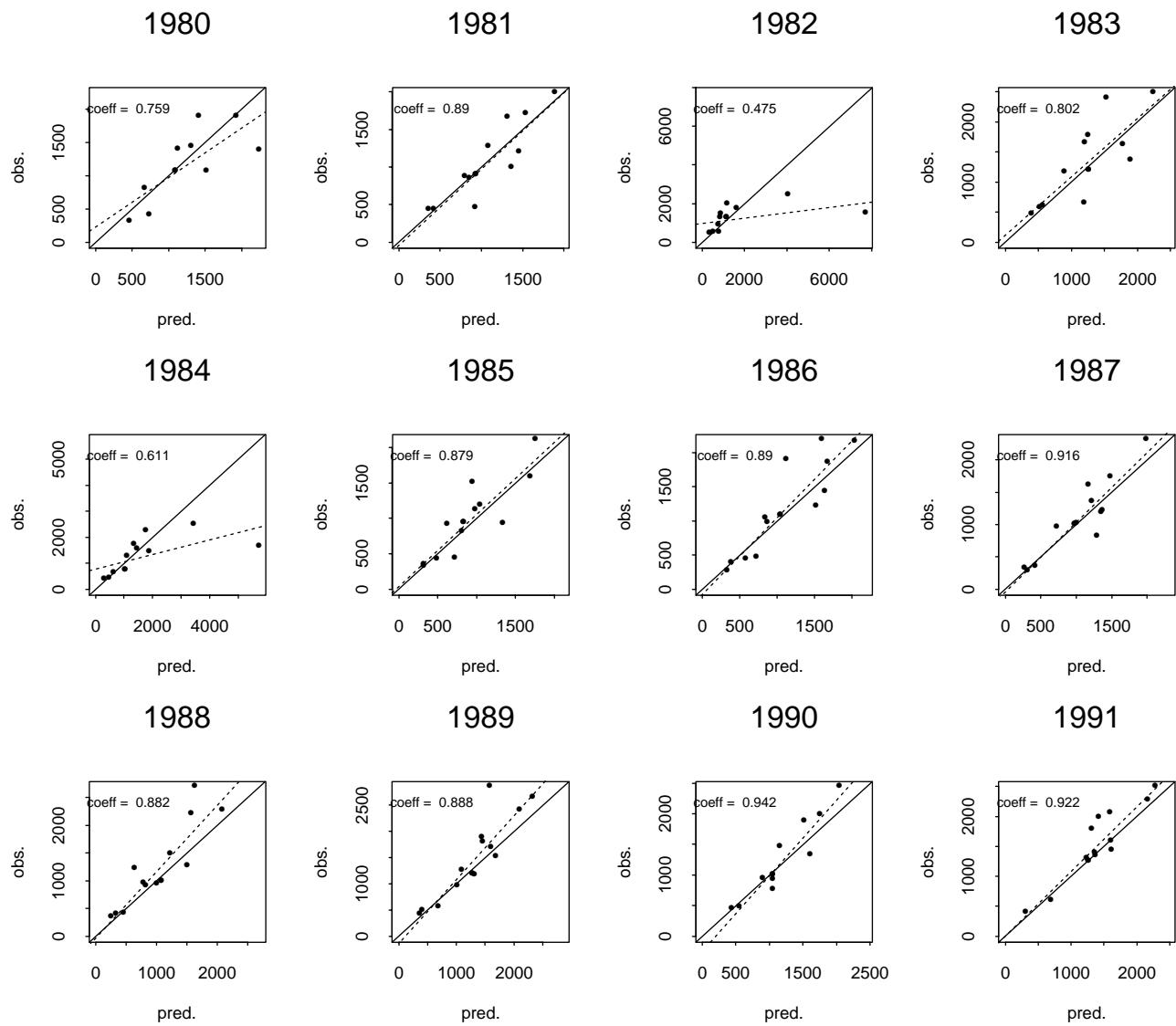


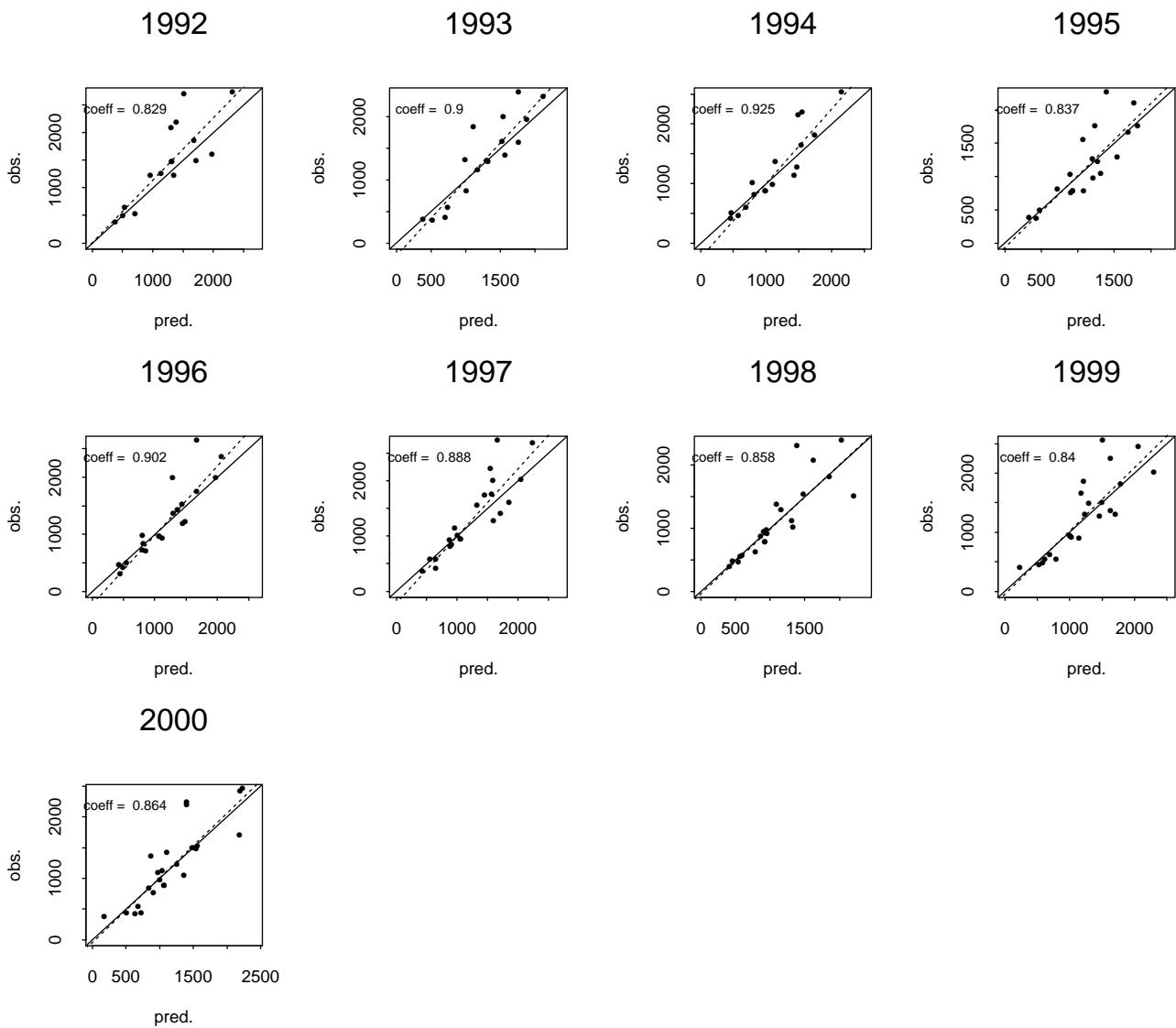
Appendix 3

Annual precipitation

Validation procedure

Scatter plots for the period 1980-2000

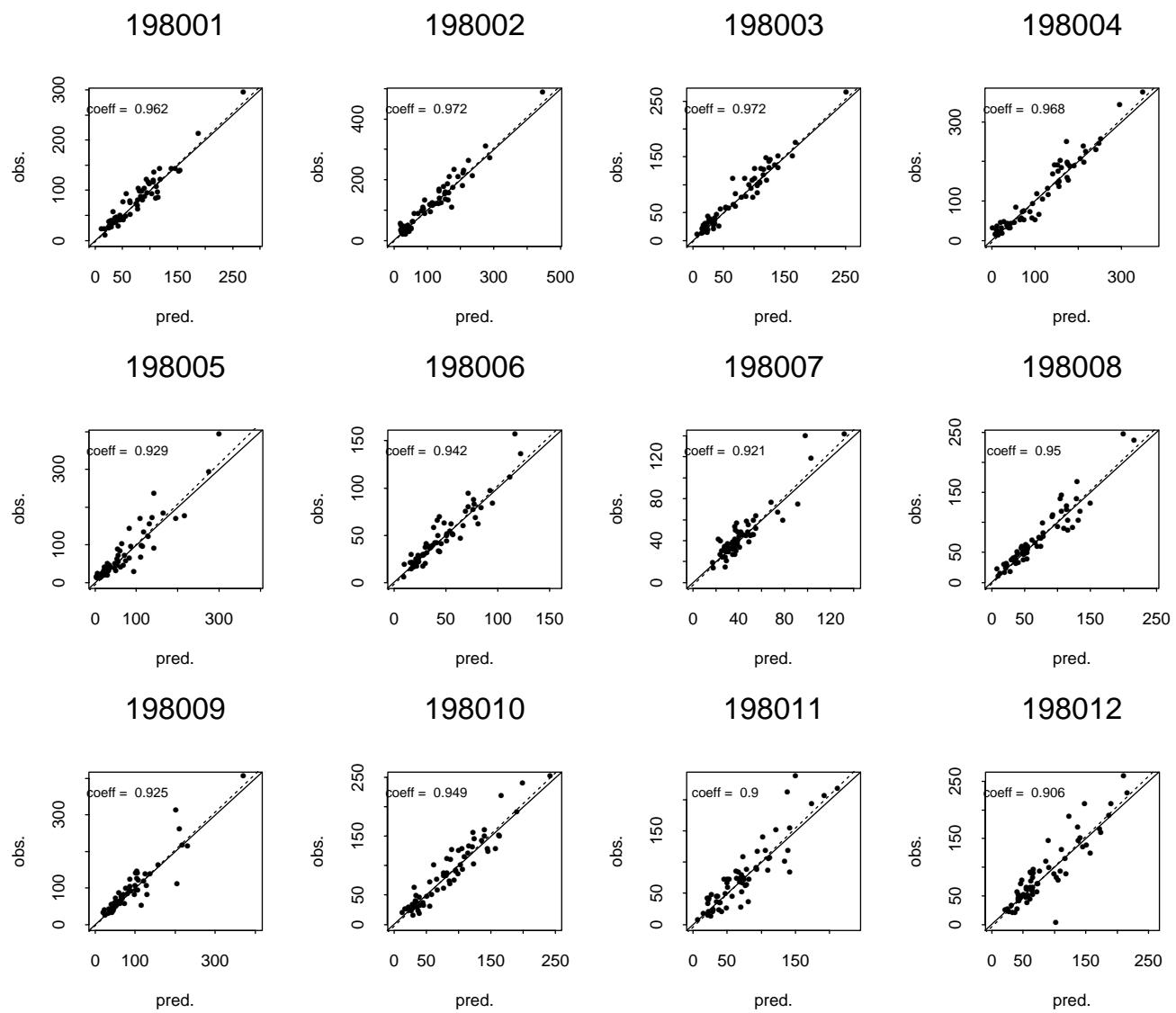


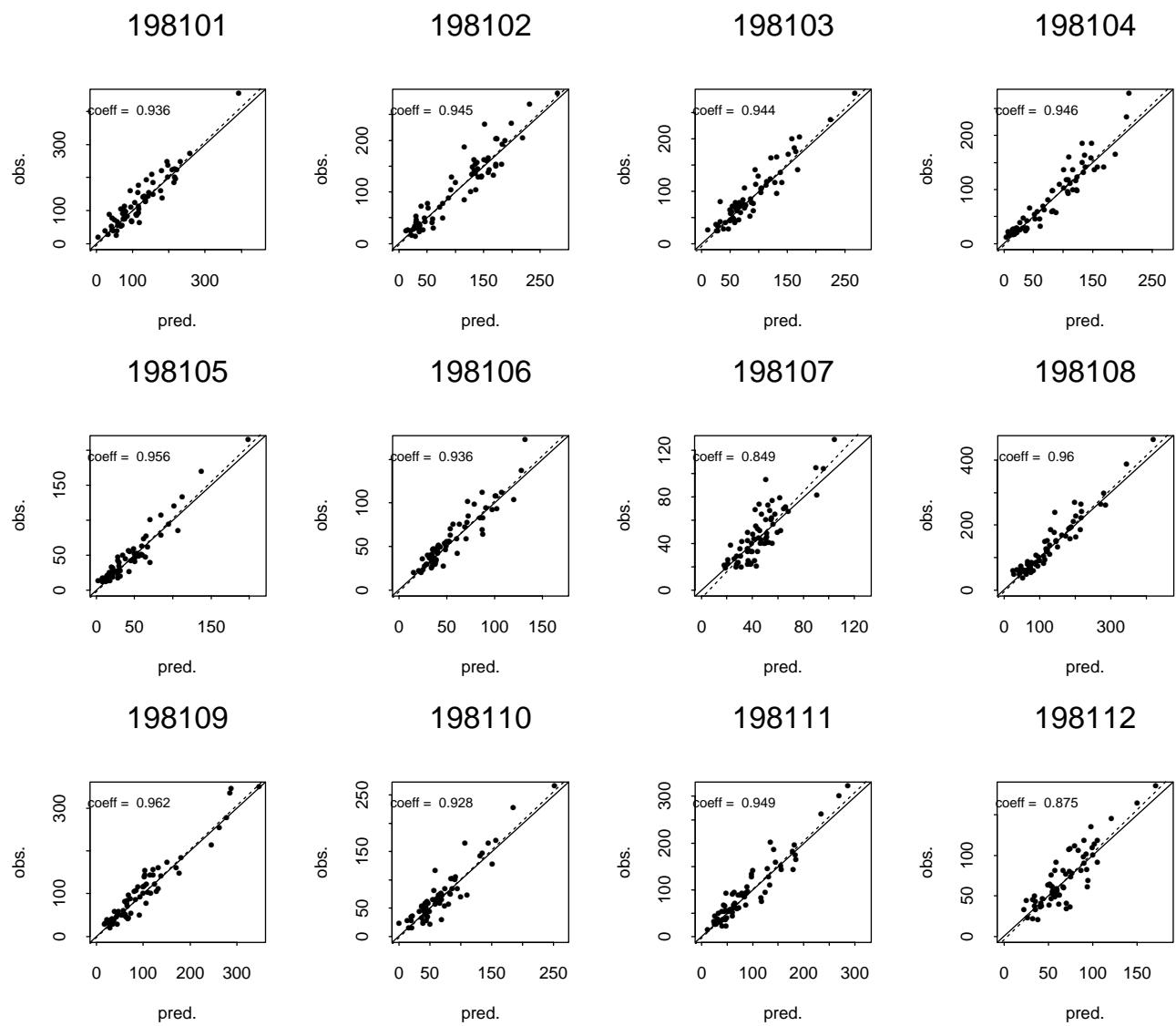


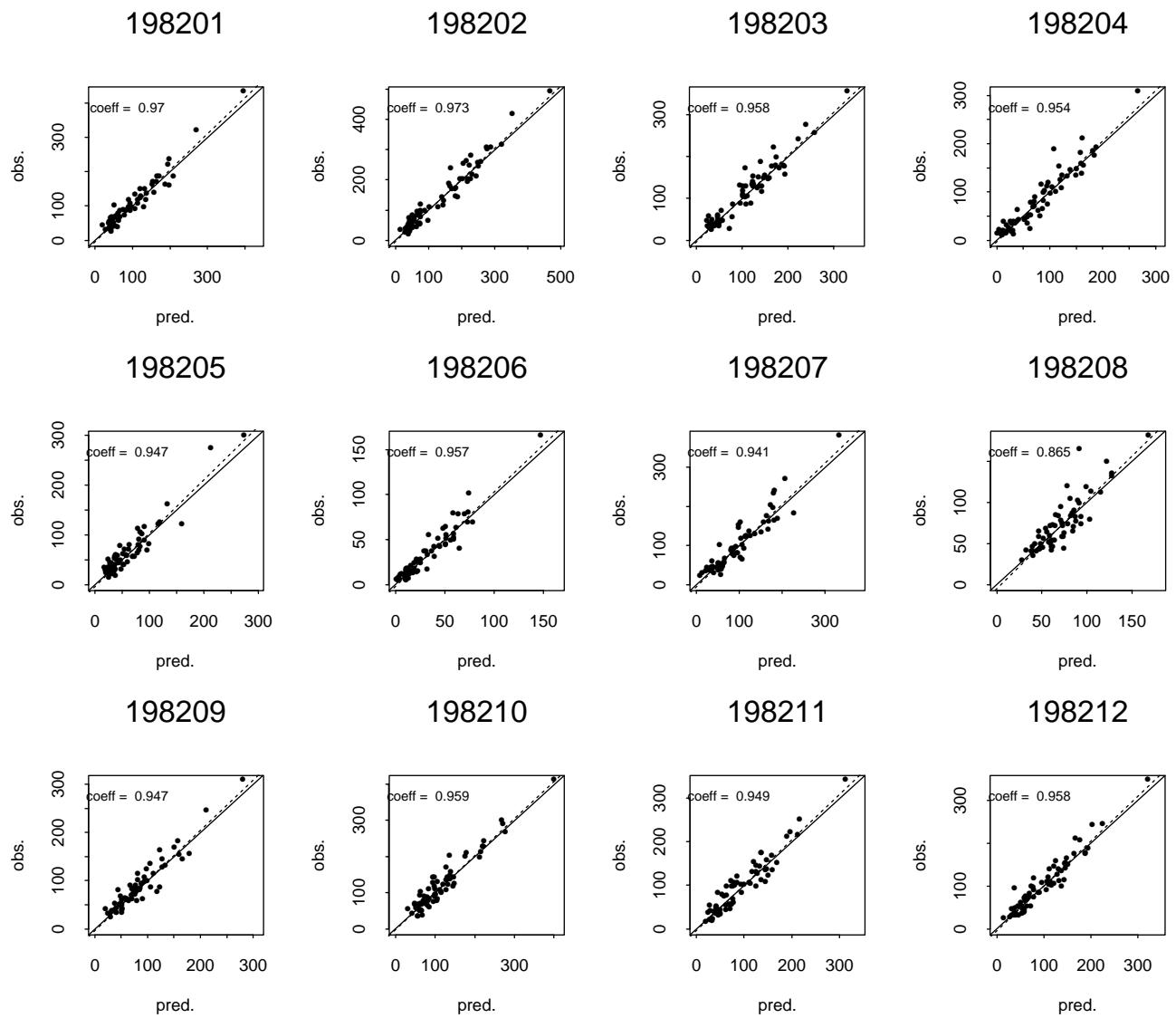
Appendix 4

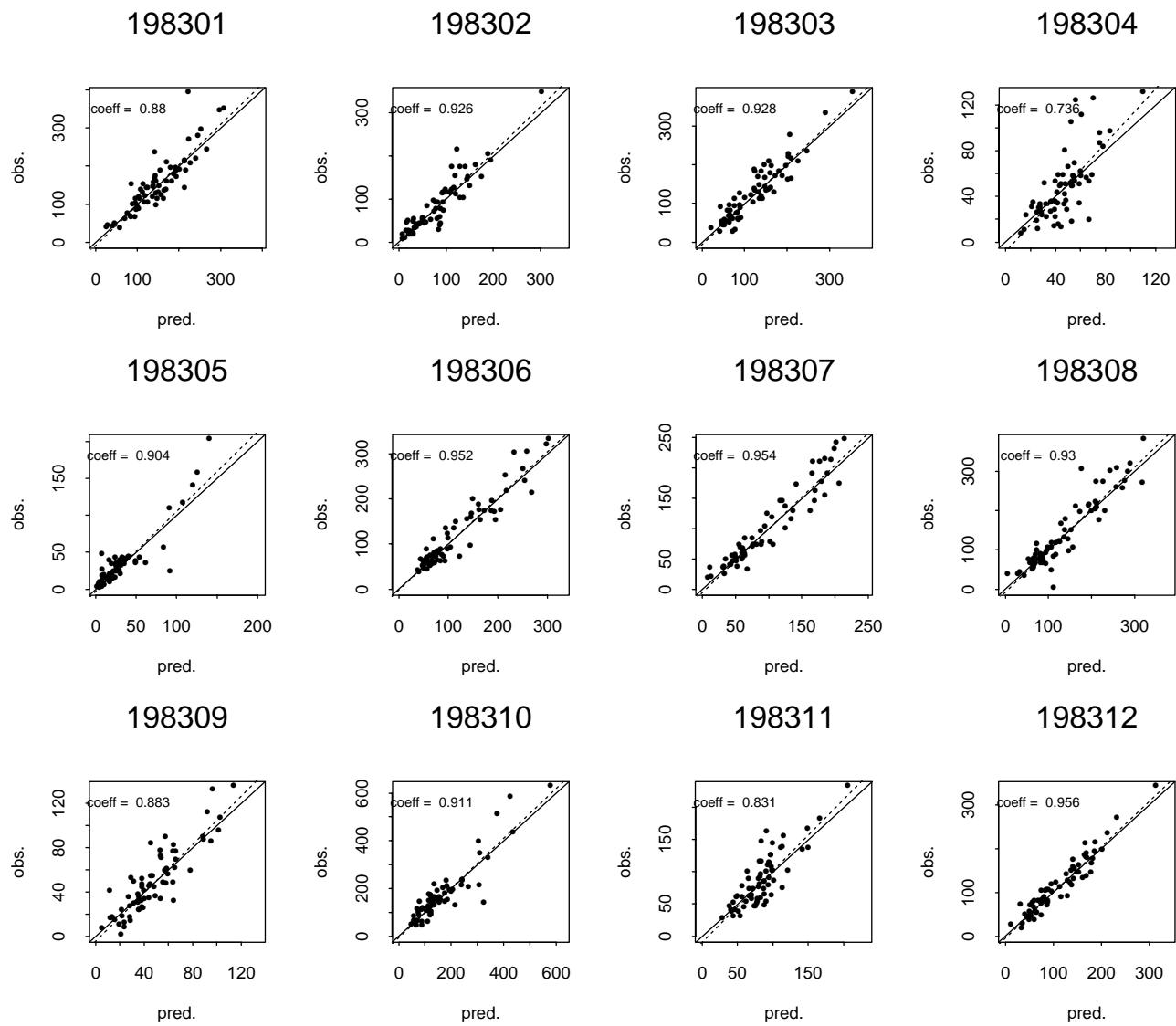
Monthly precipitation

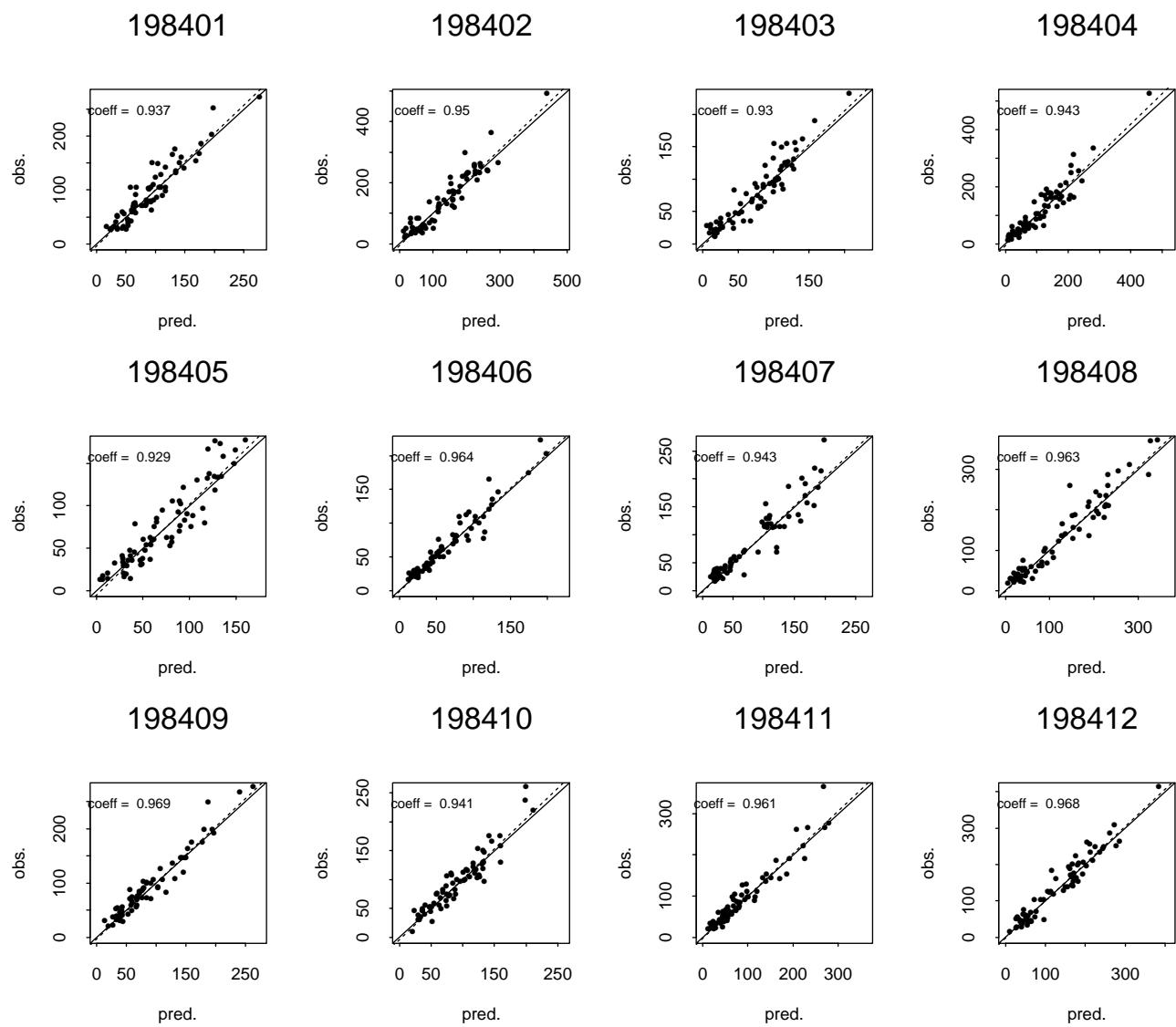
Scatter plots for the period 1980-2000

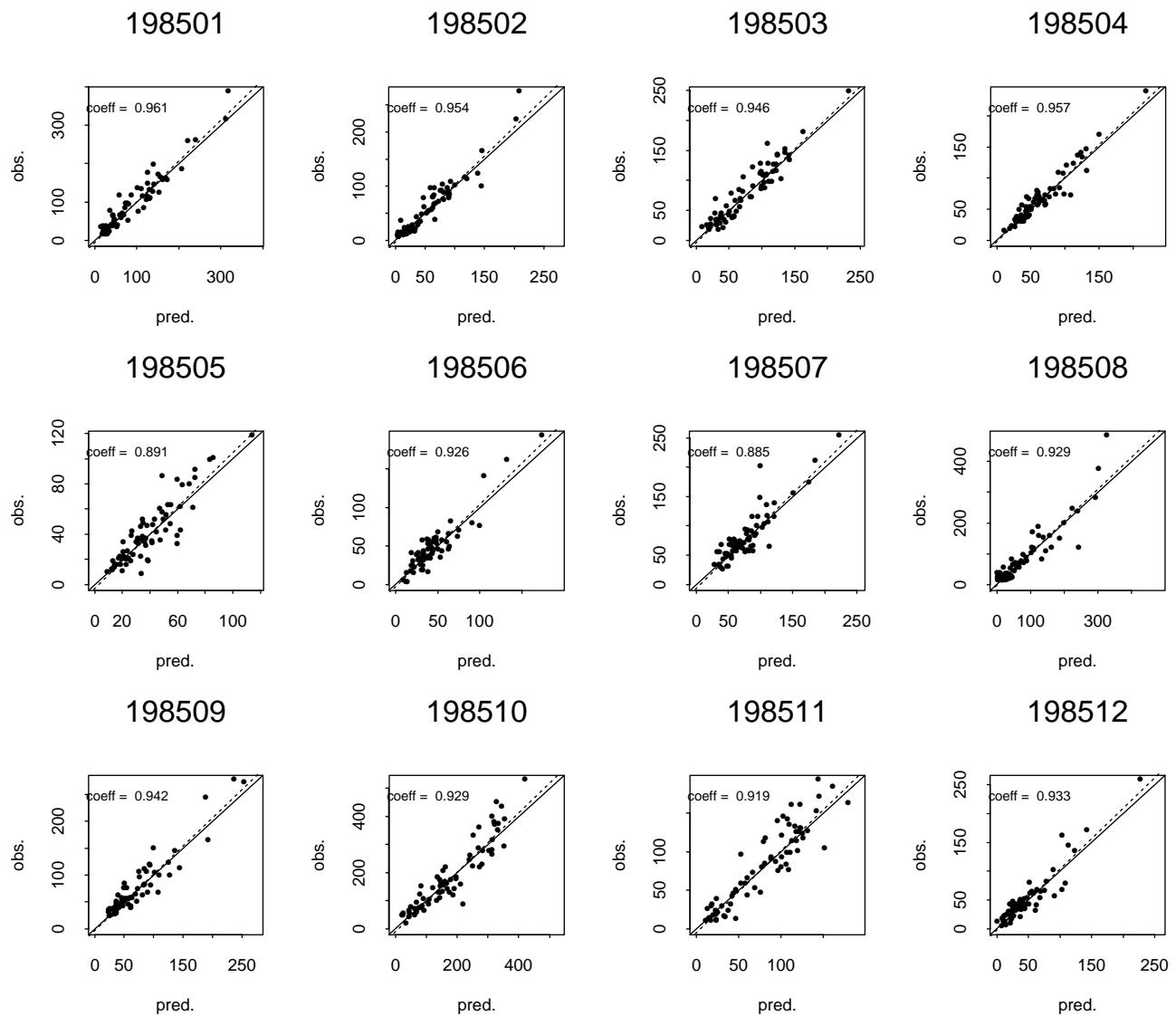


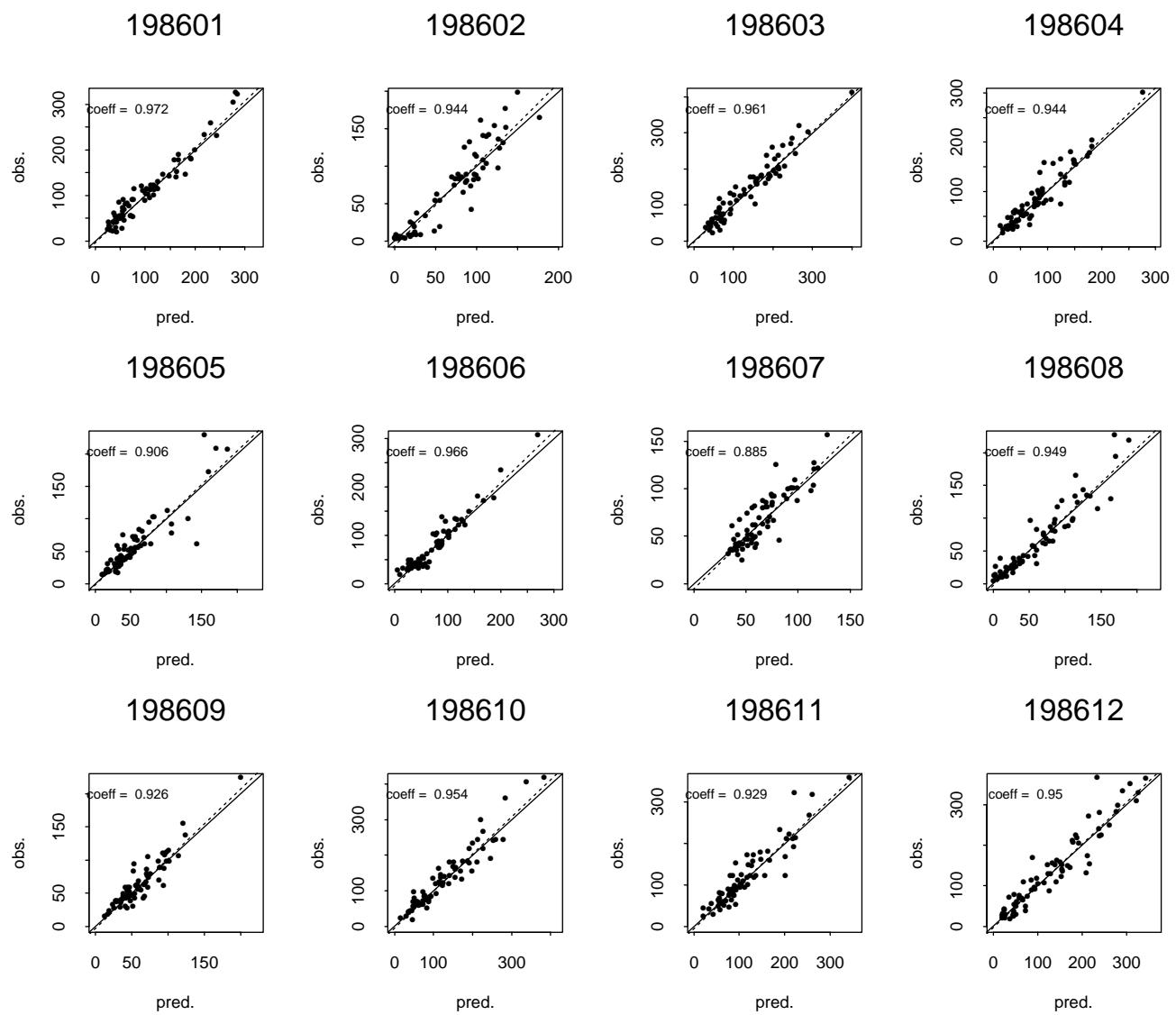


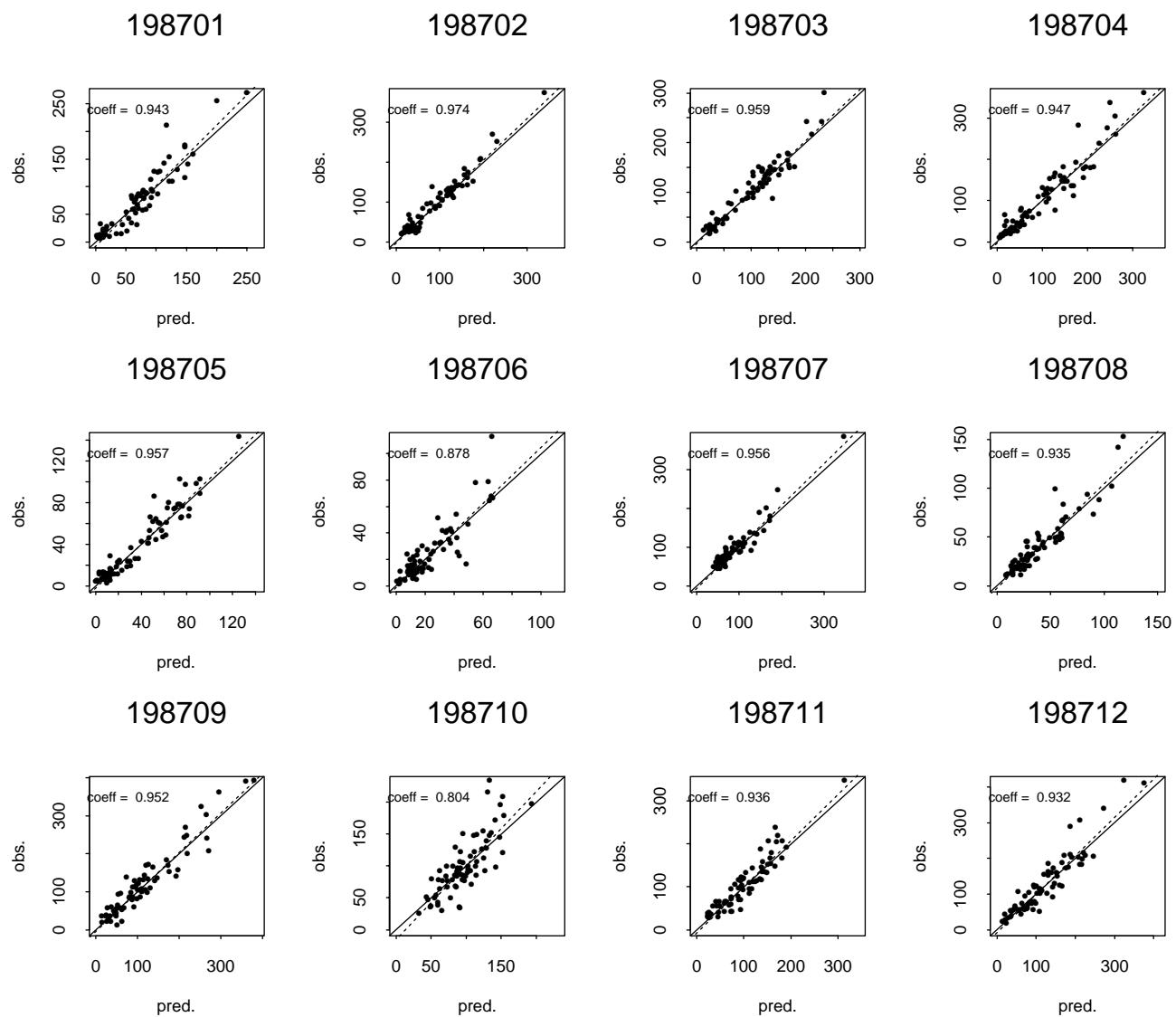


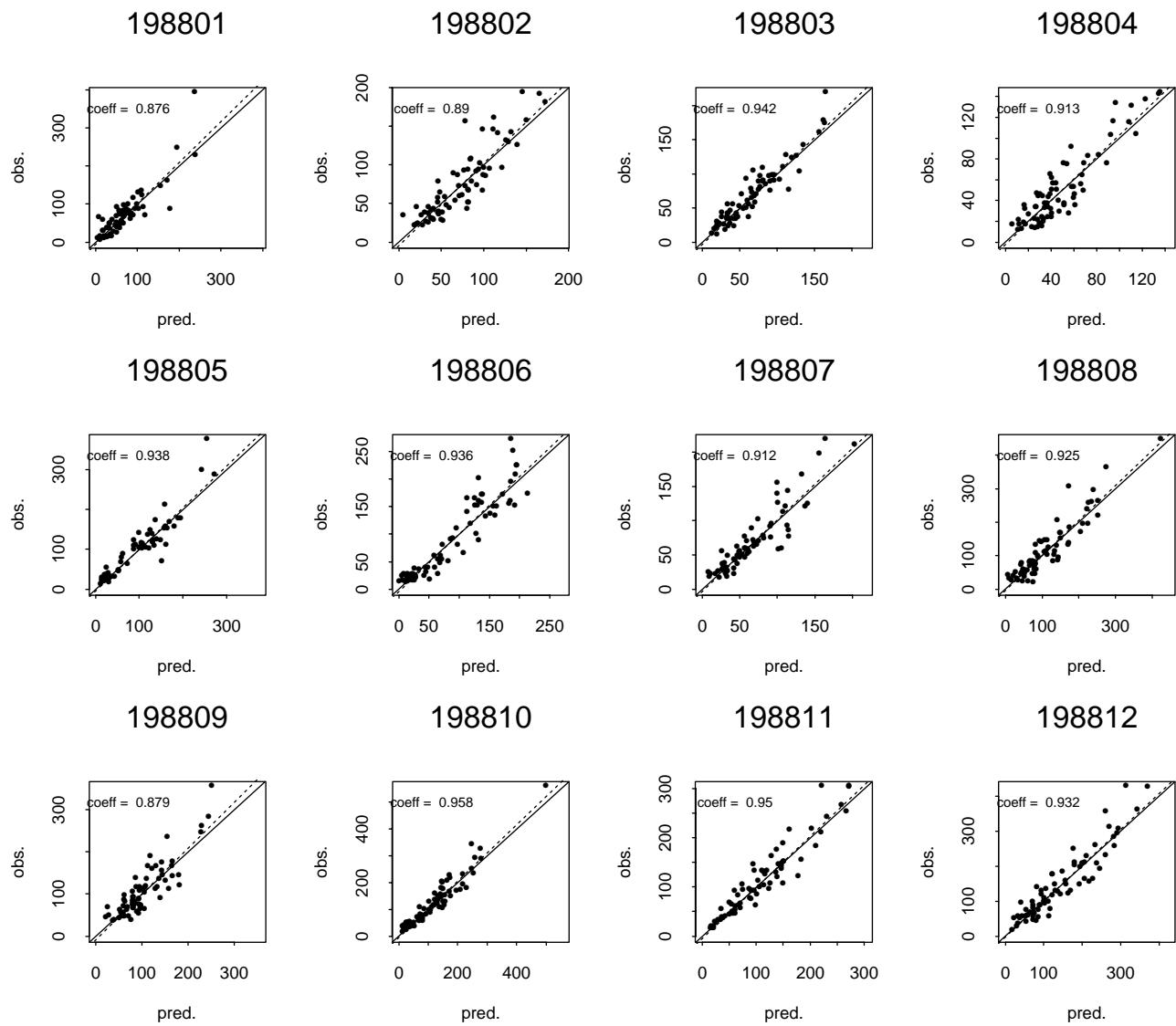


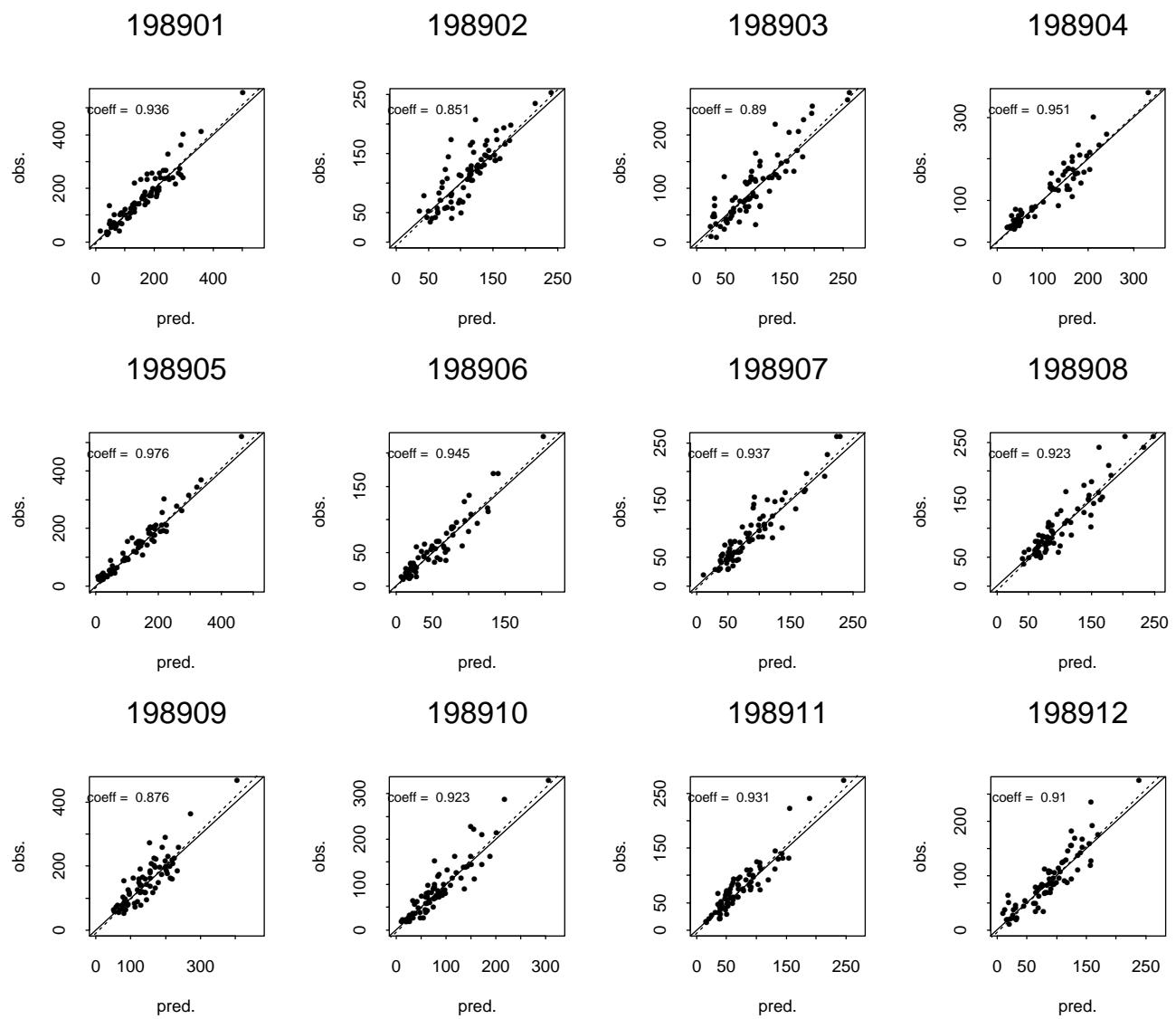


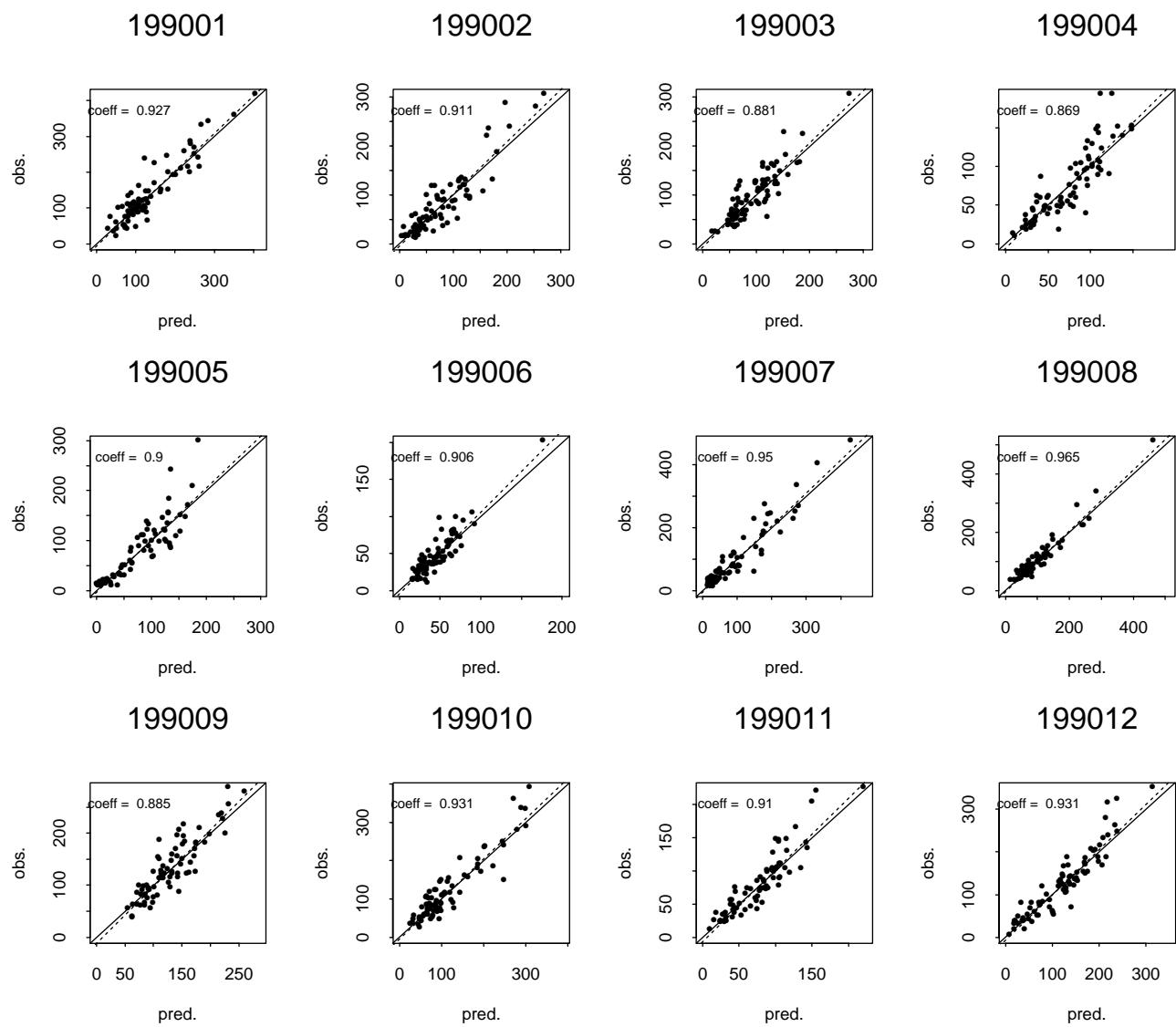


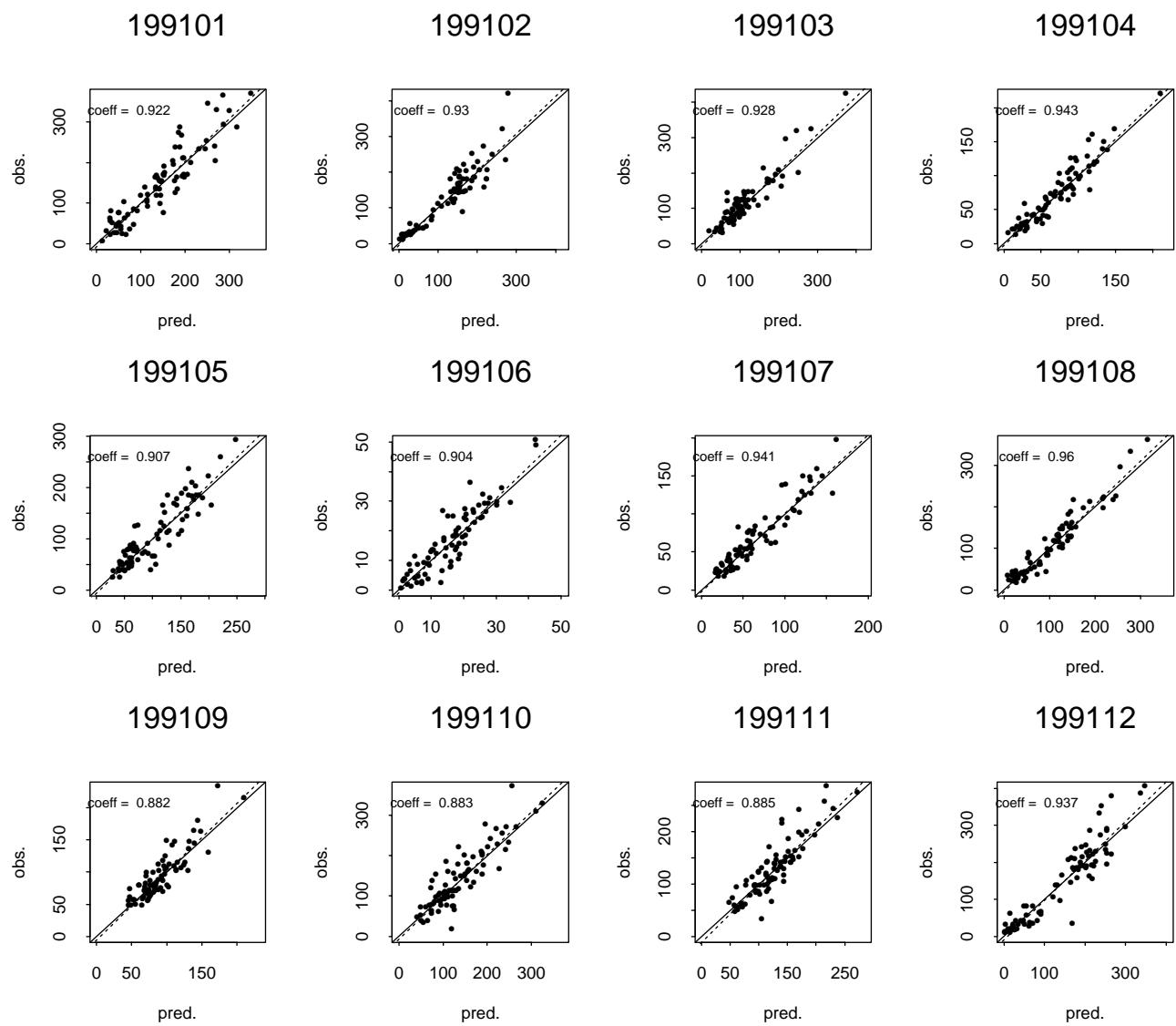


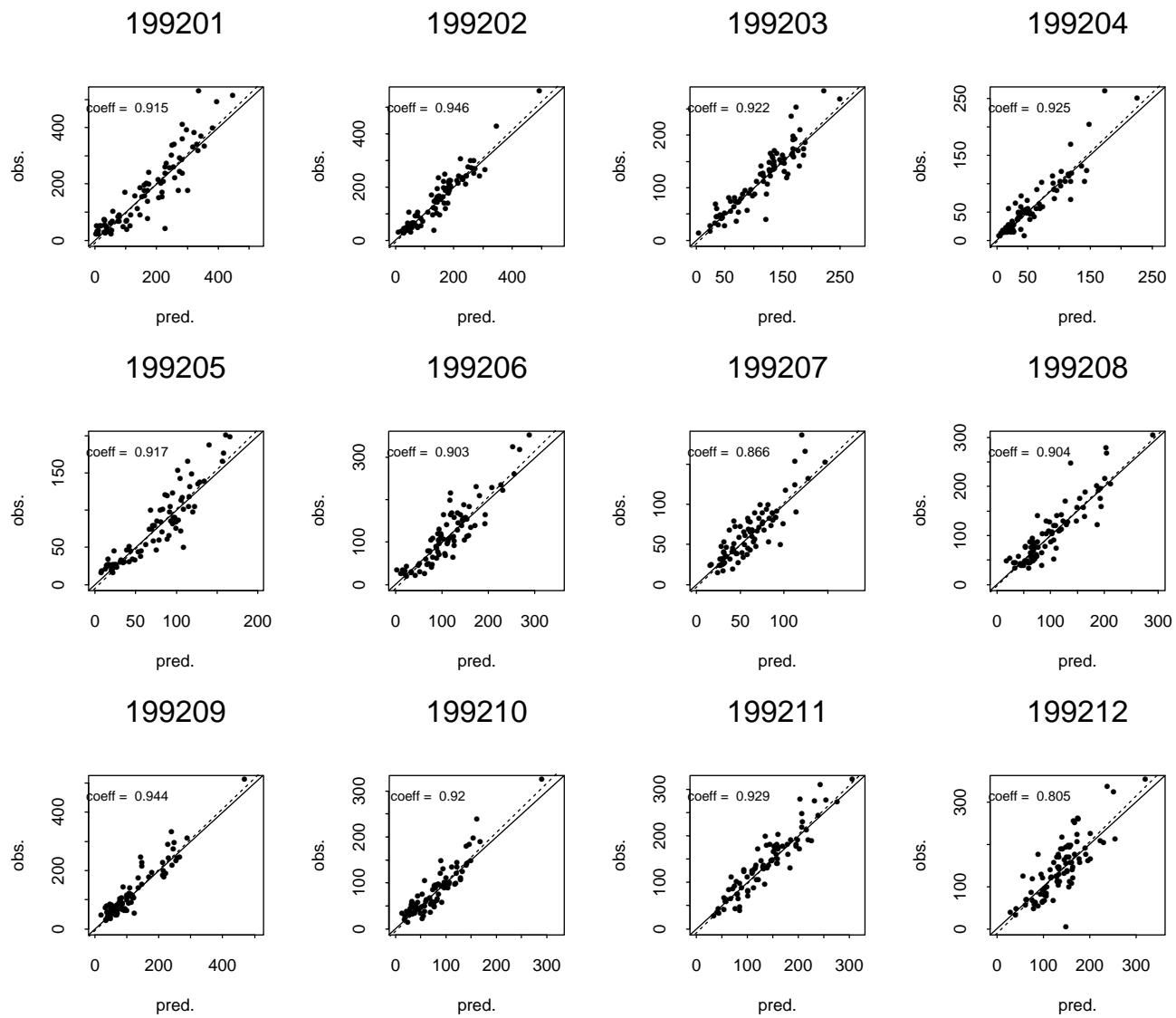


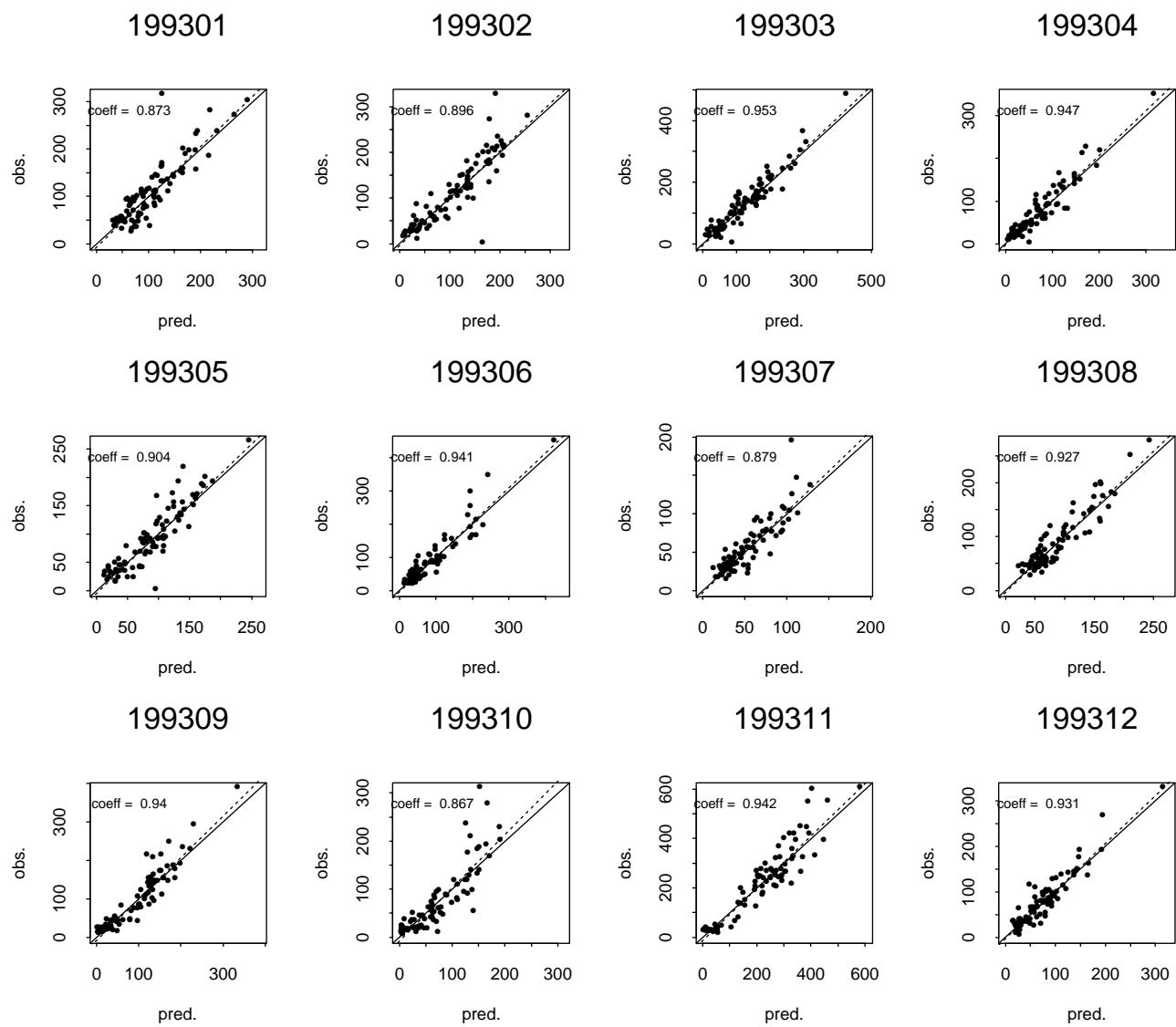


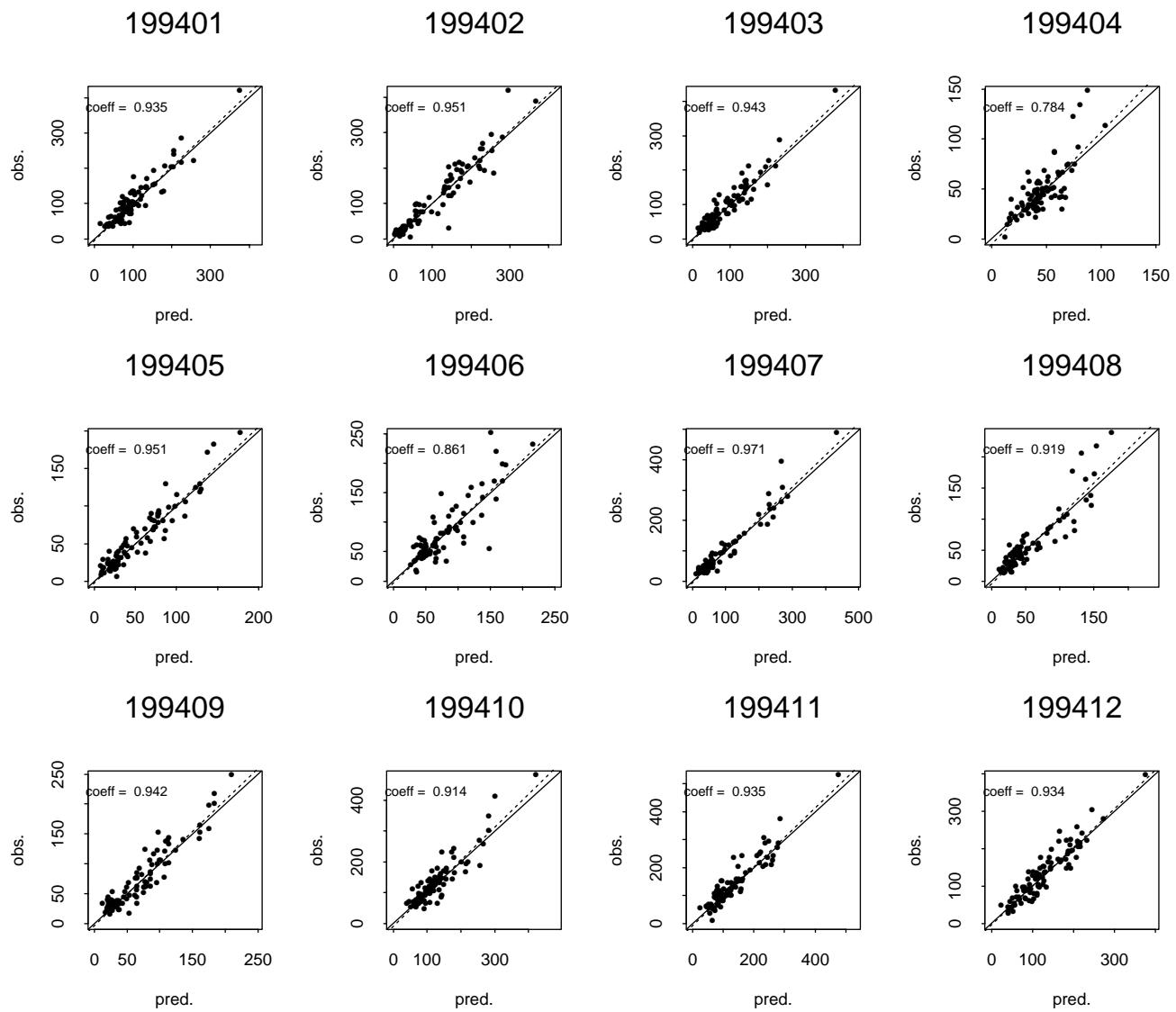


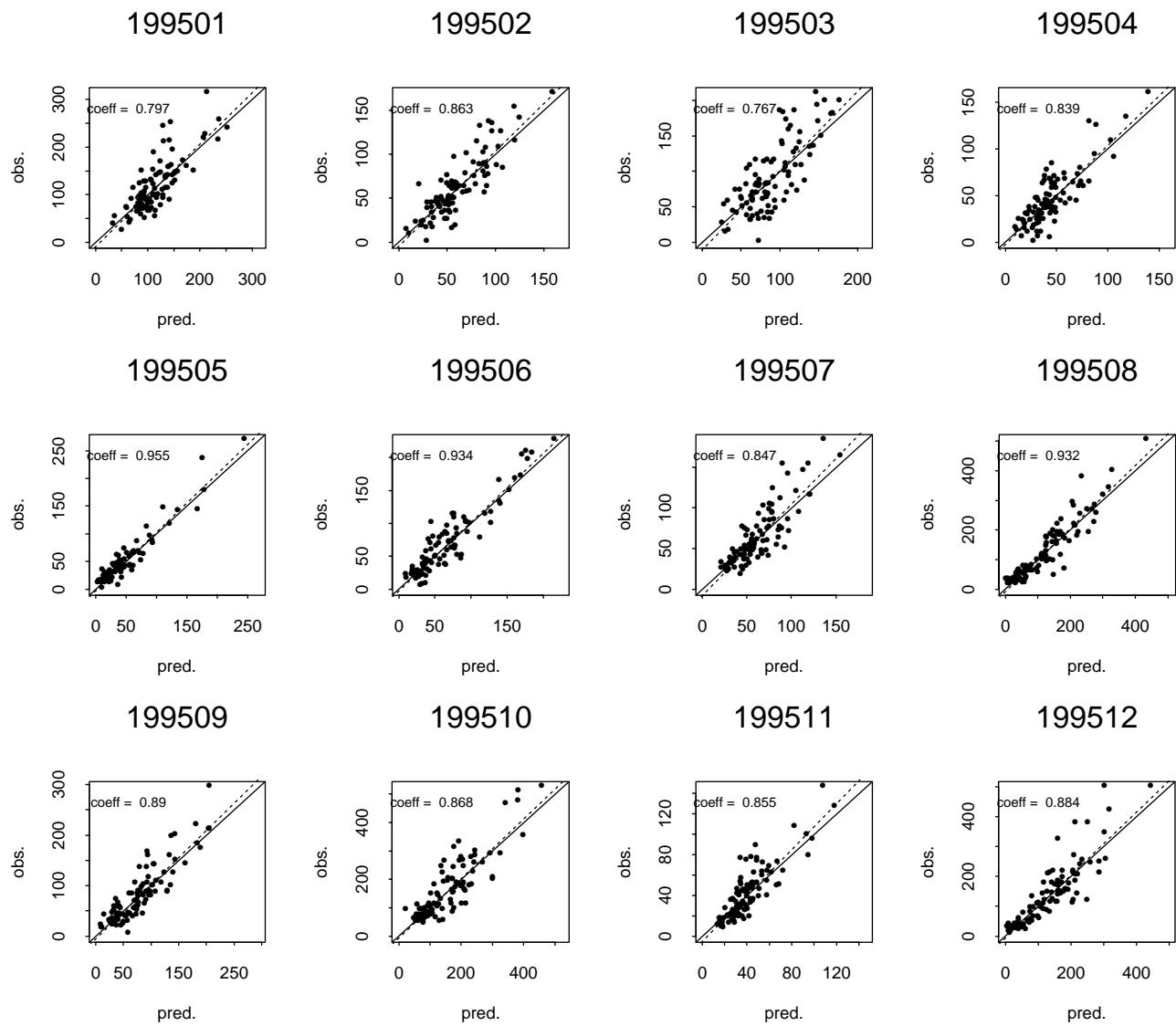


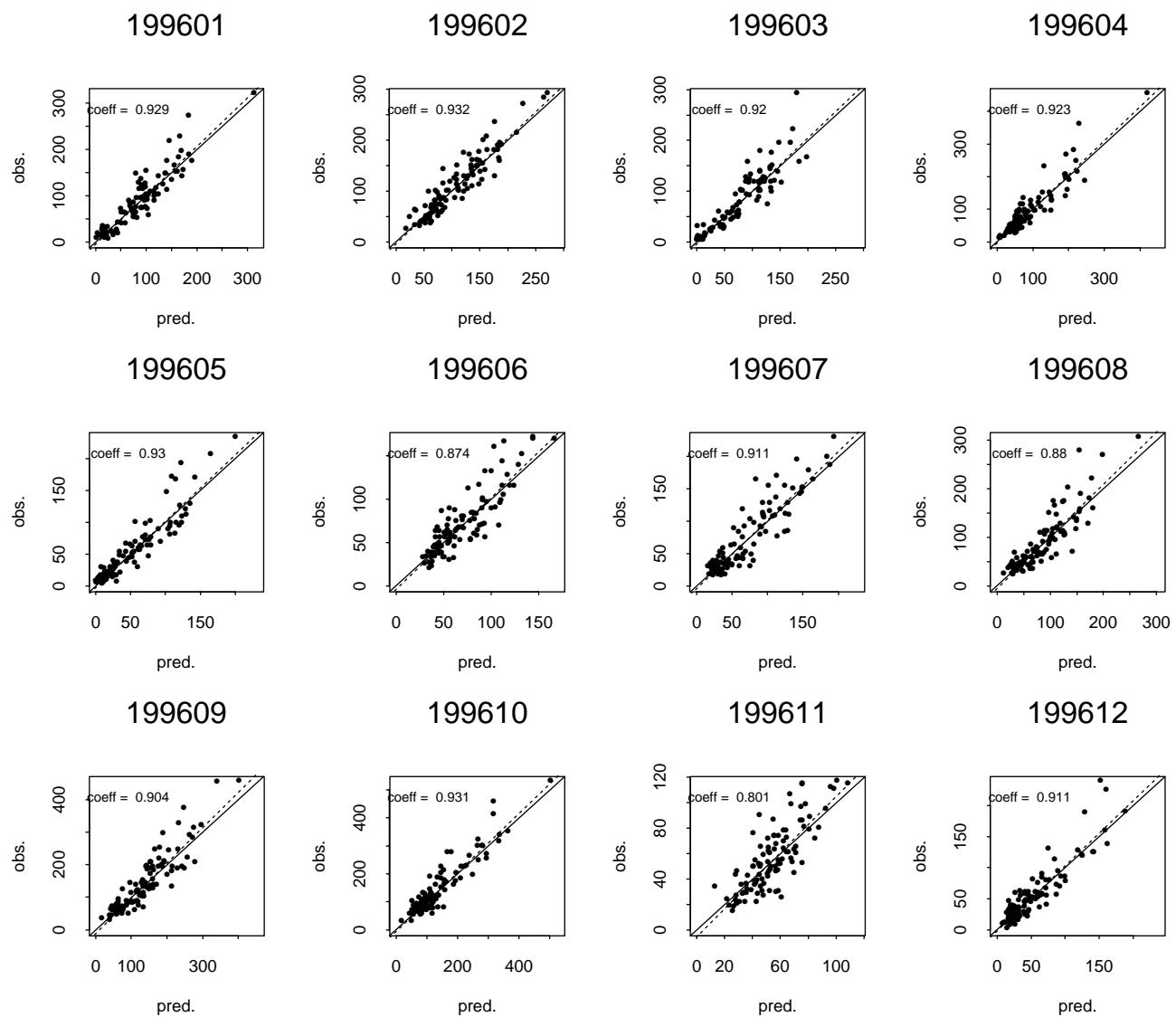


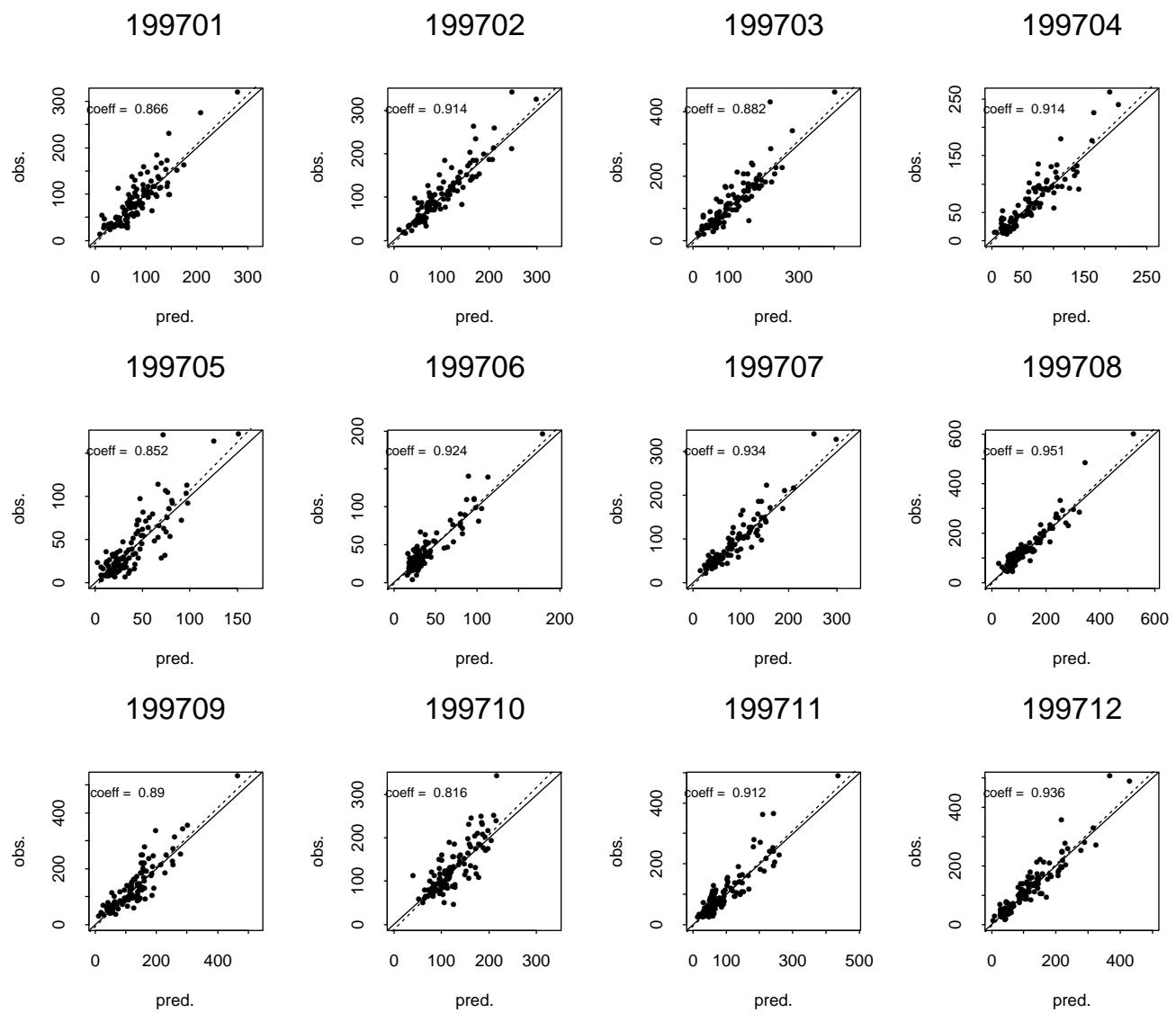


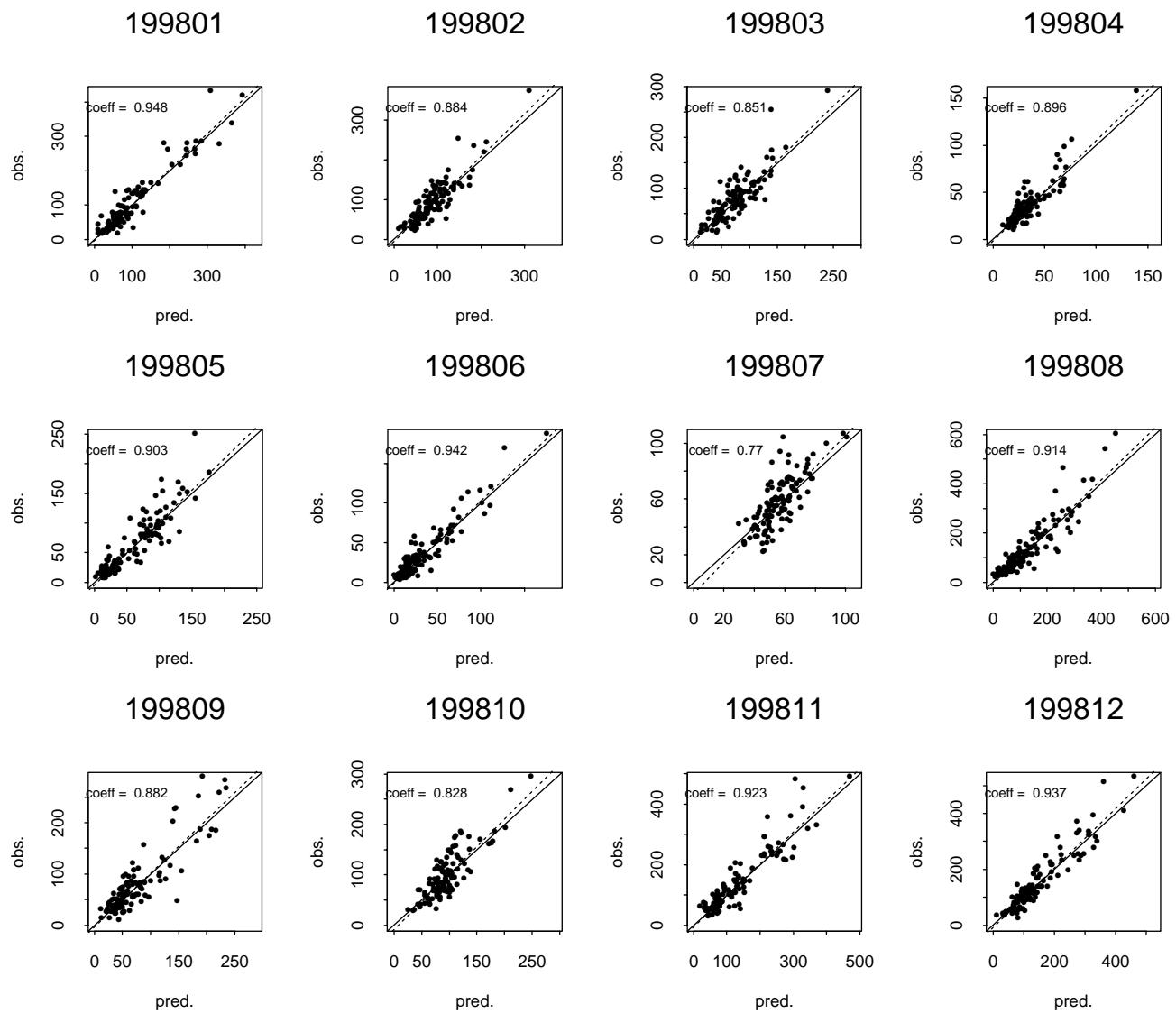


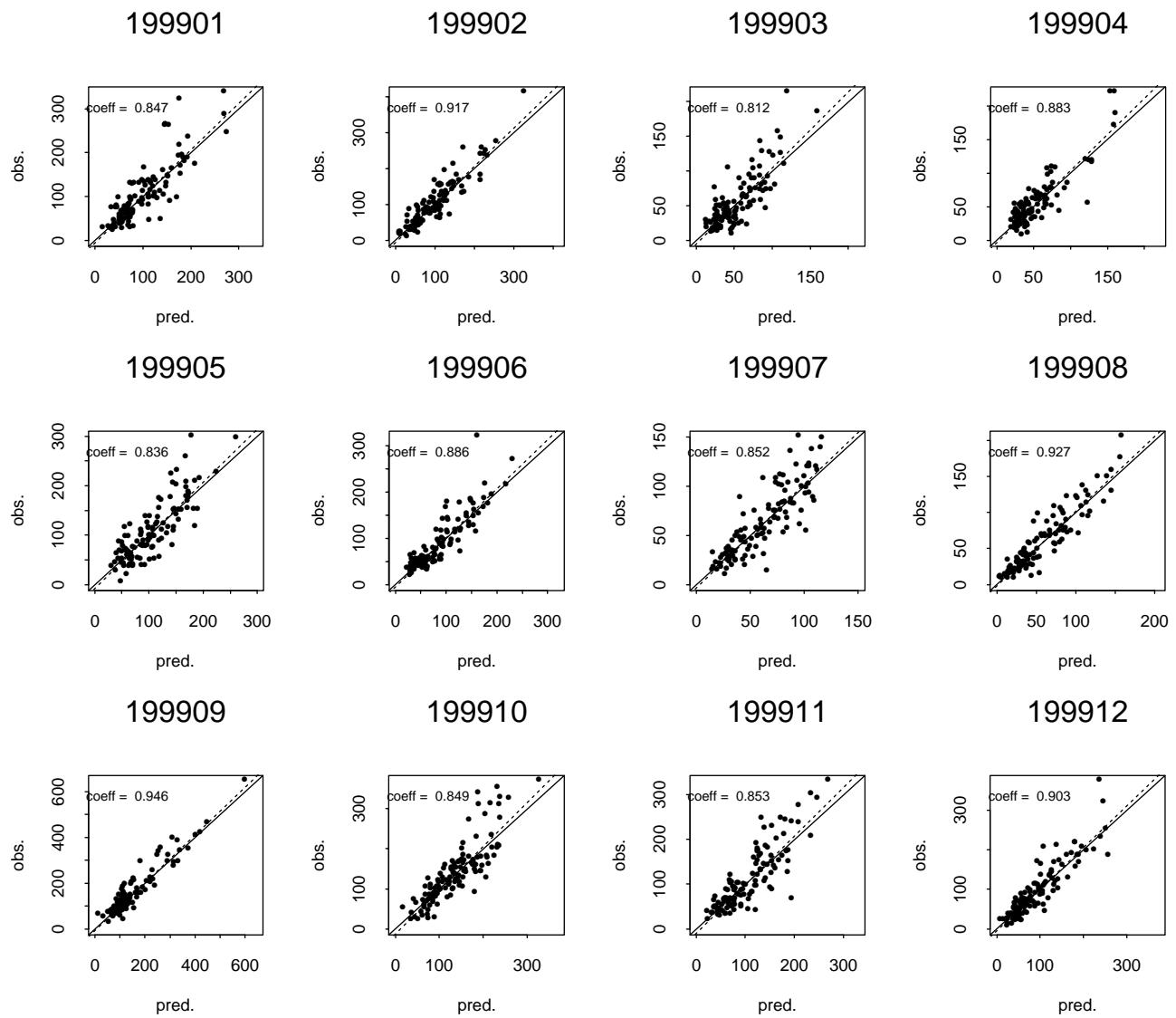


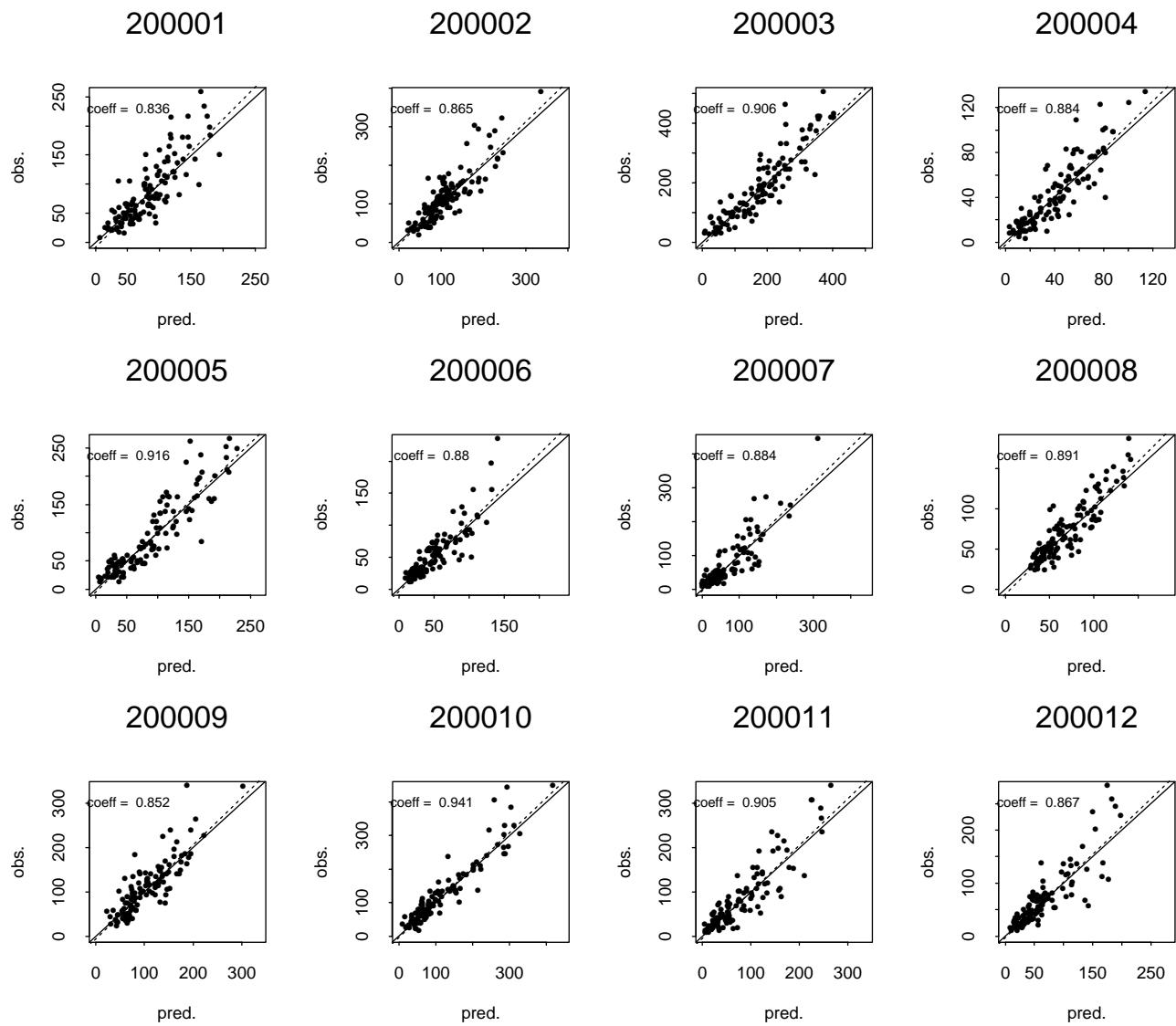








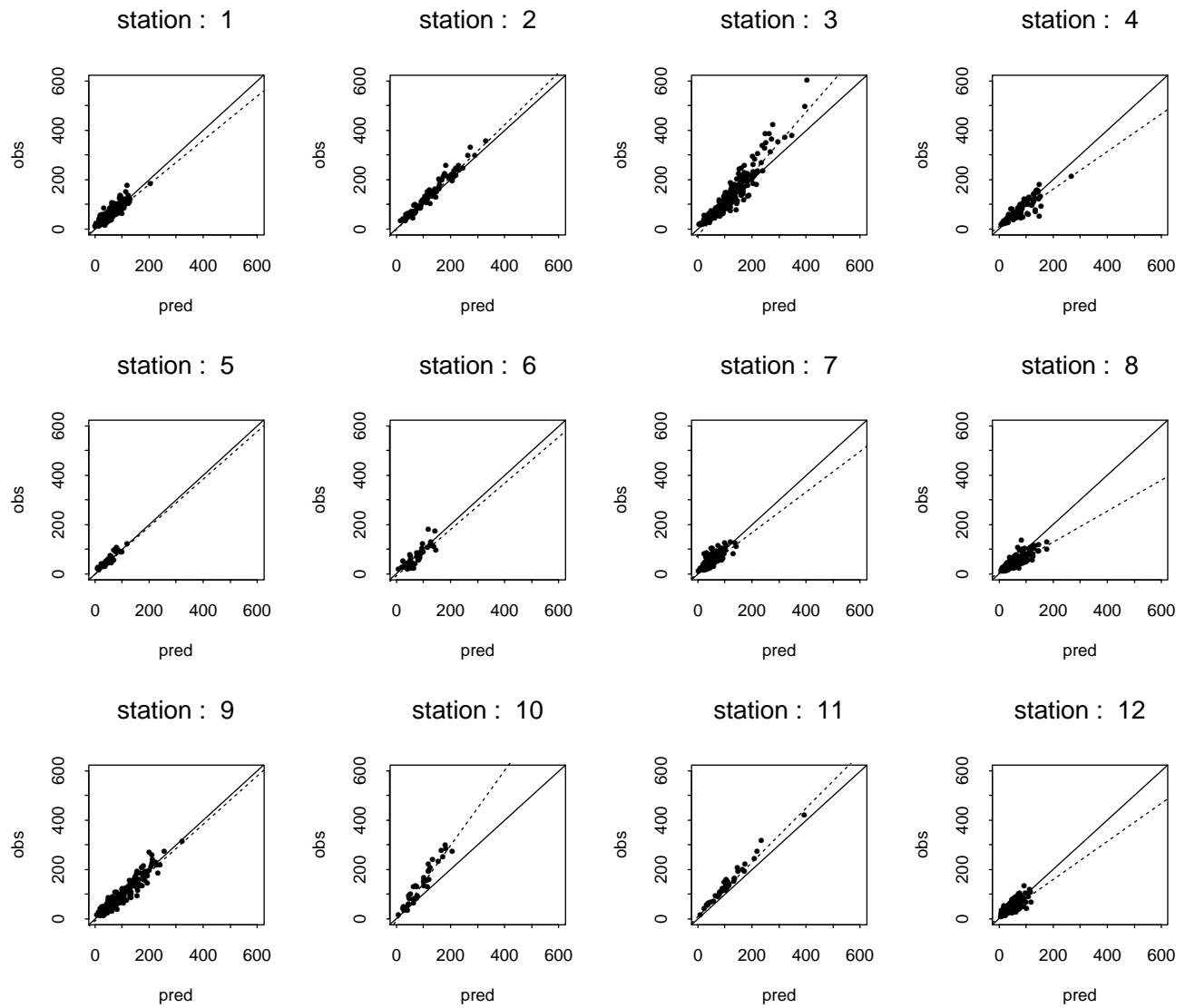


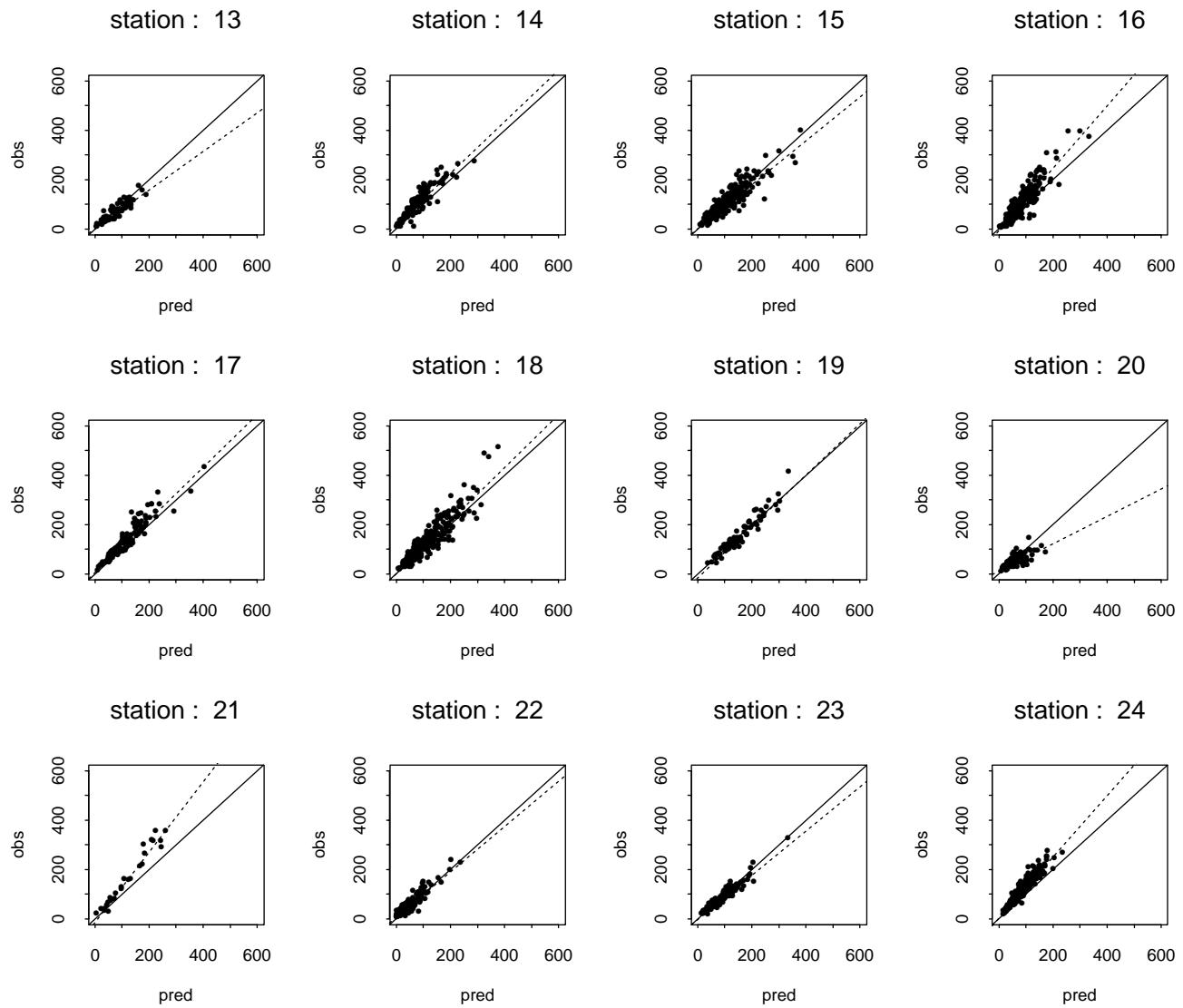


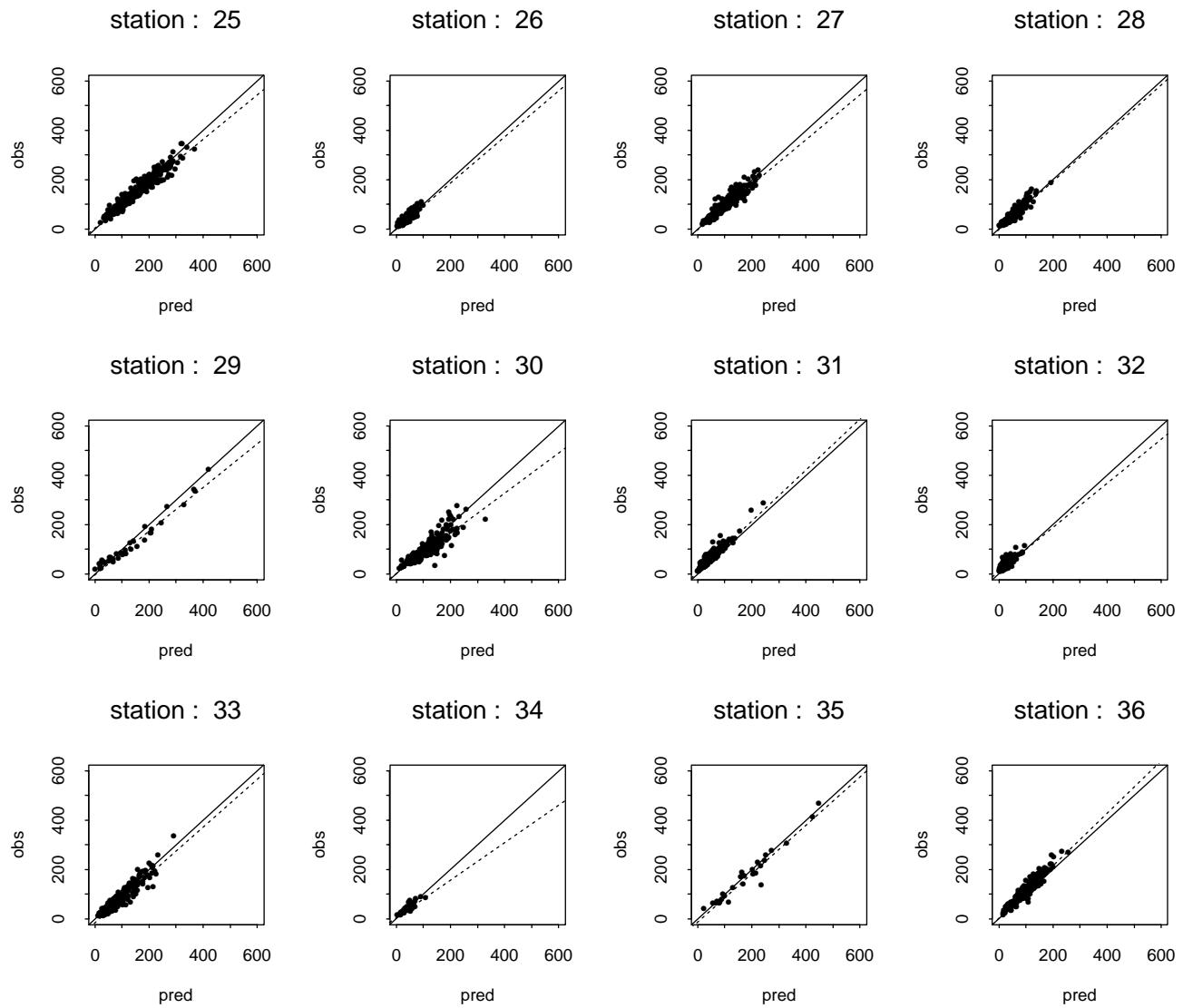
Appendix 5

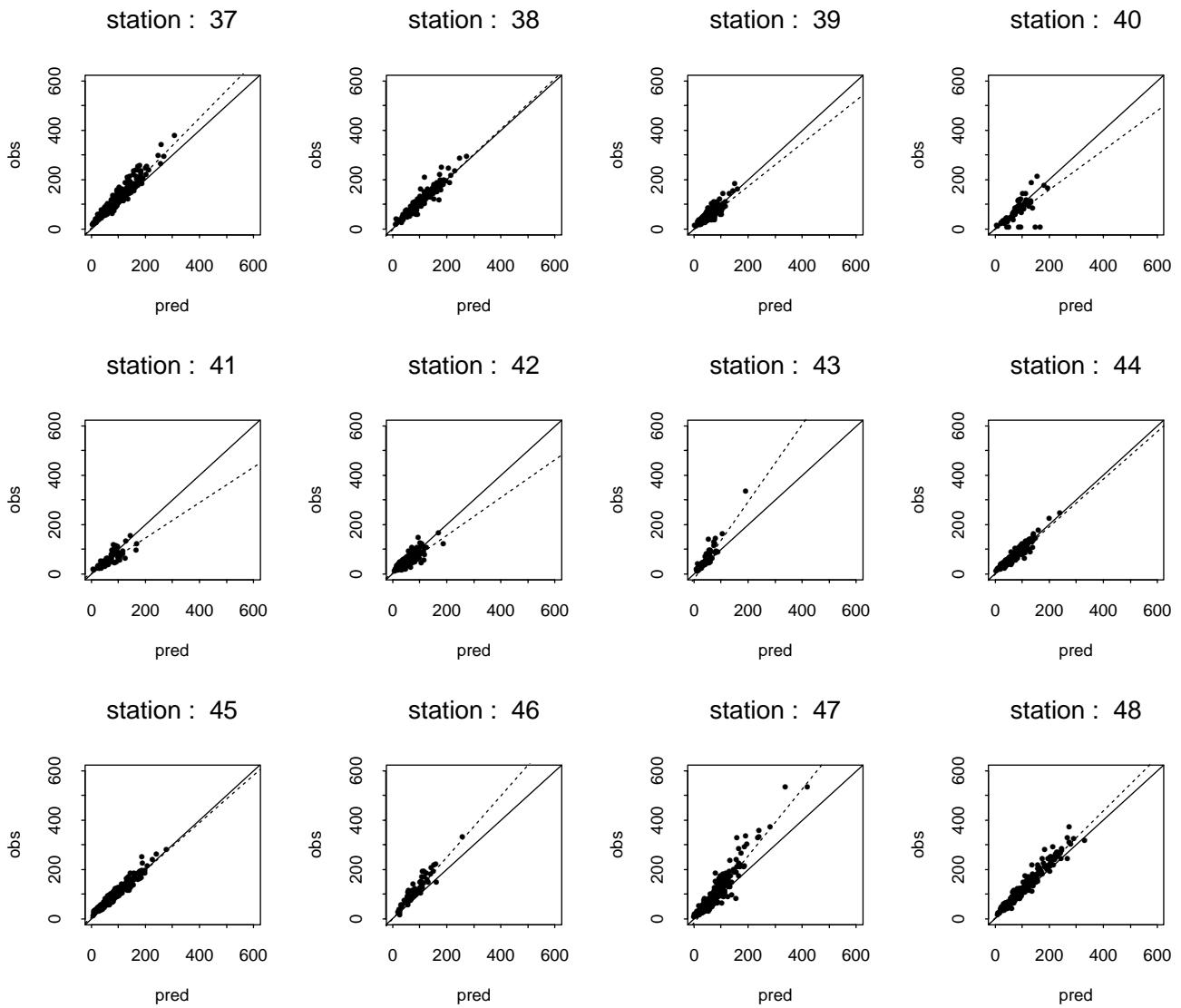
Monthly precipitation

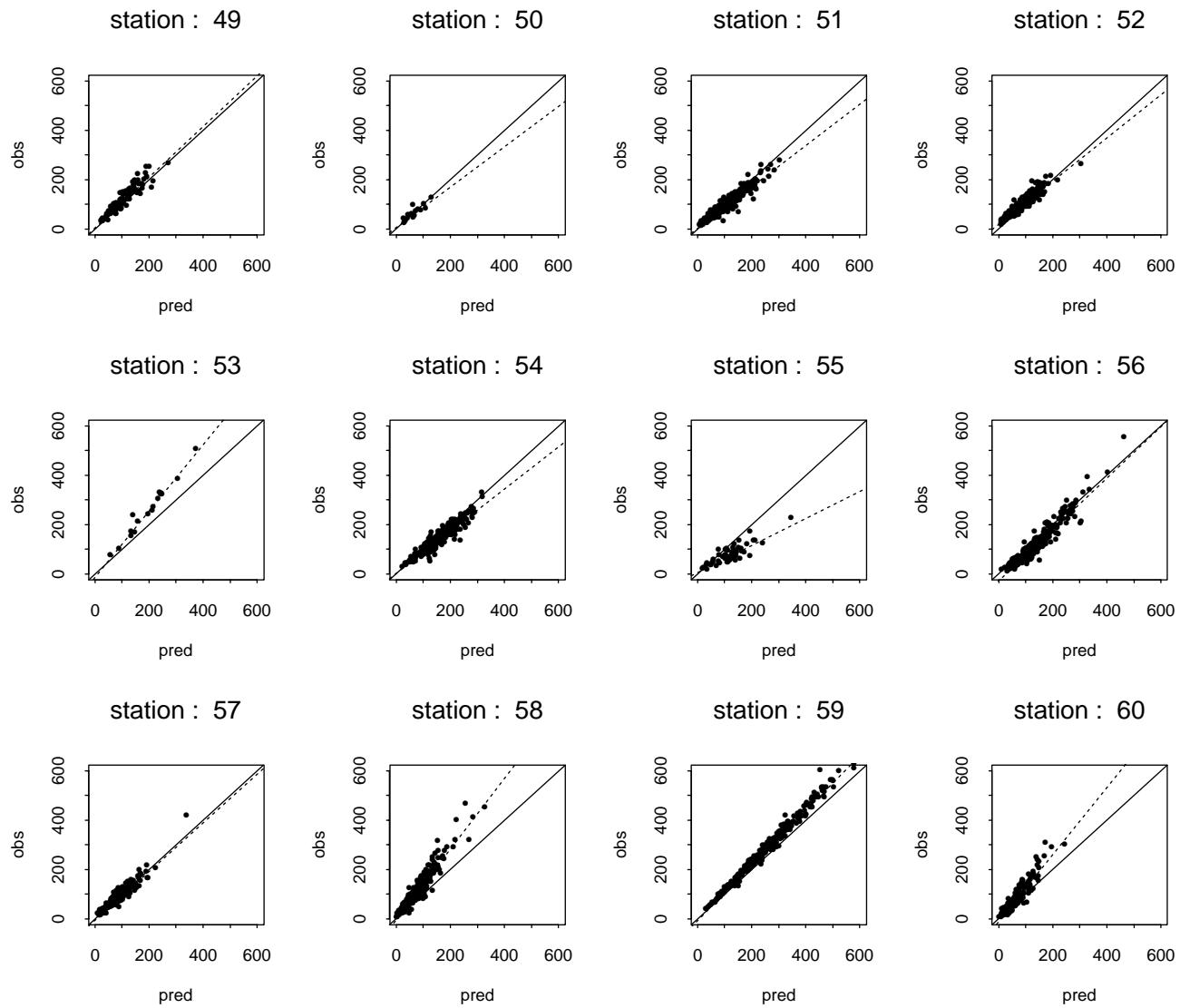
Scatter plots for each station (period 1980-2000)

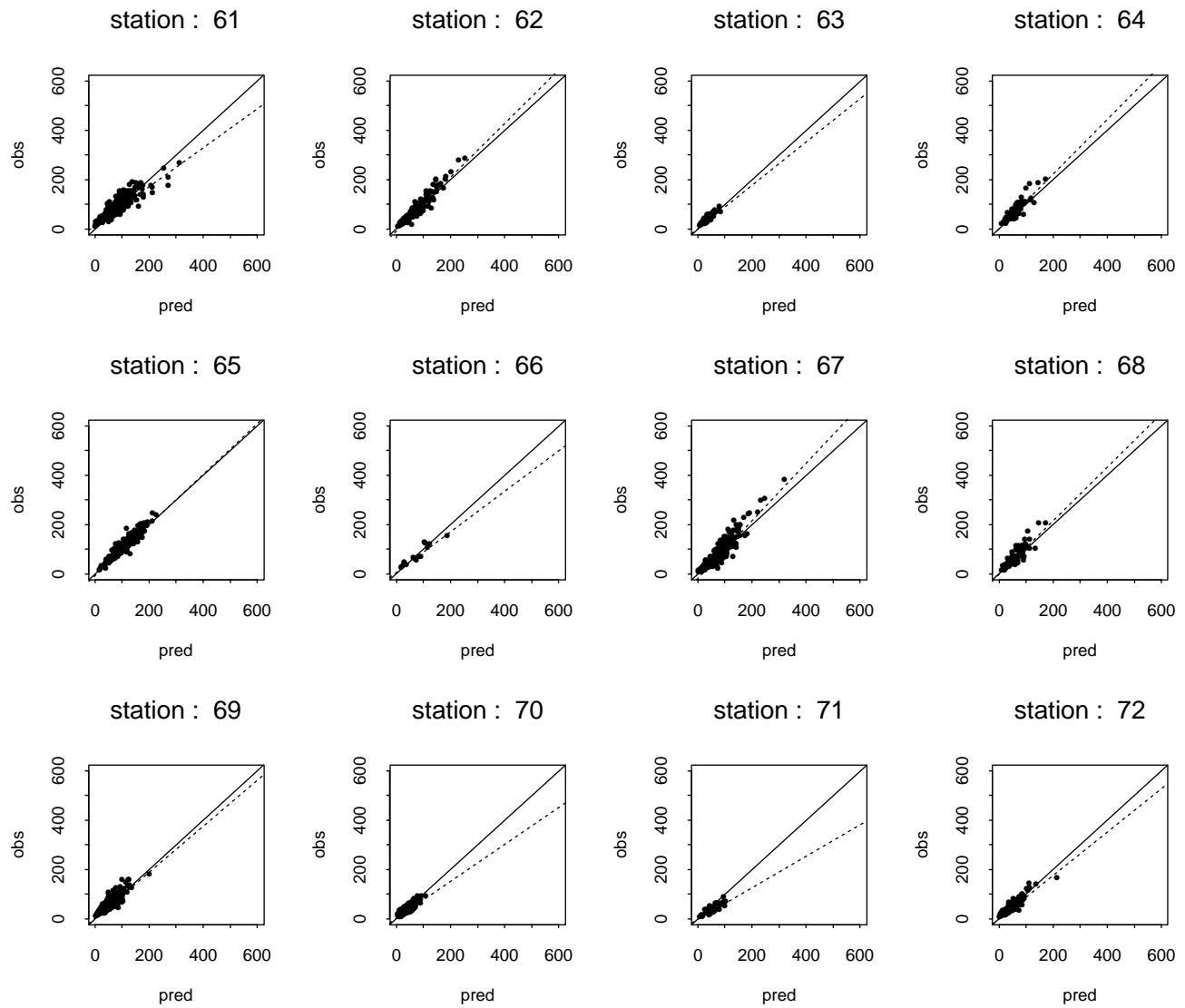


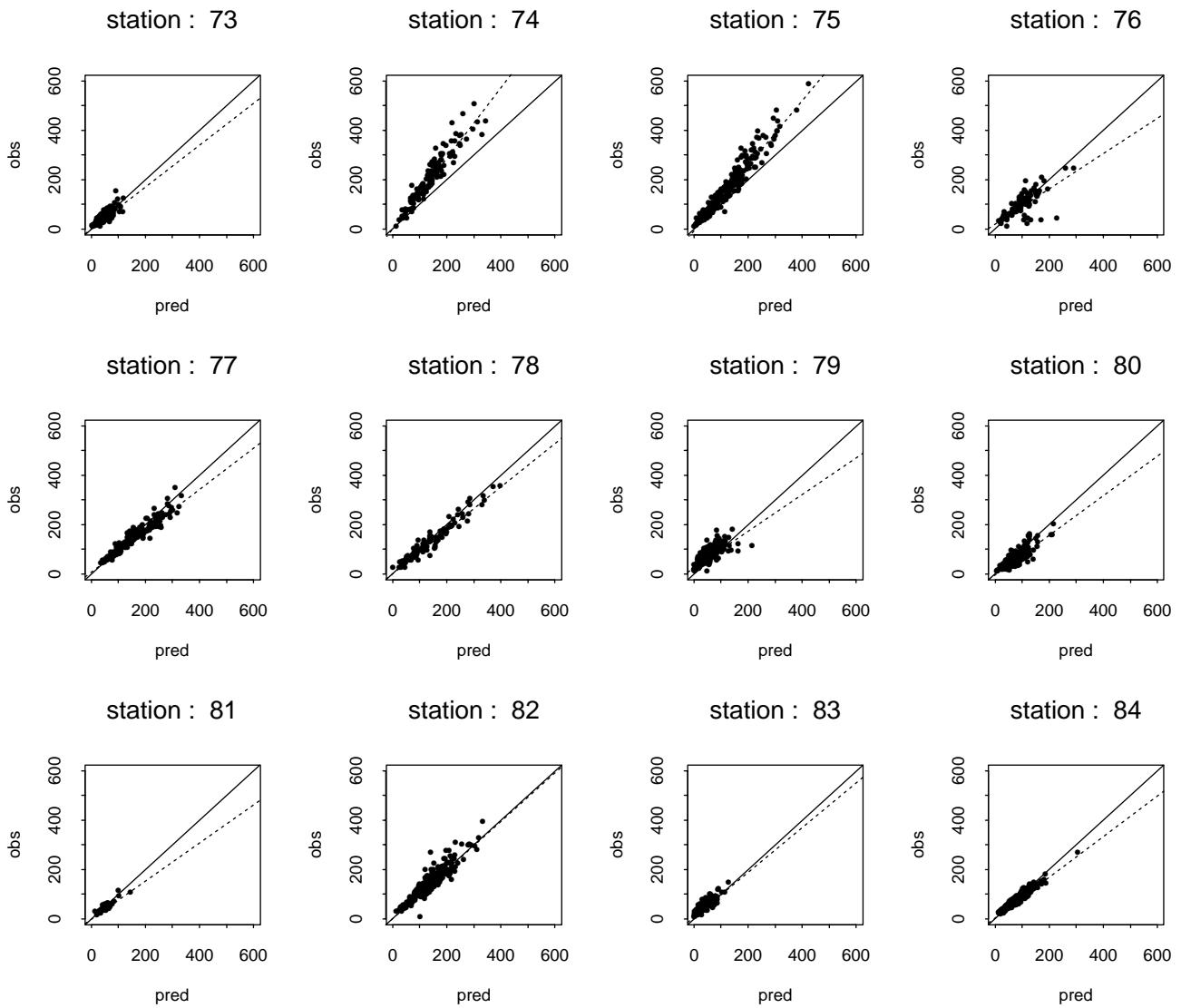


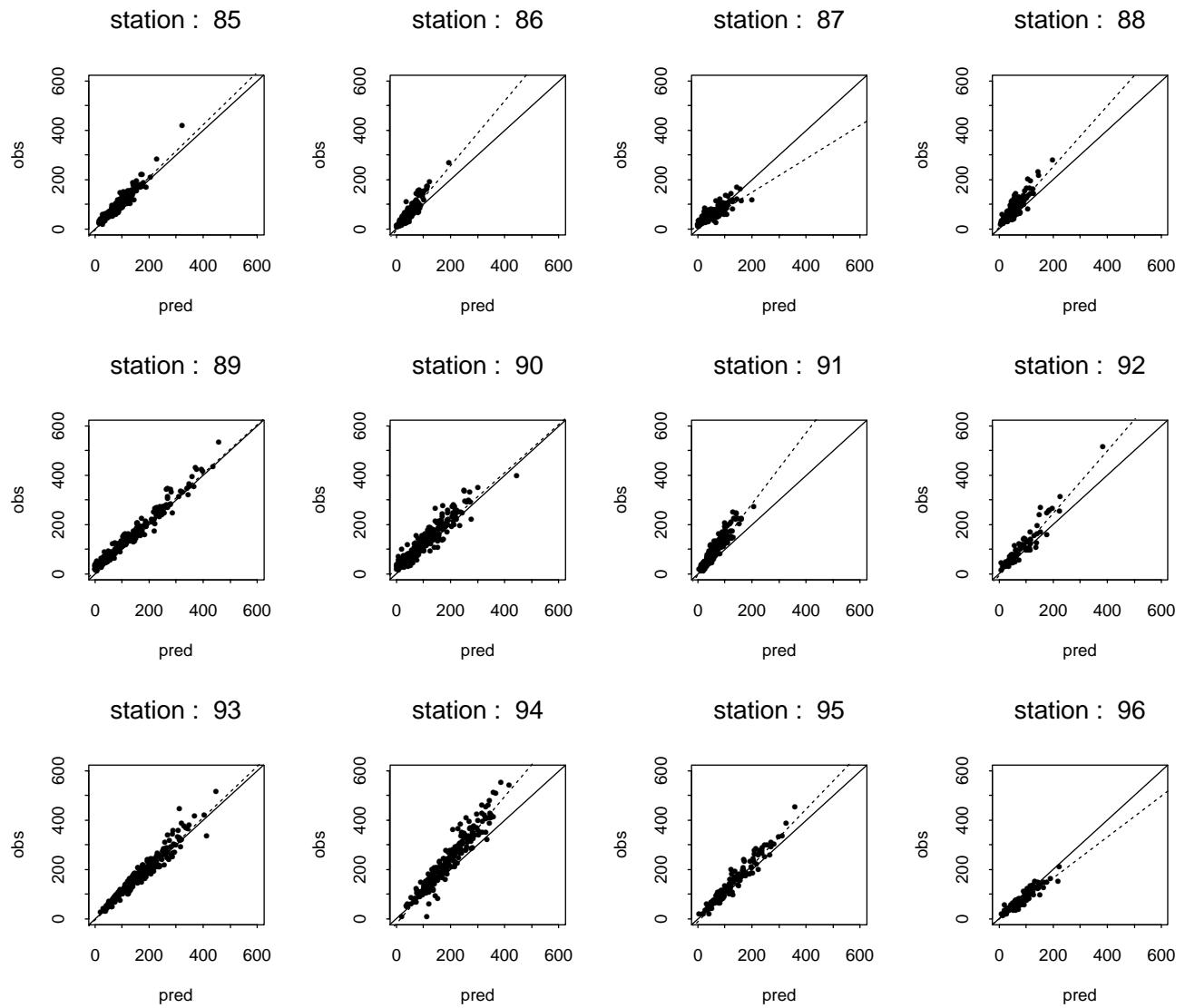


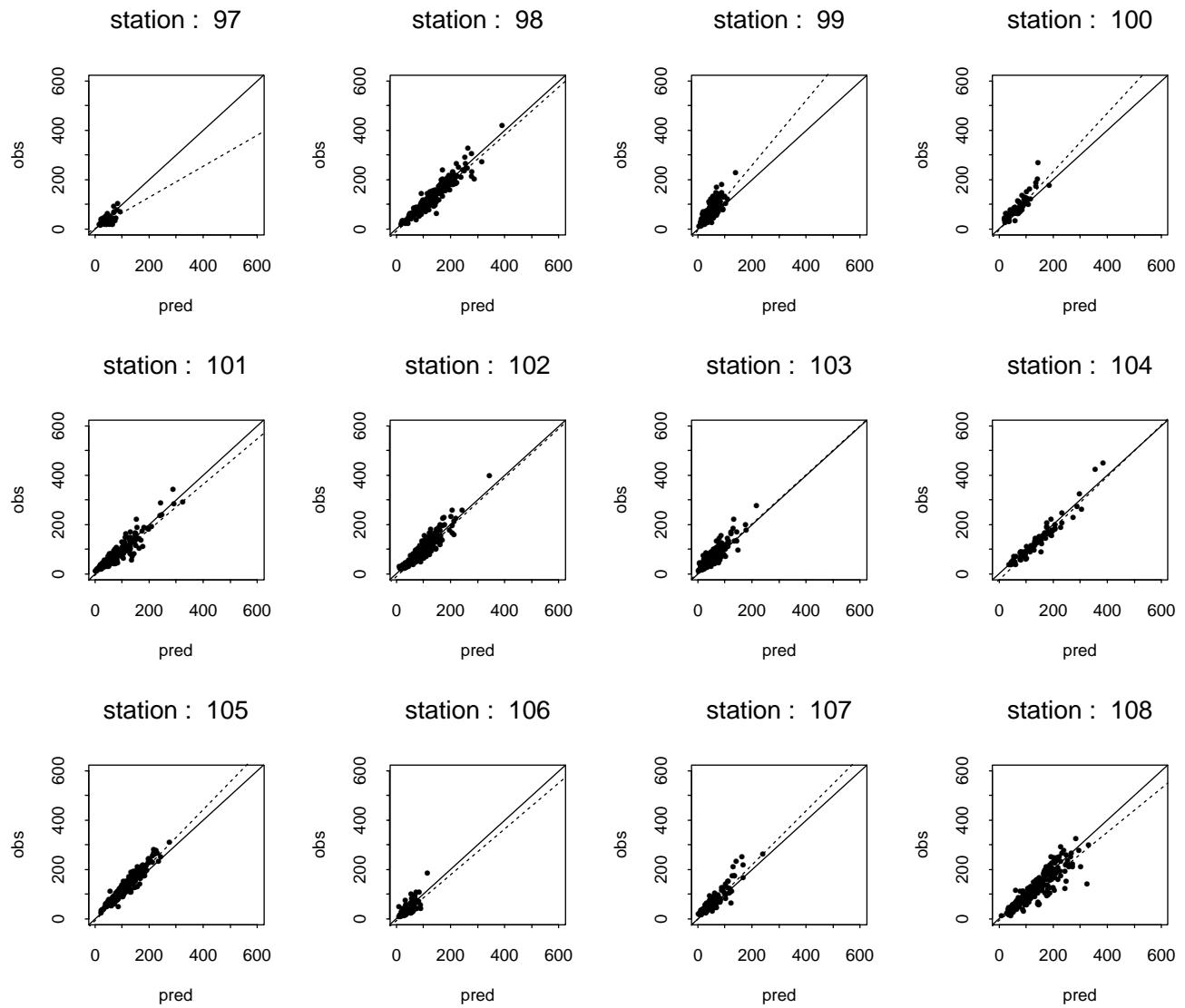


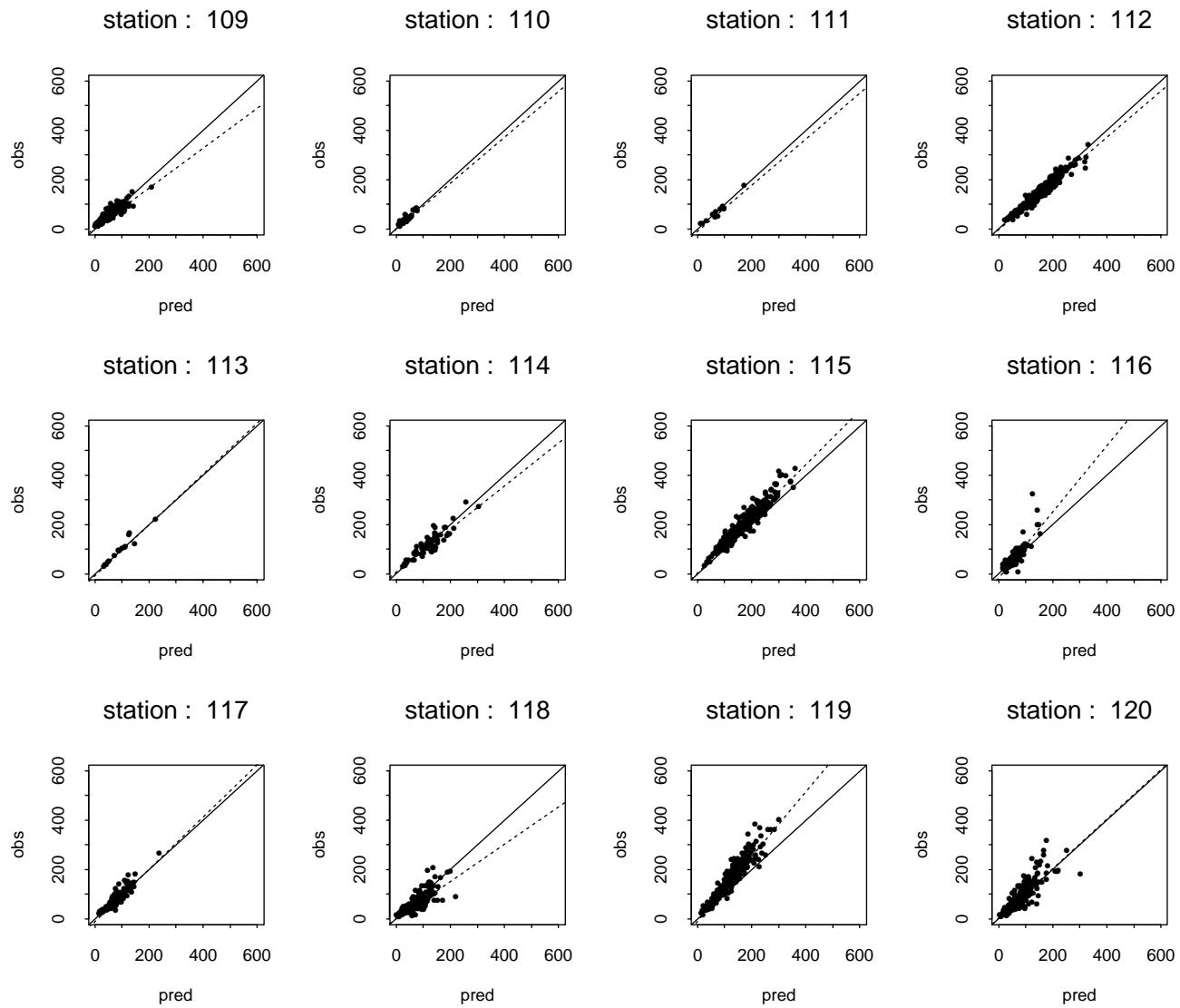


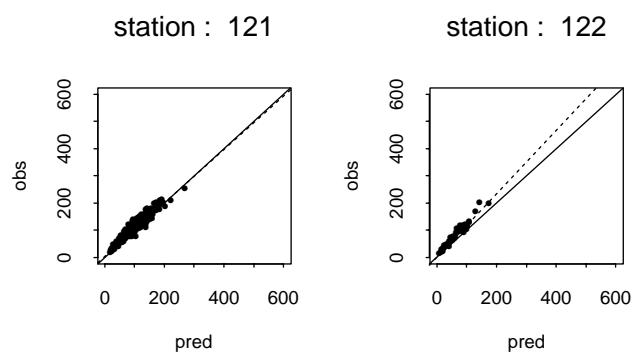










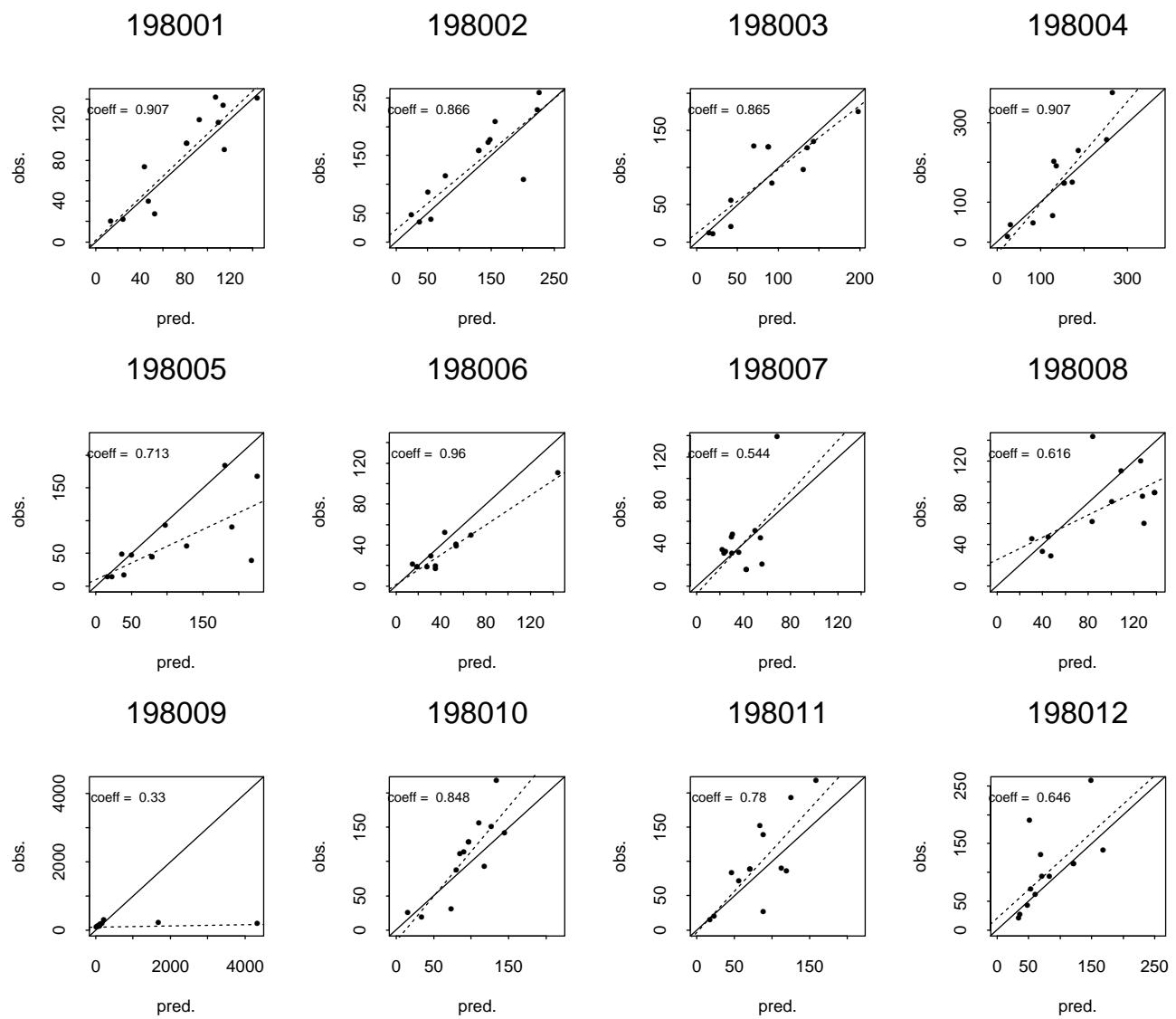


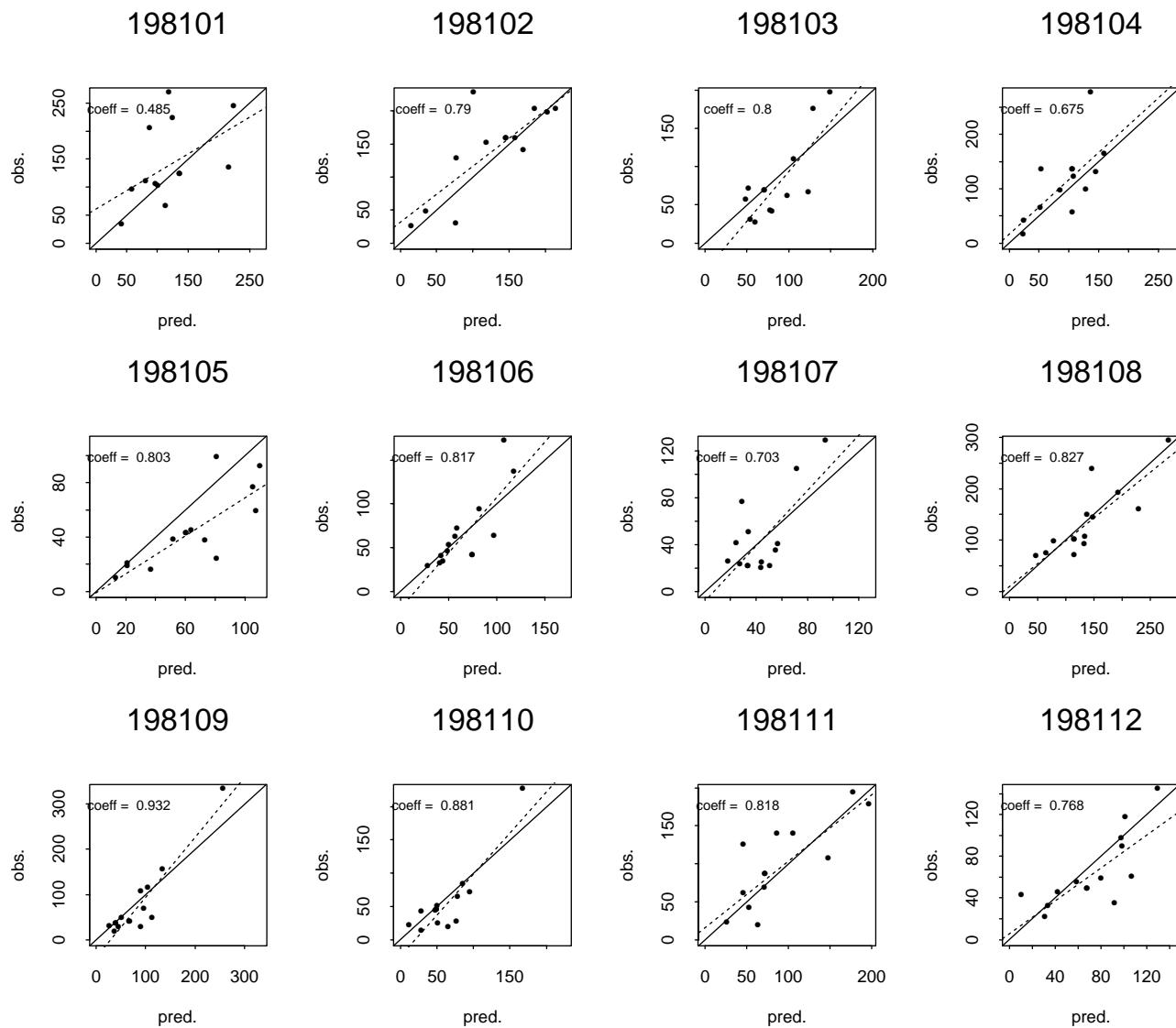
Appendix 6

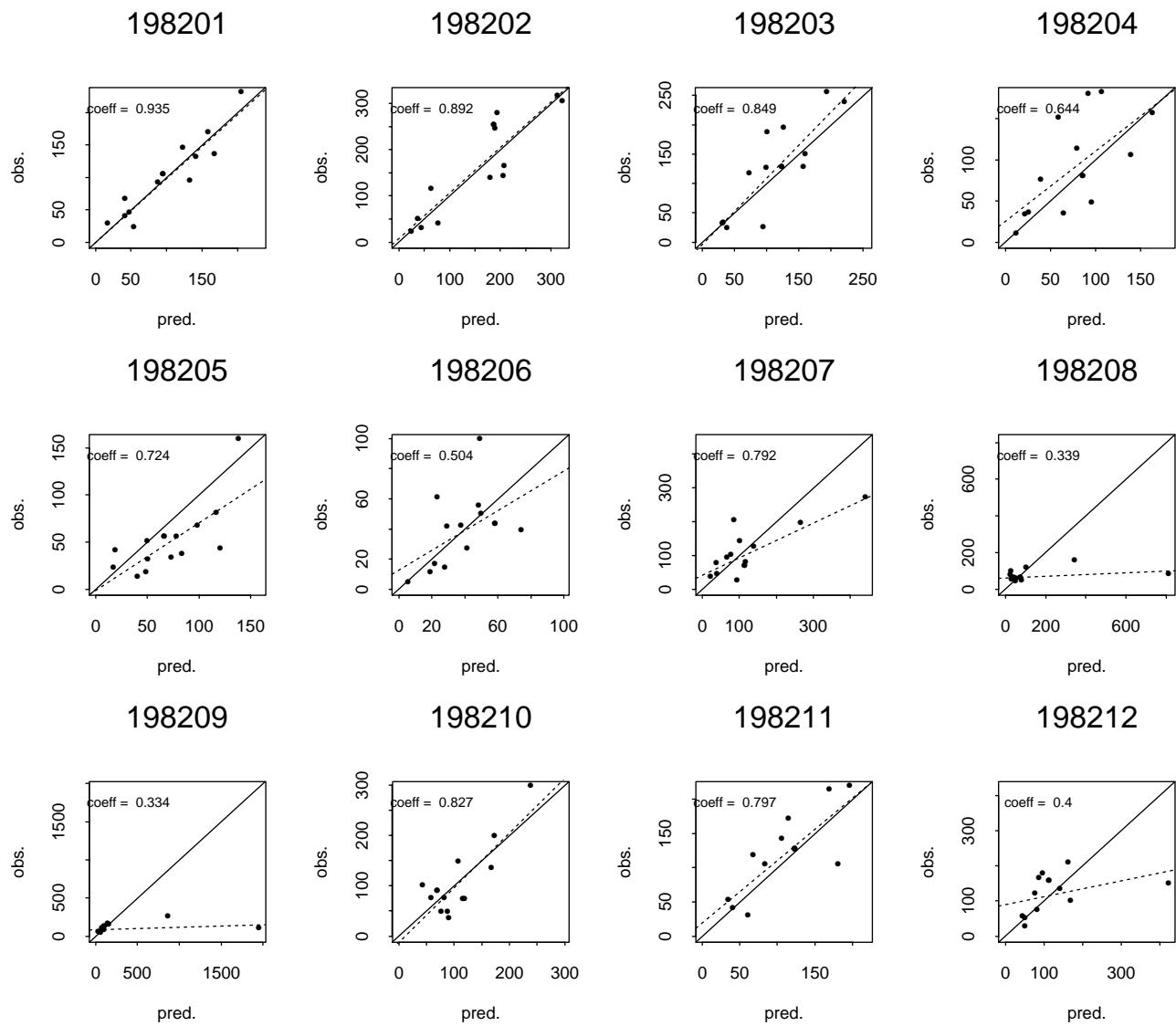
Monthly precipitation

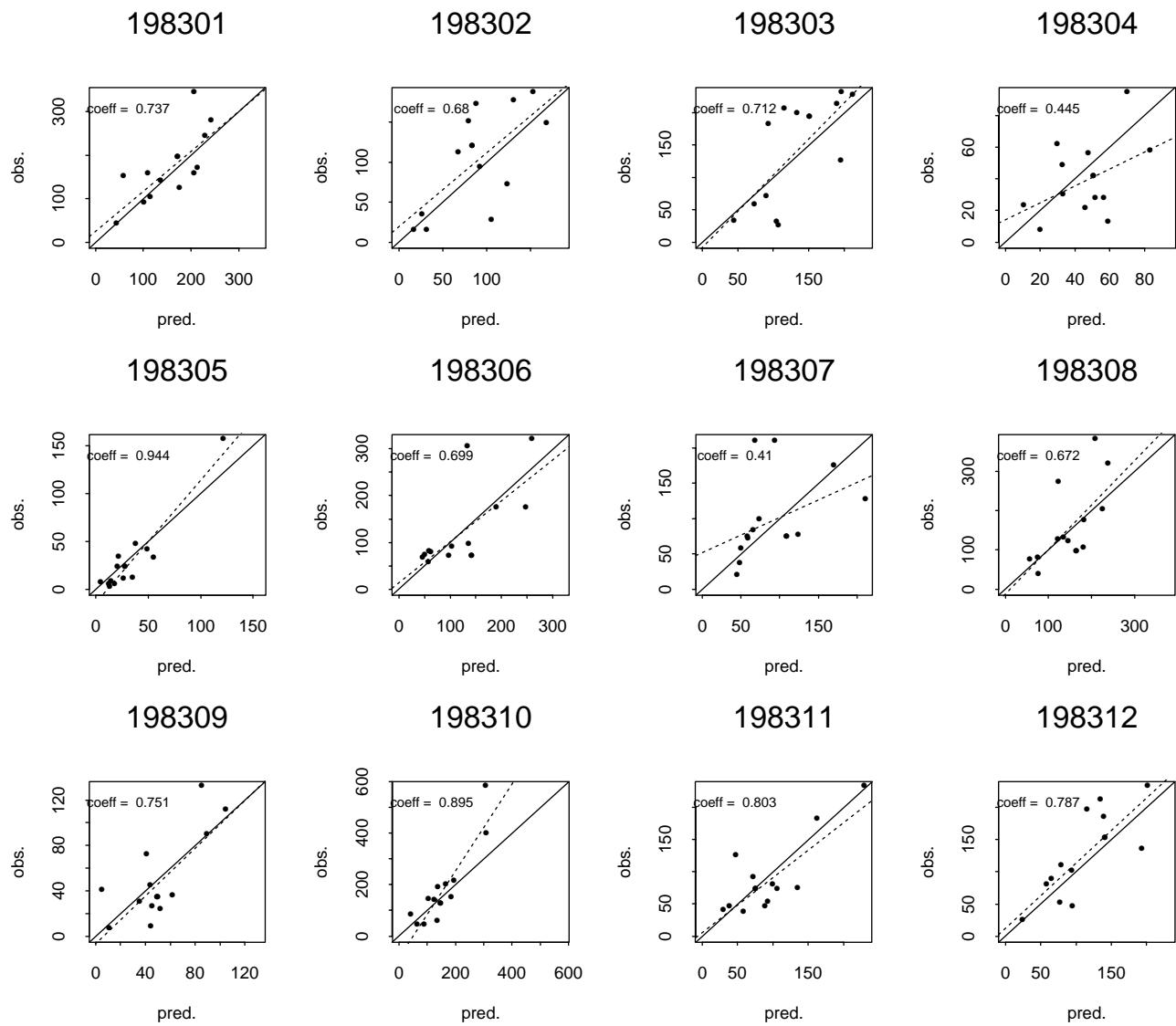
Validation procedure

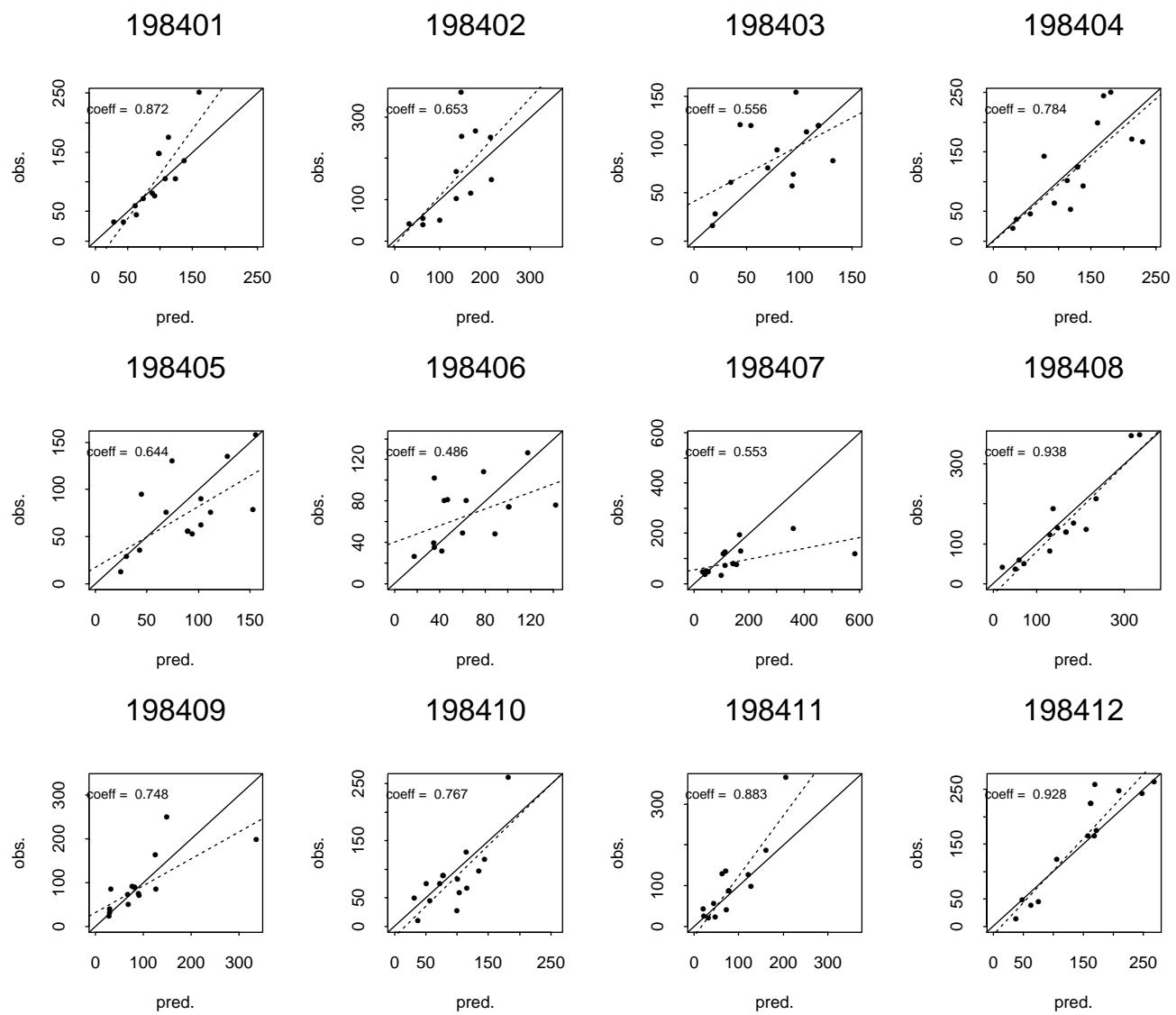
Scatter plots for the period 1980-2000

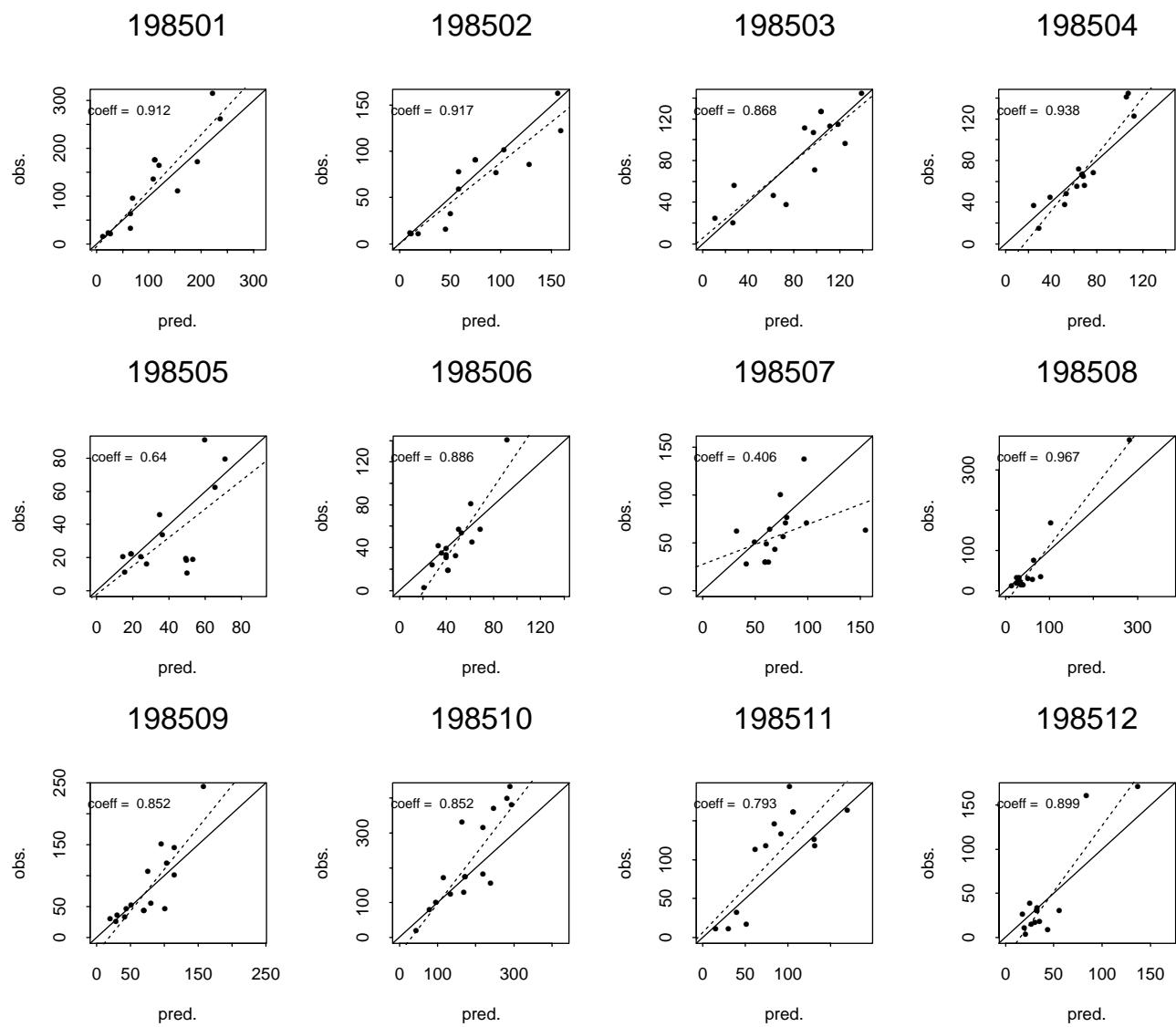


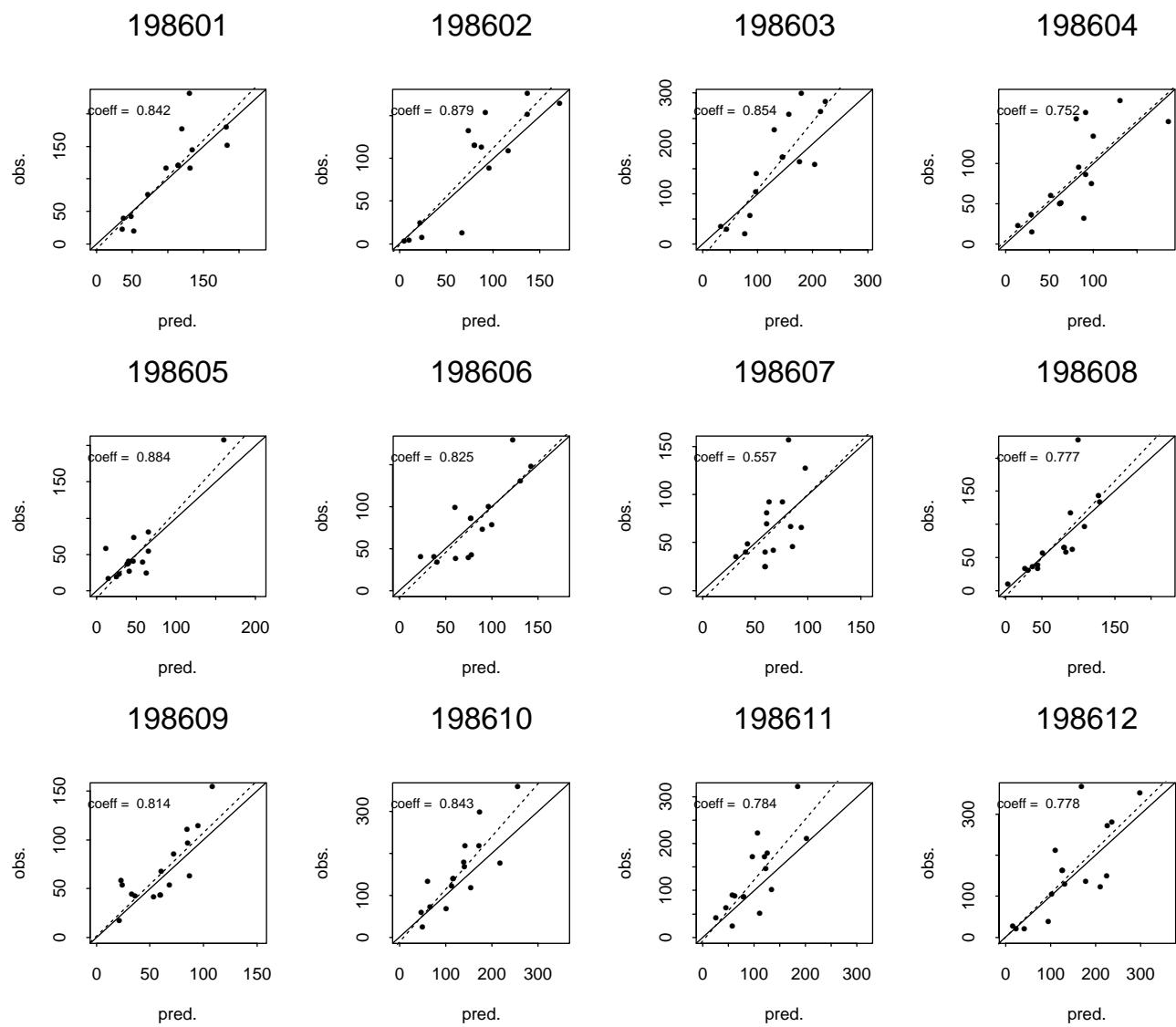


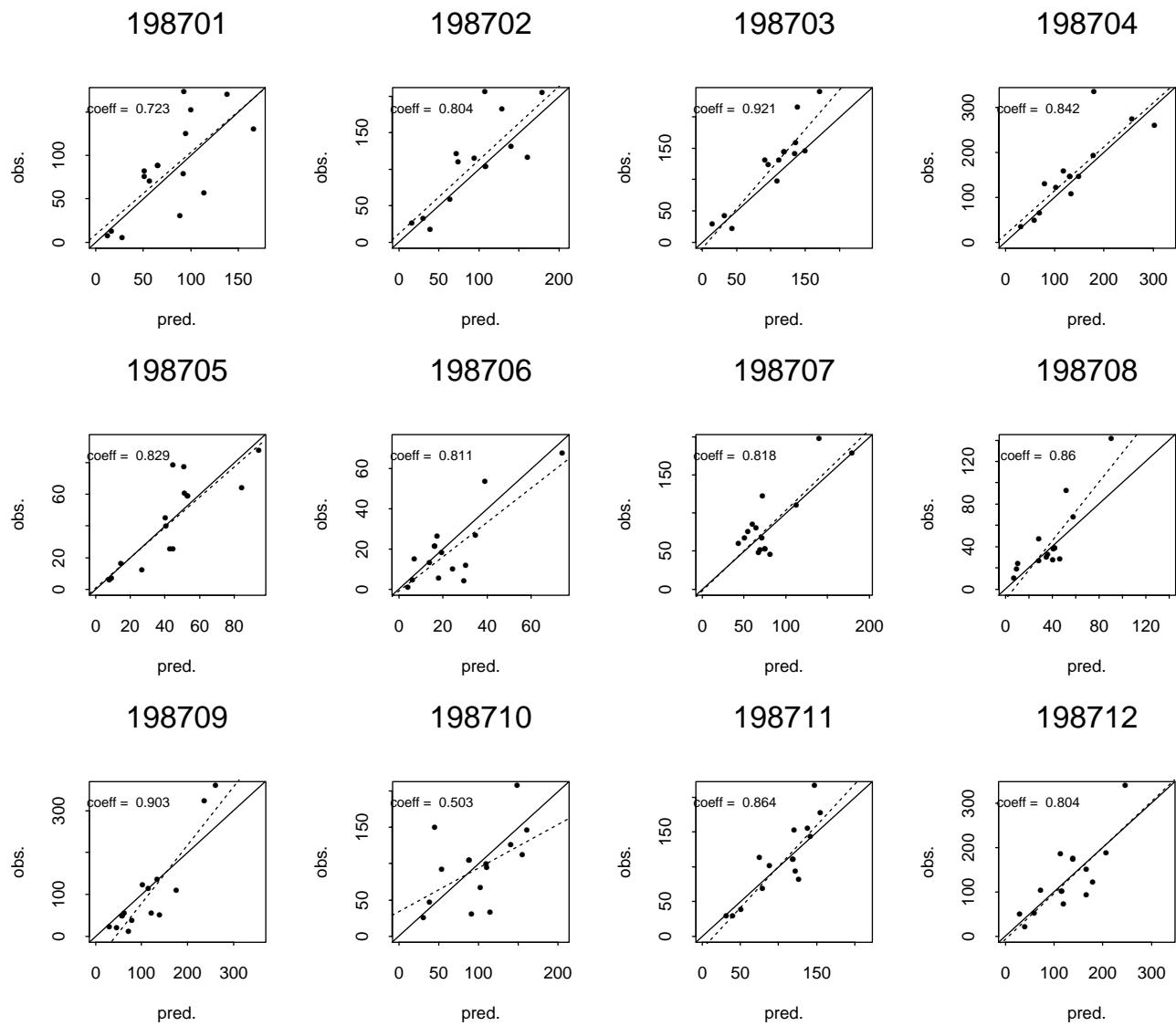


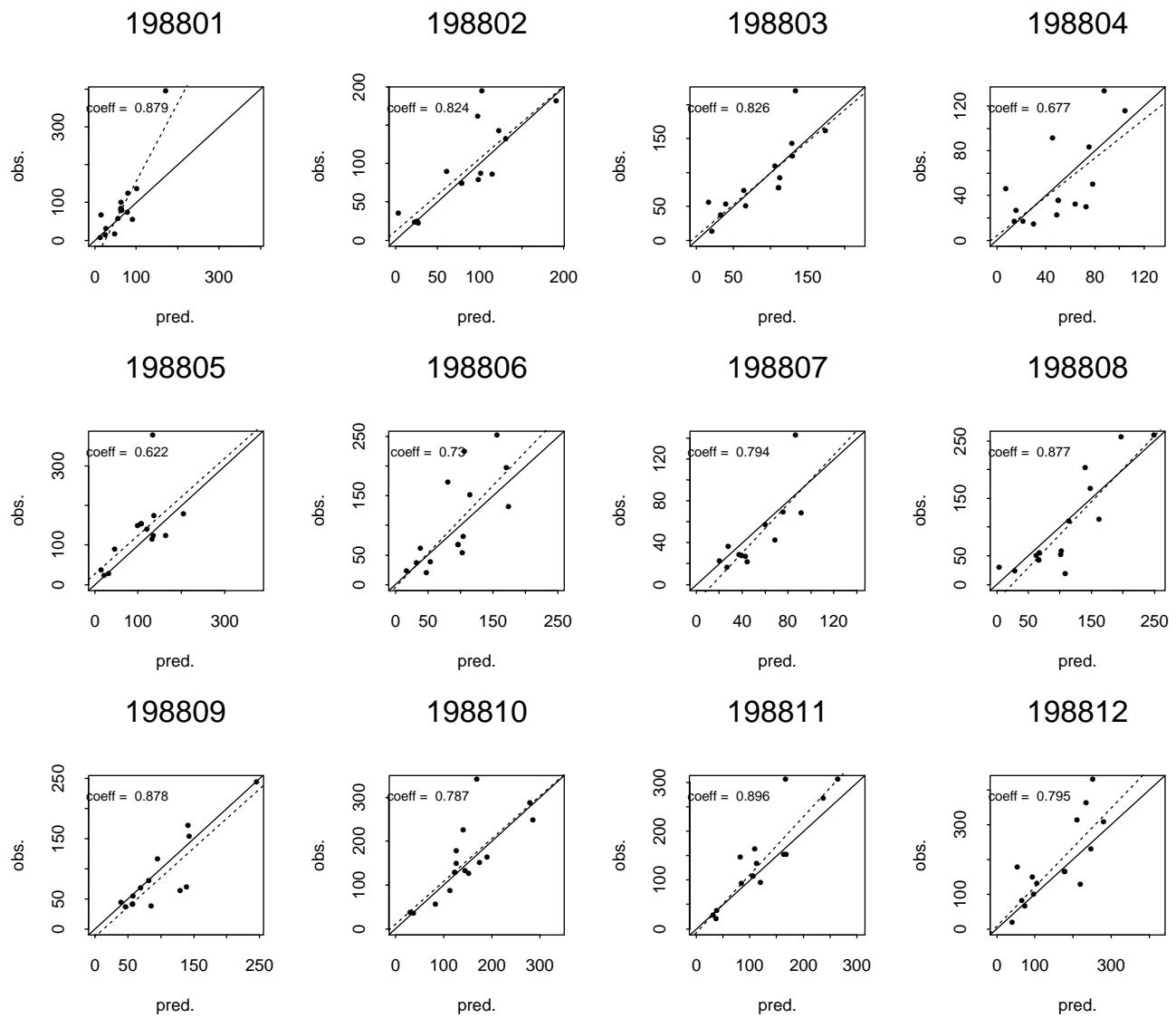


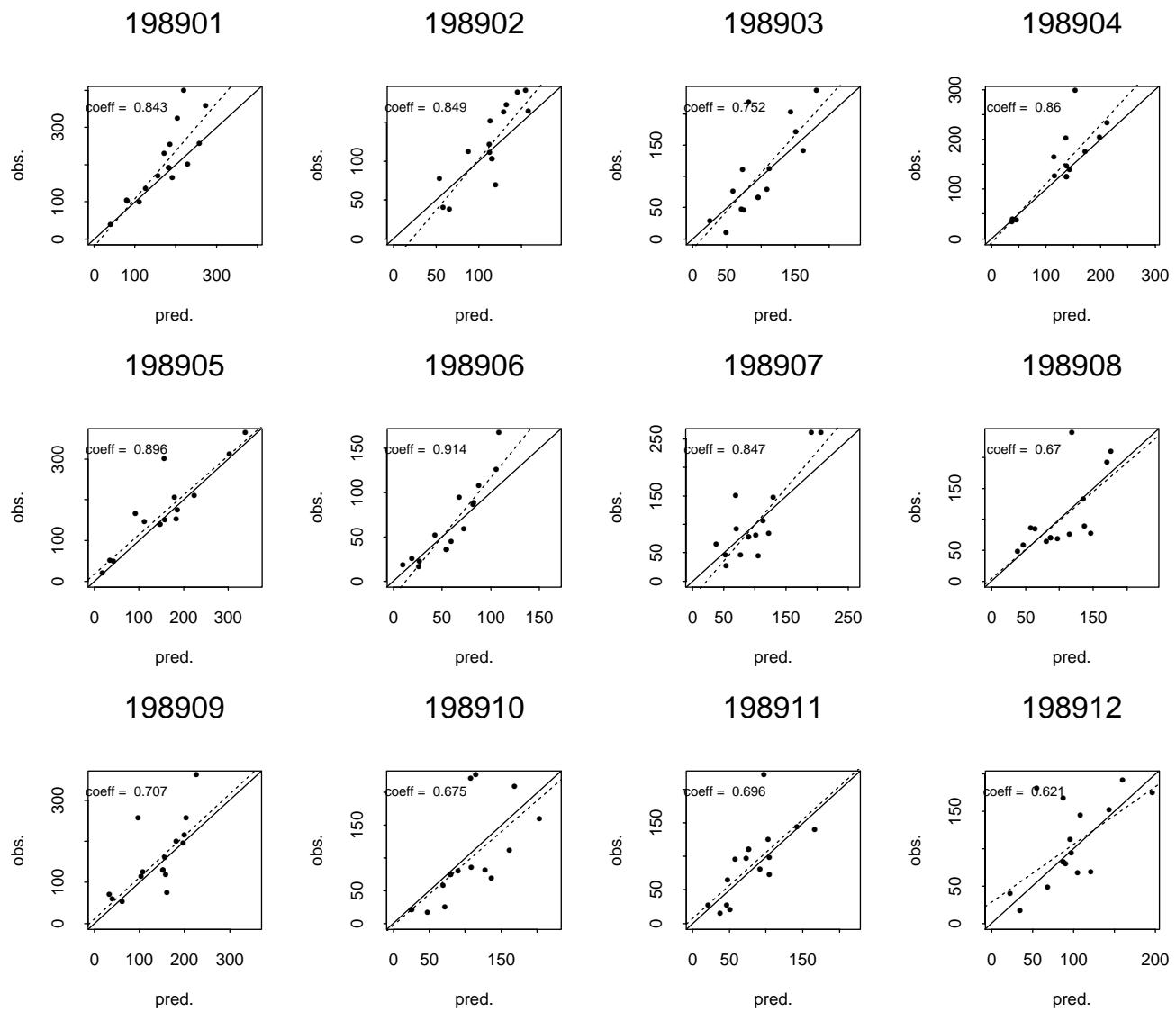


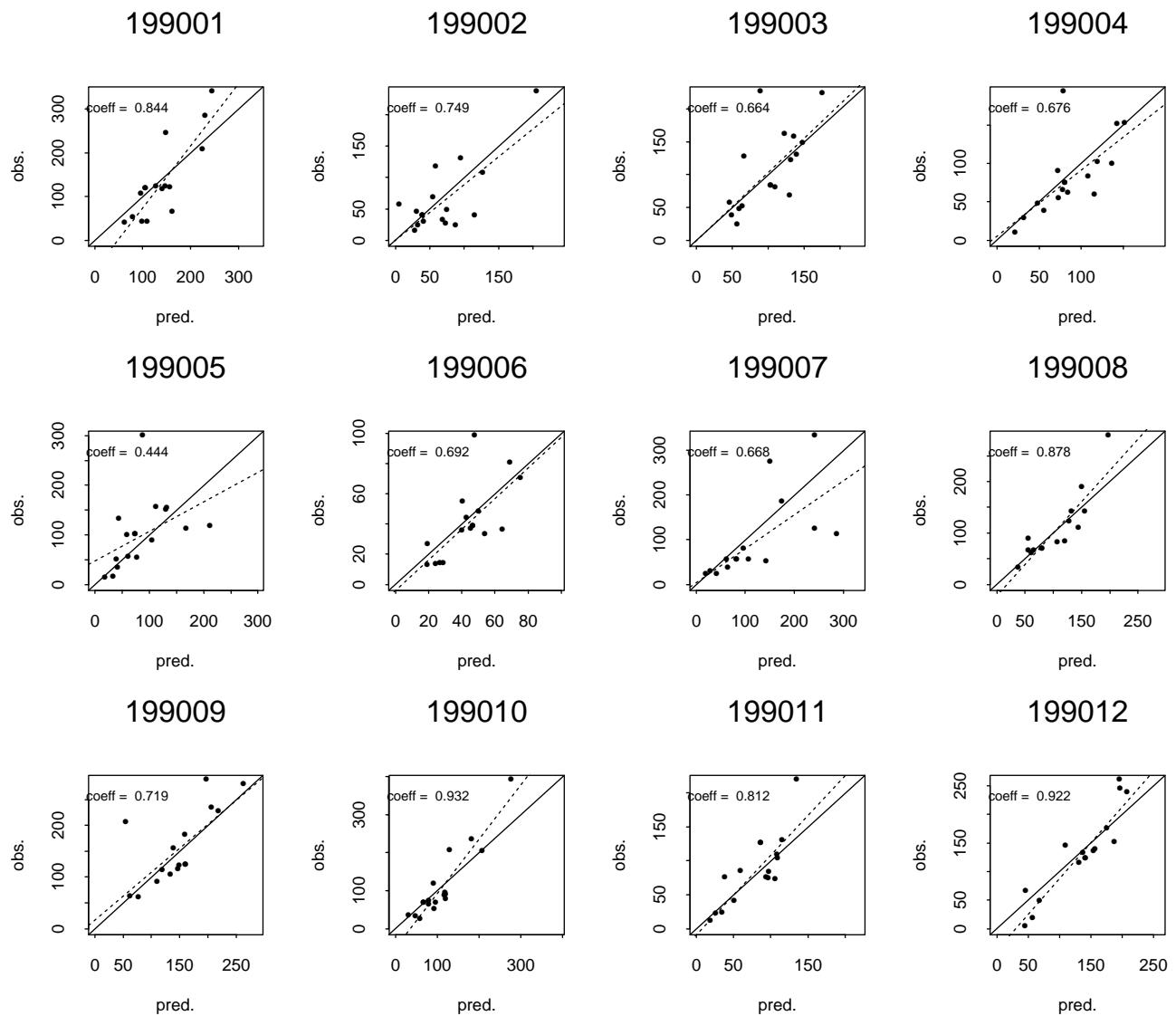


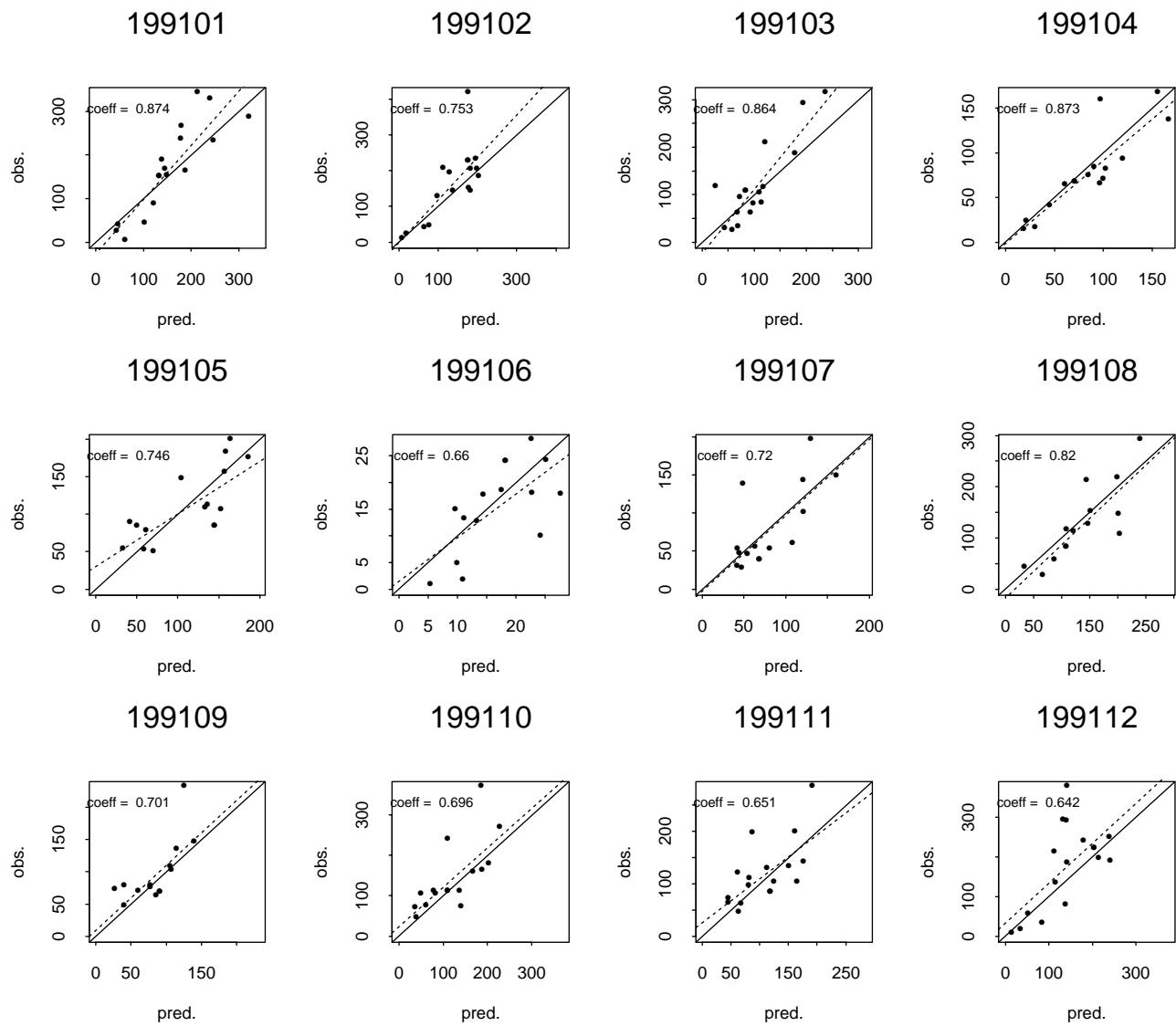


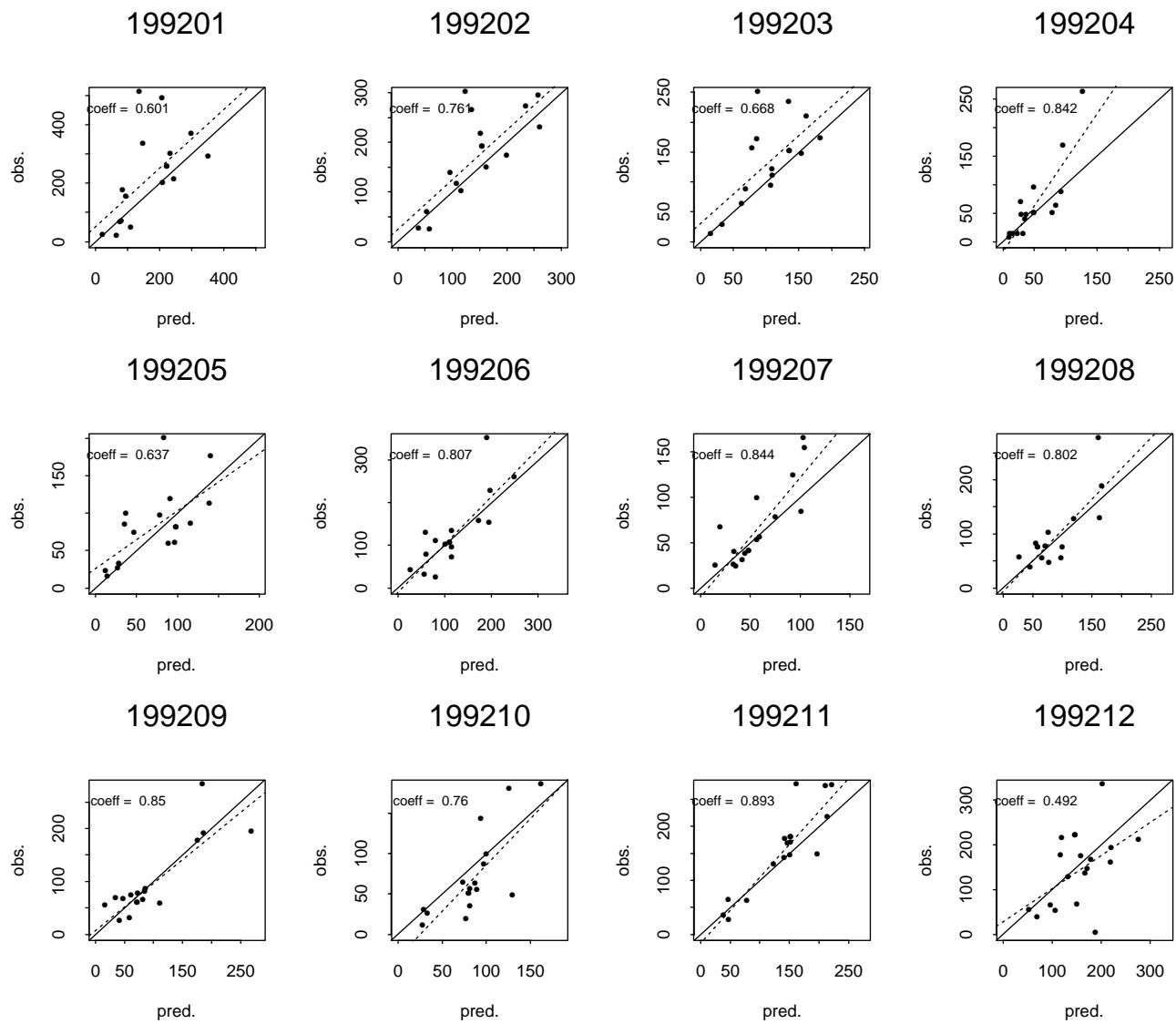


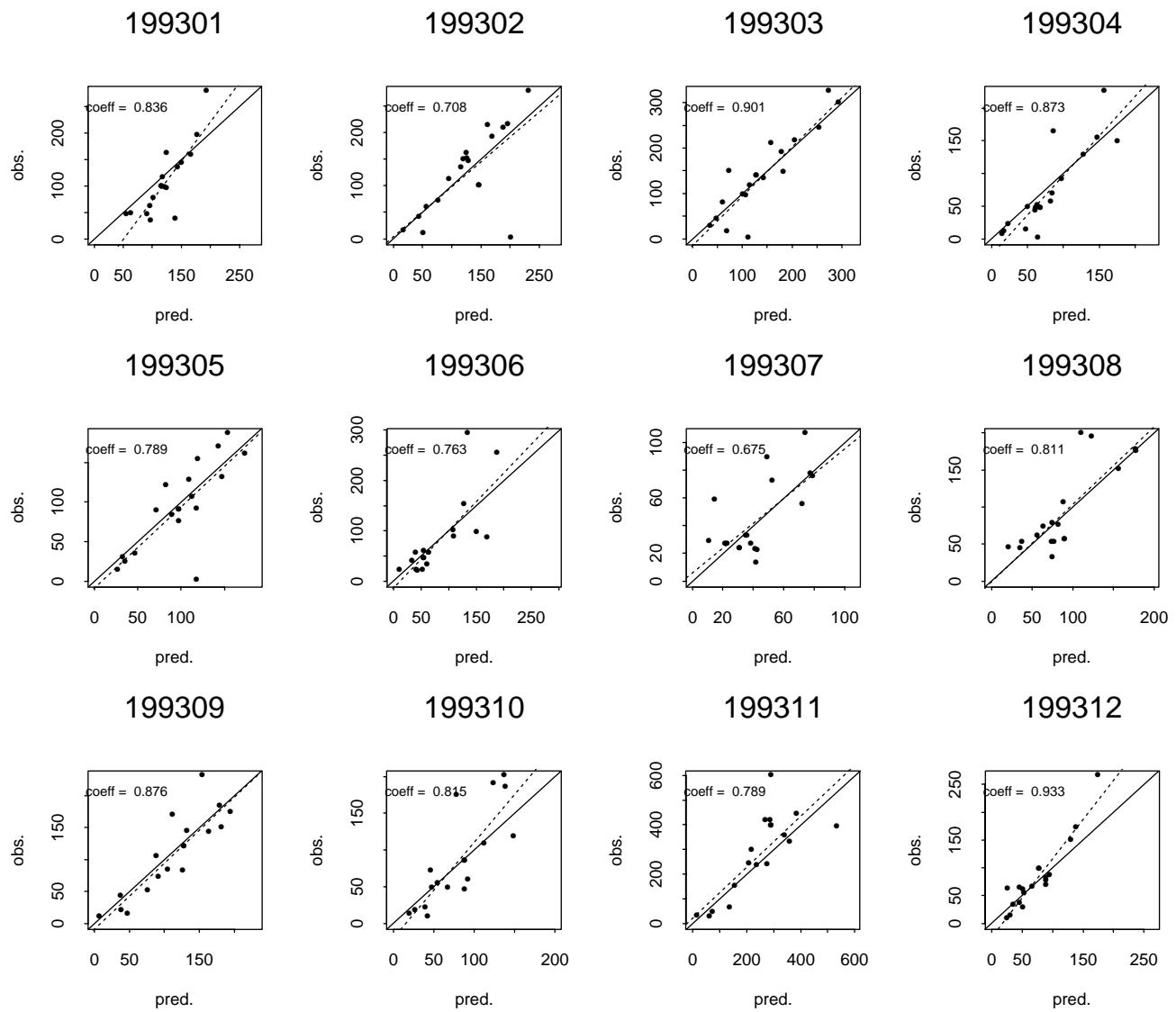


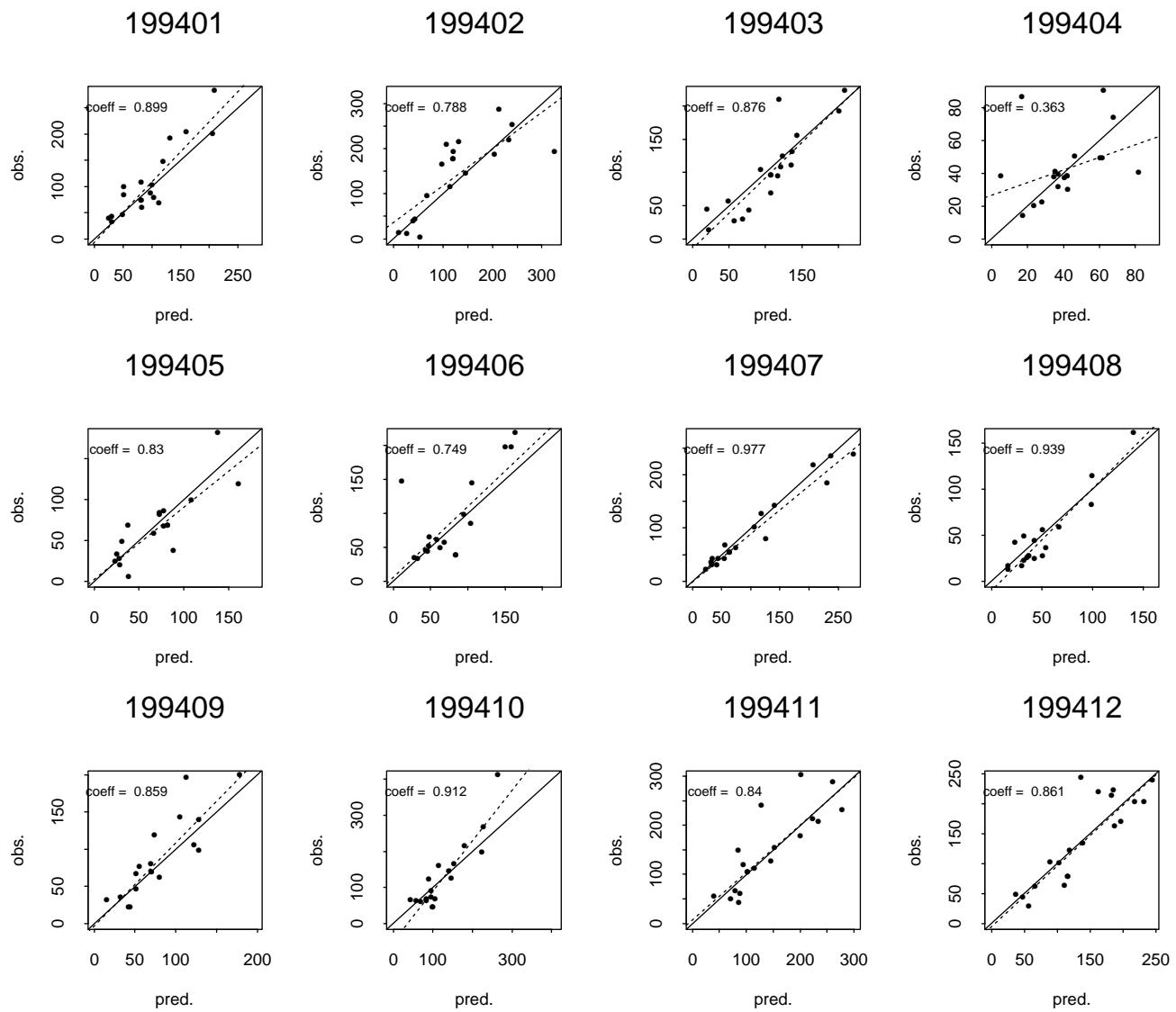


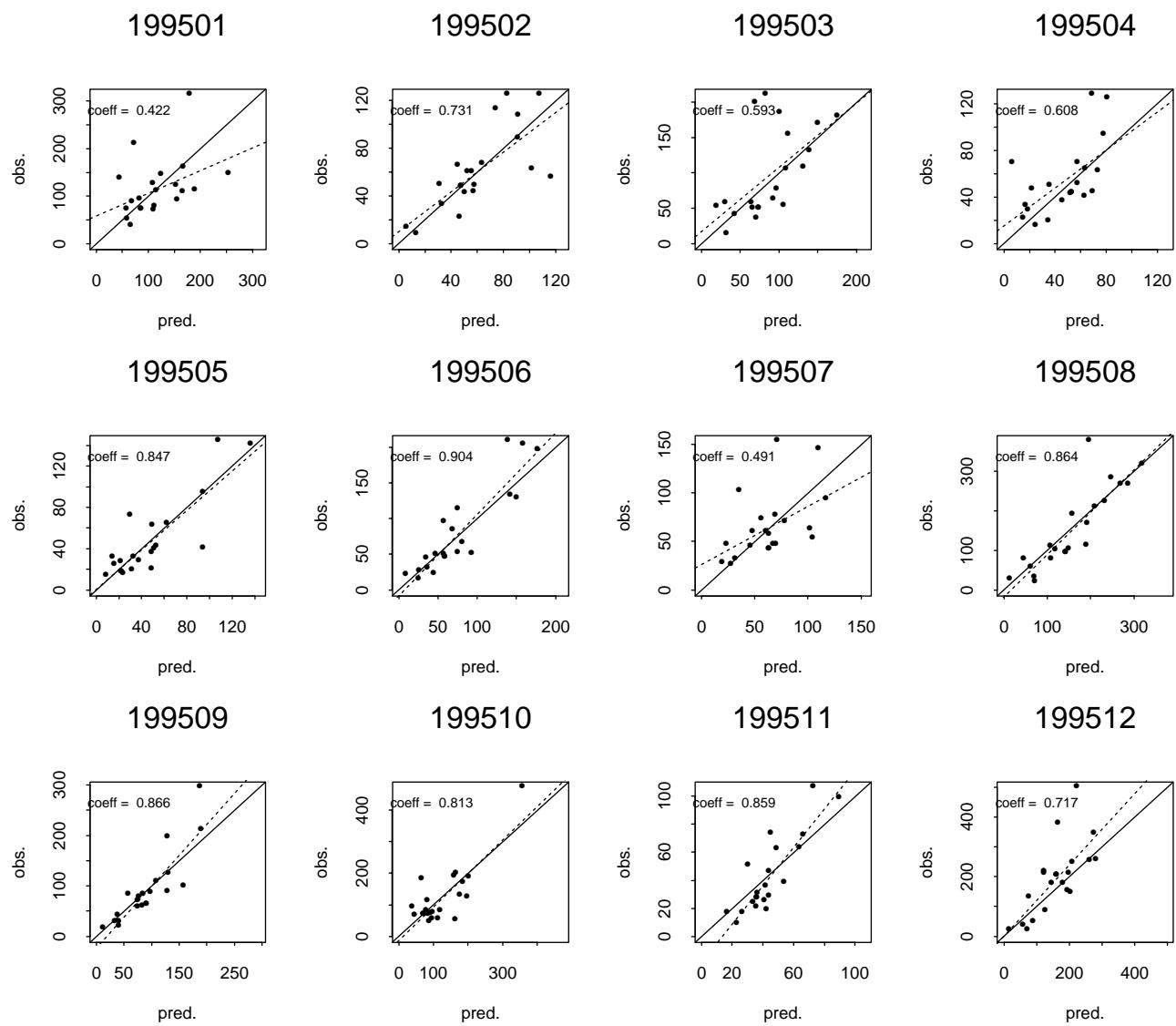


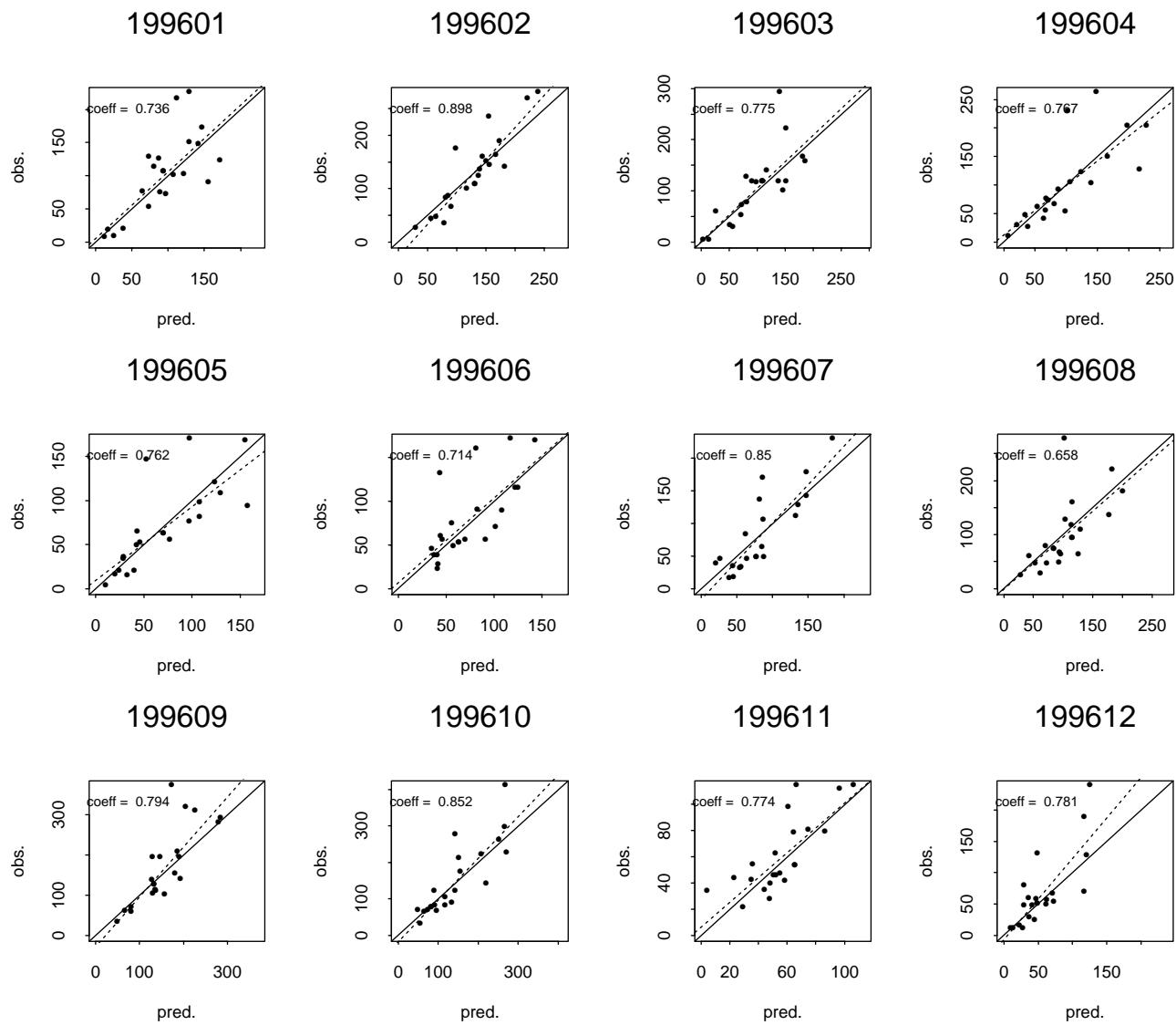


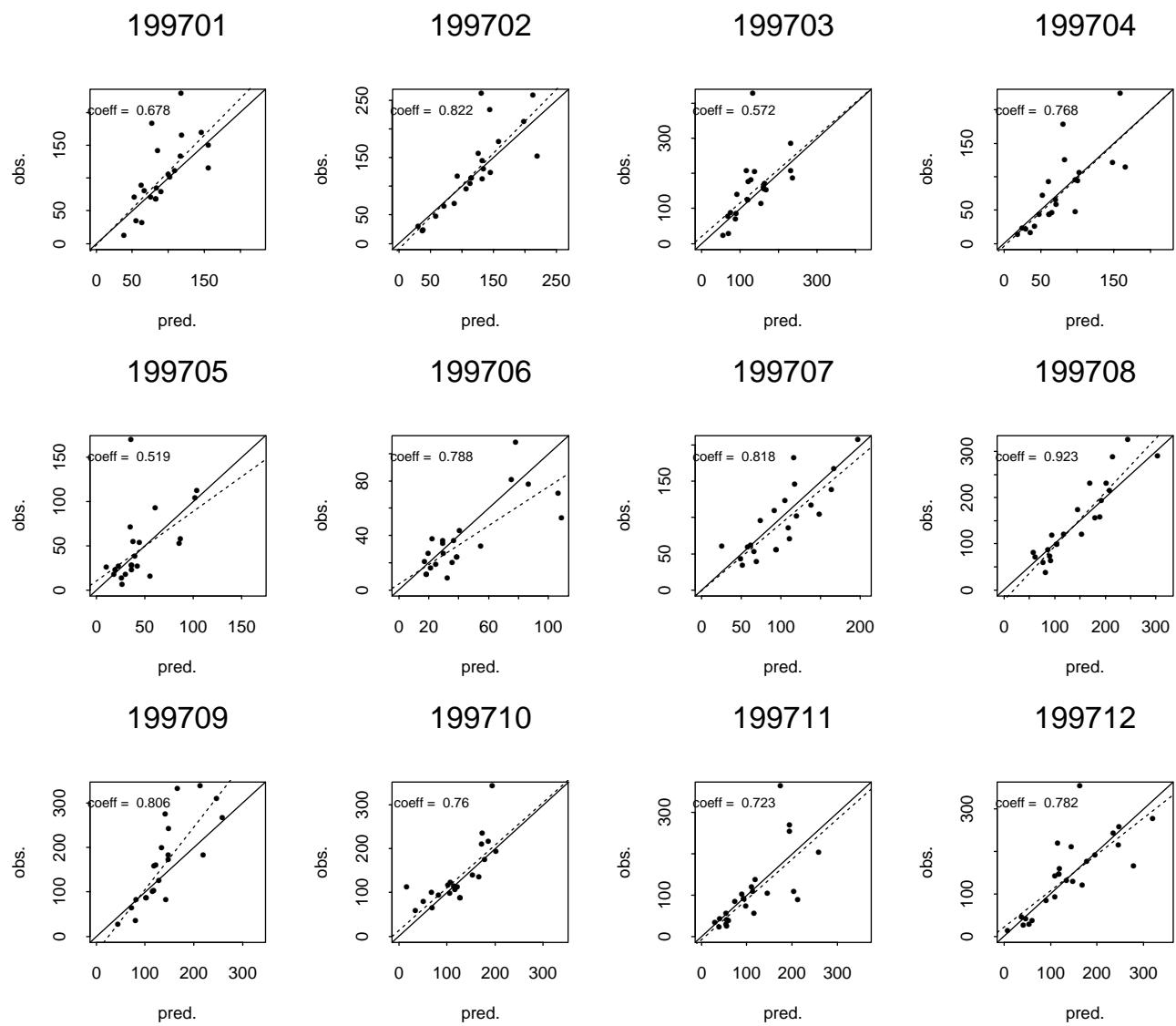


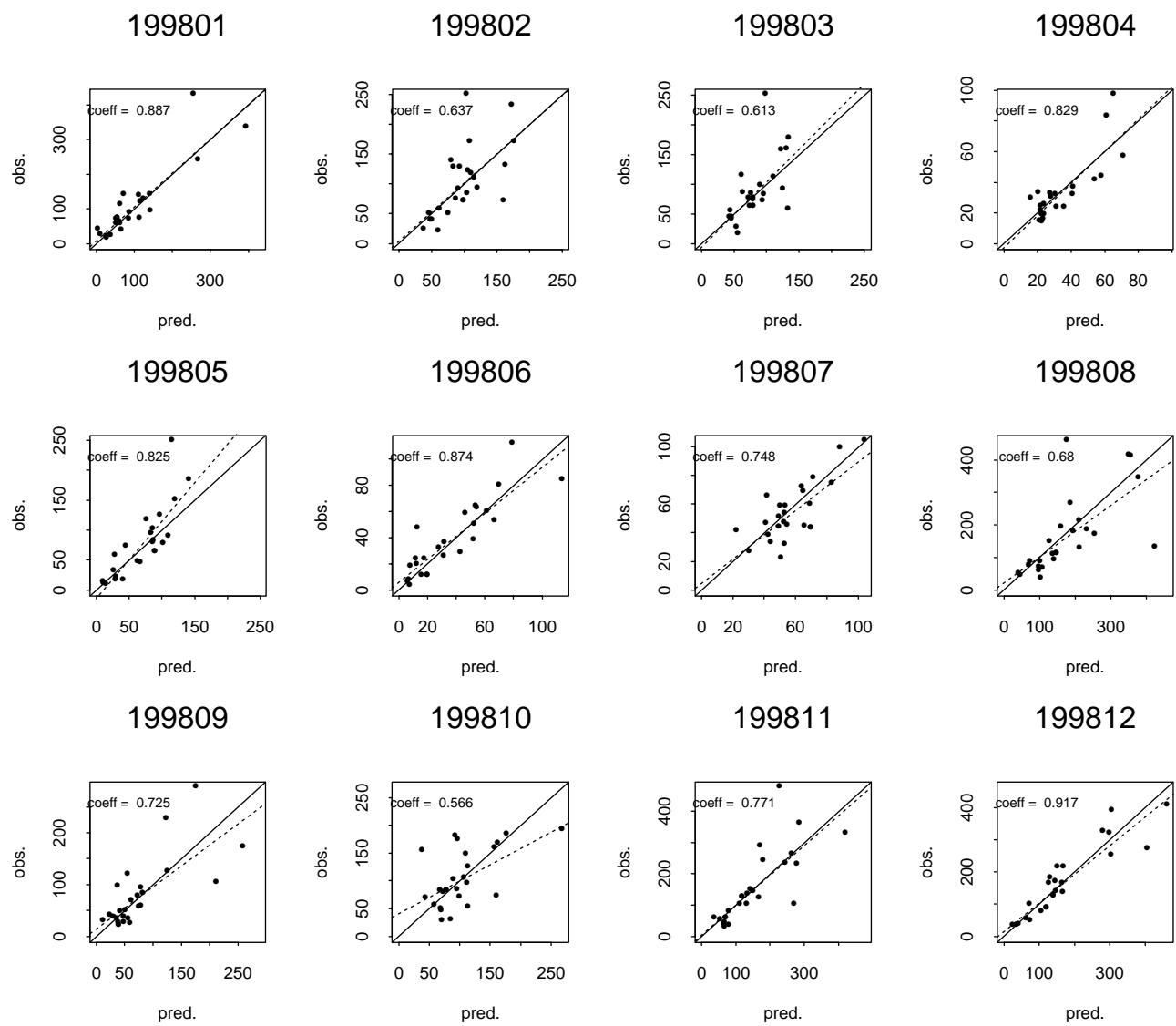


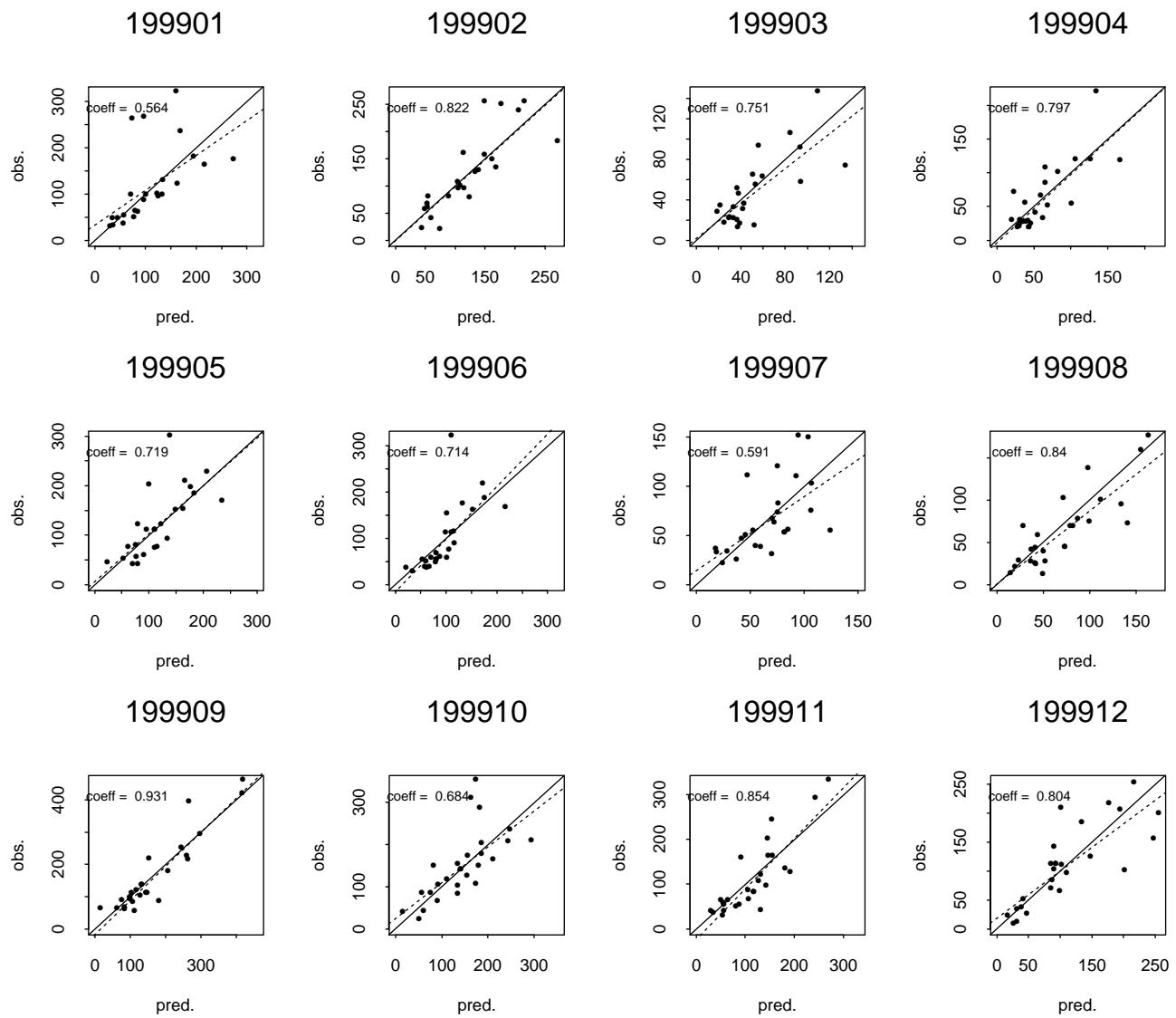


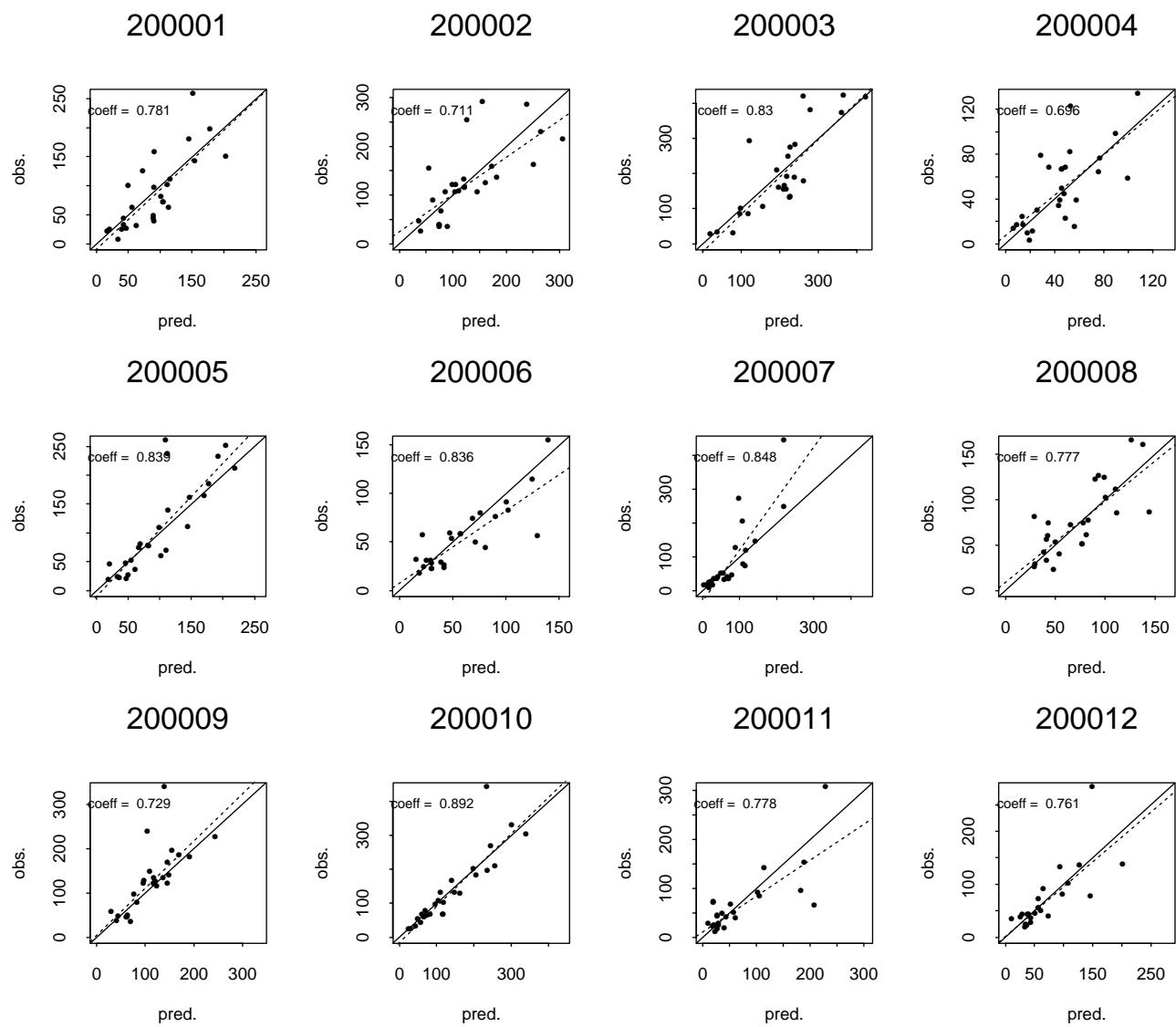








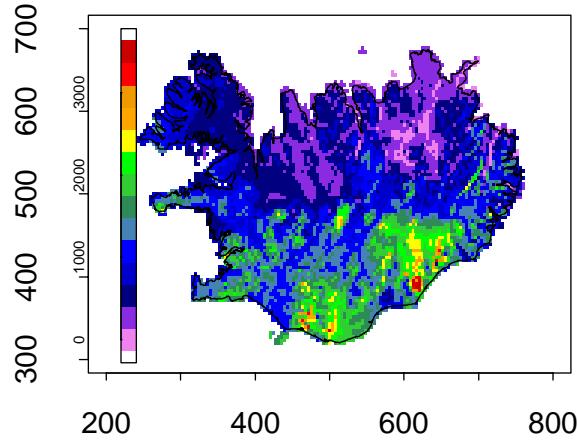




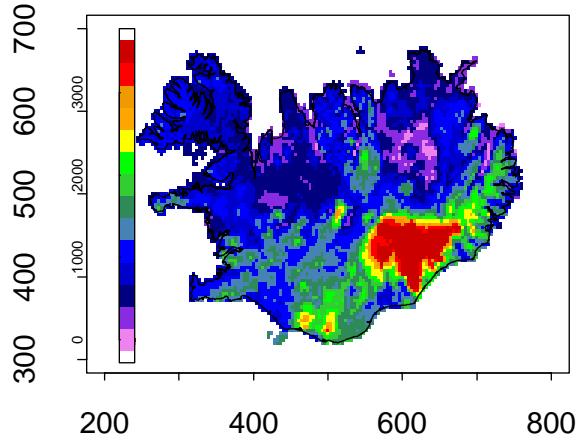
Appendix 7

Annual precipitation maps from 1980 to 2000
(4 km resolution)

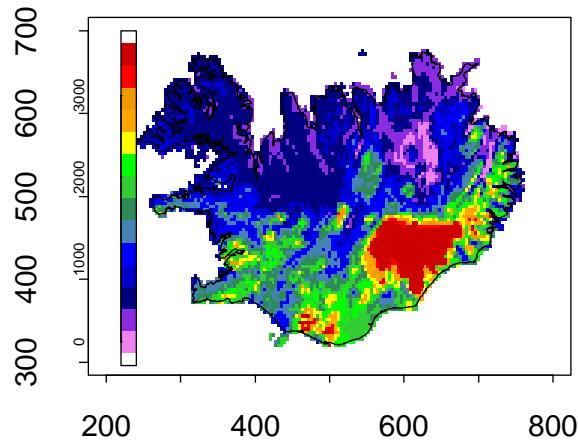
1980



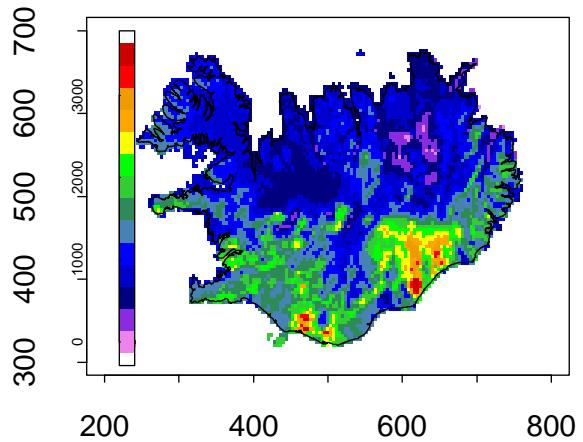
1981



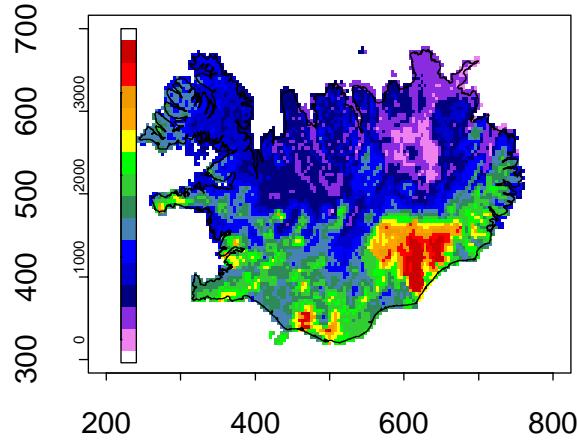
1982



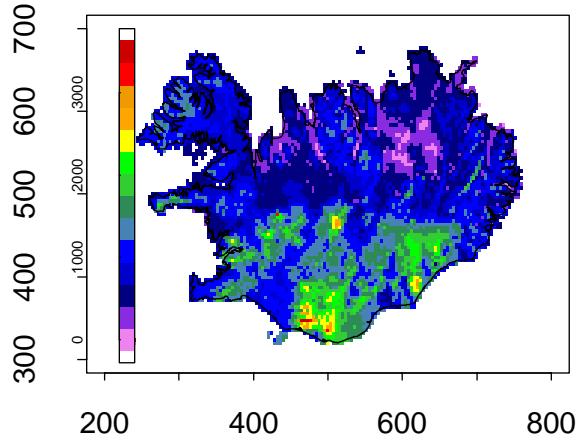
1983



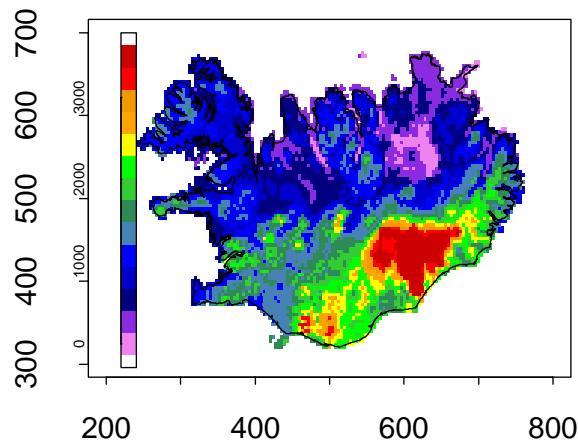
1984



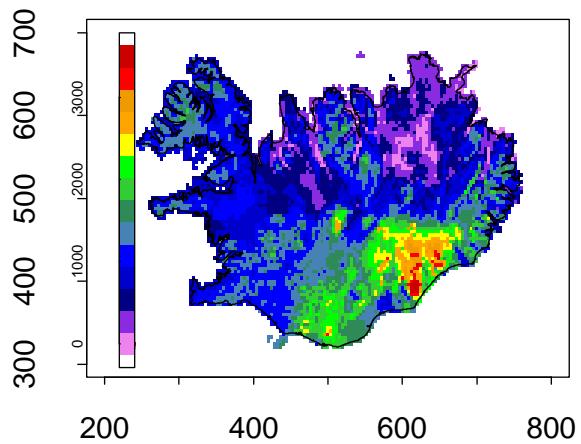
1985



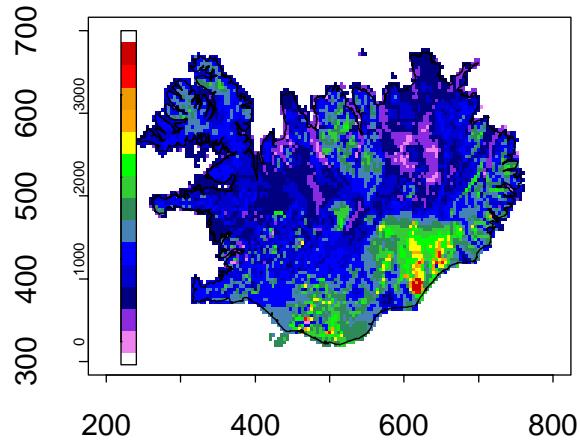
1986



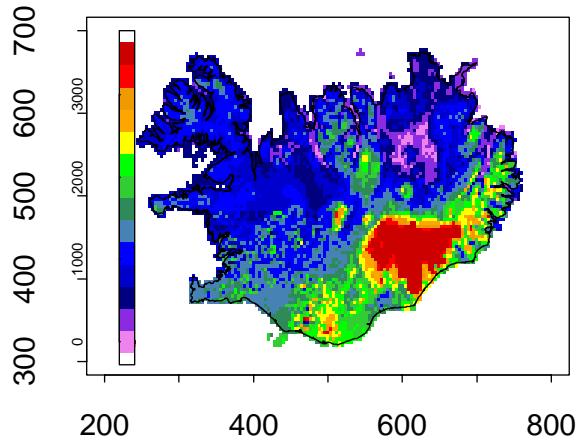
1987



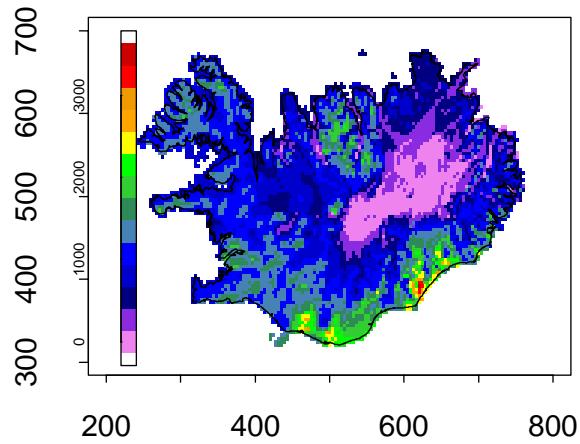
1988



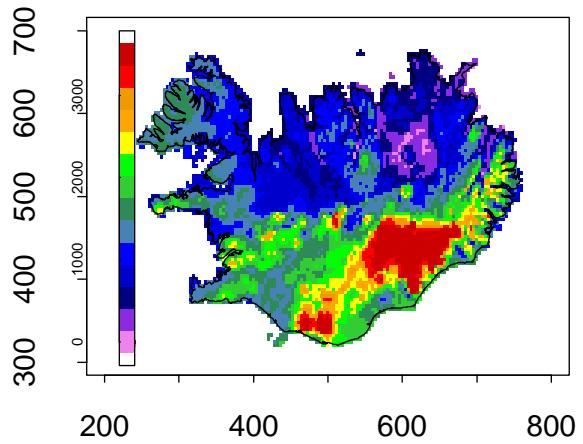
1989



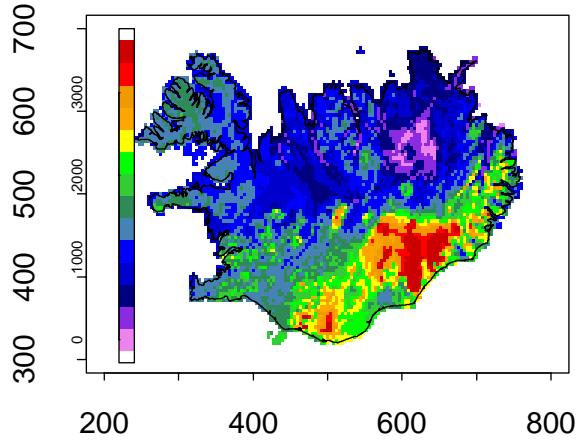
1990



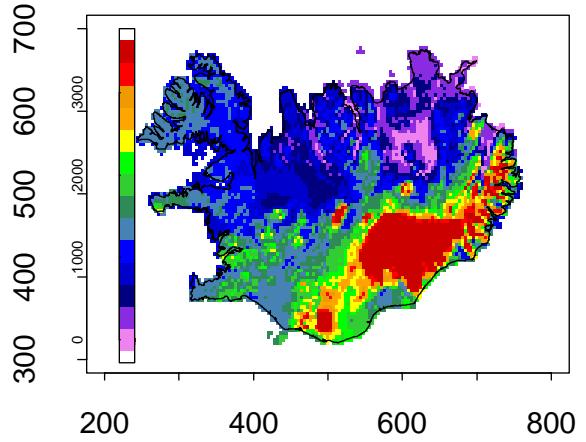
1991



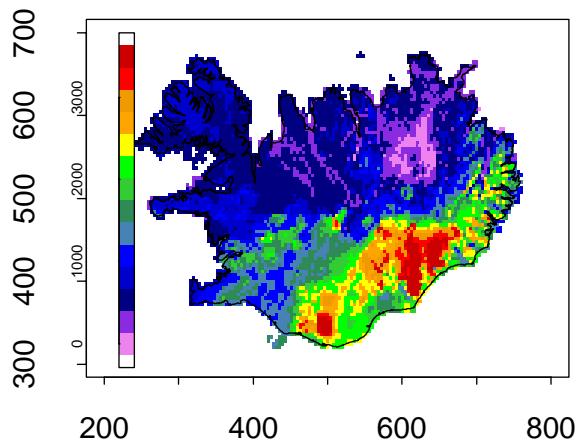
1992



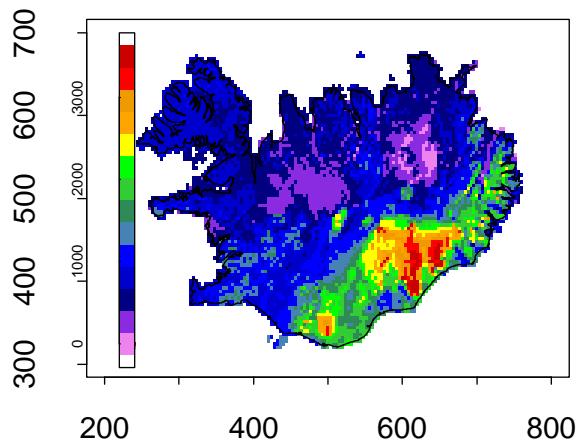
1993

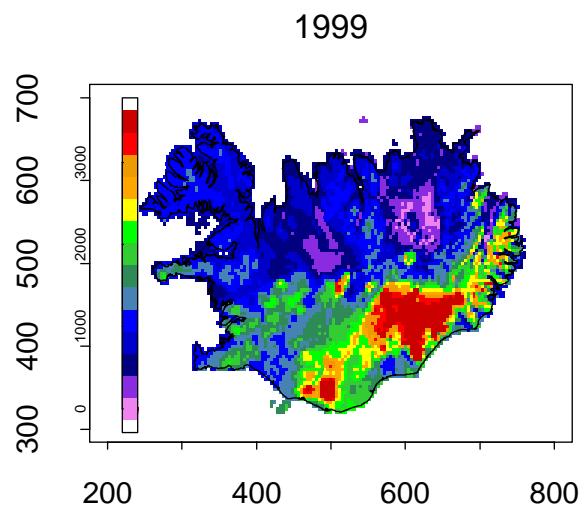
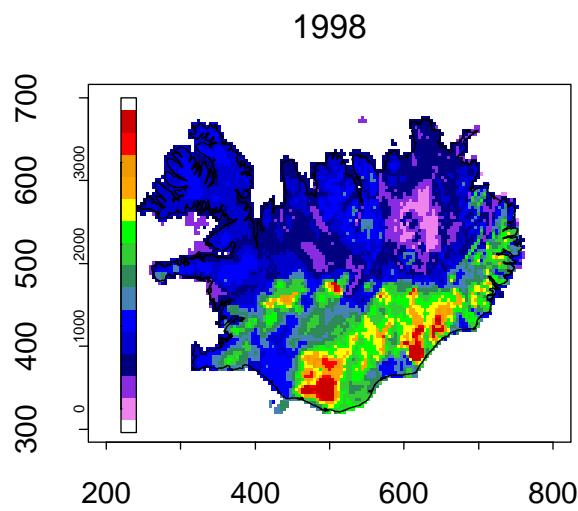
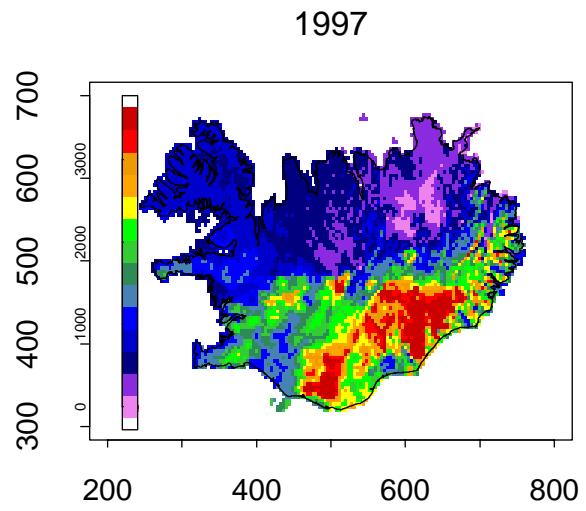
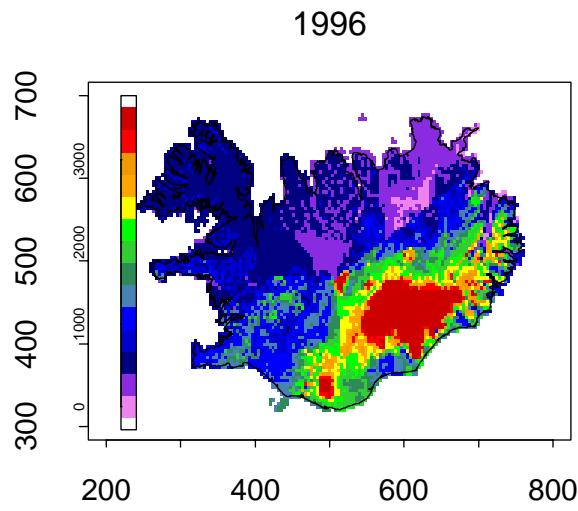


1994

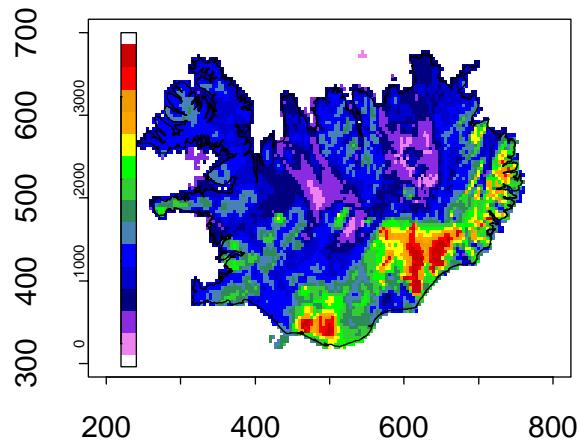


1995





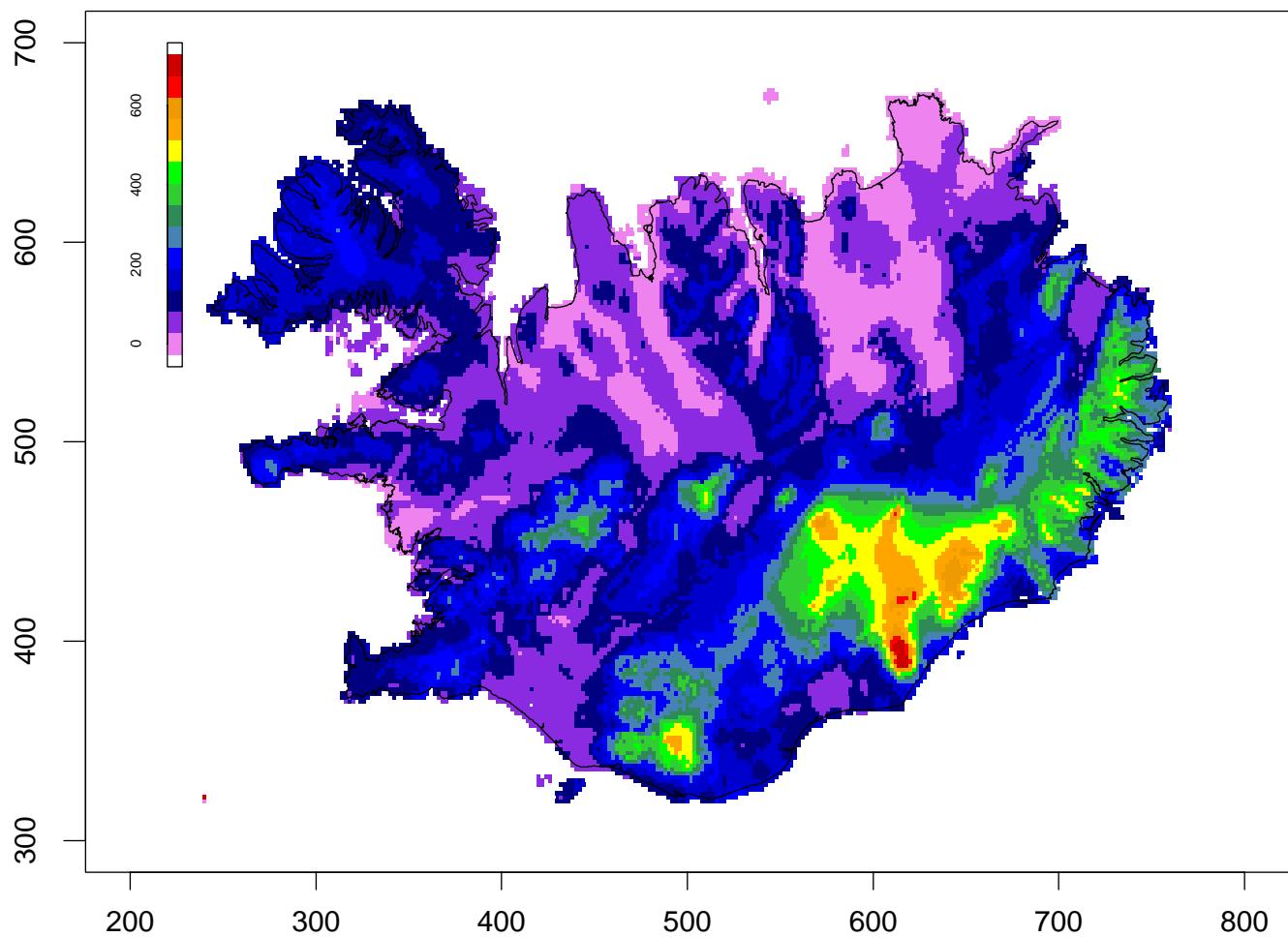
2000



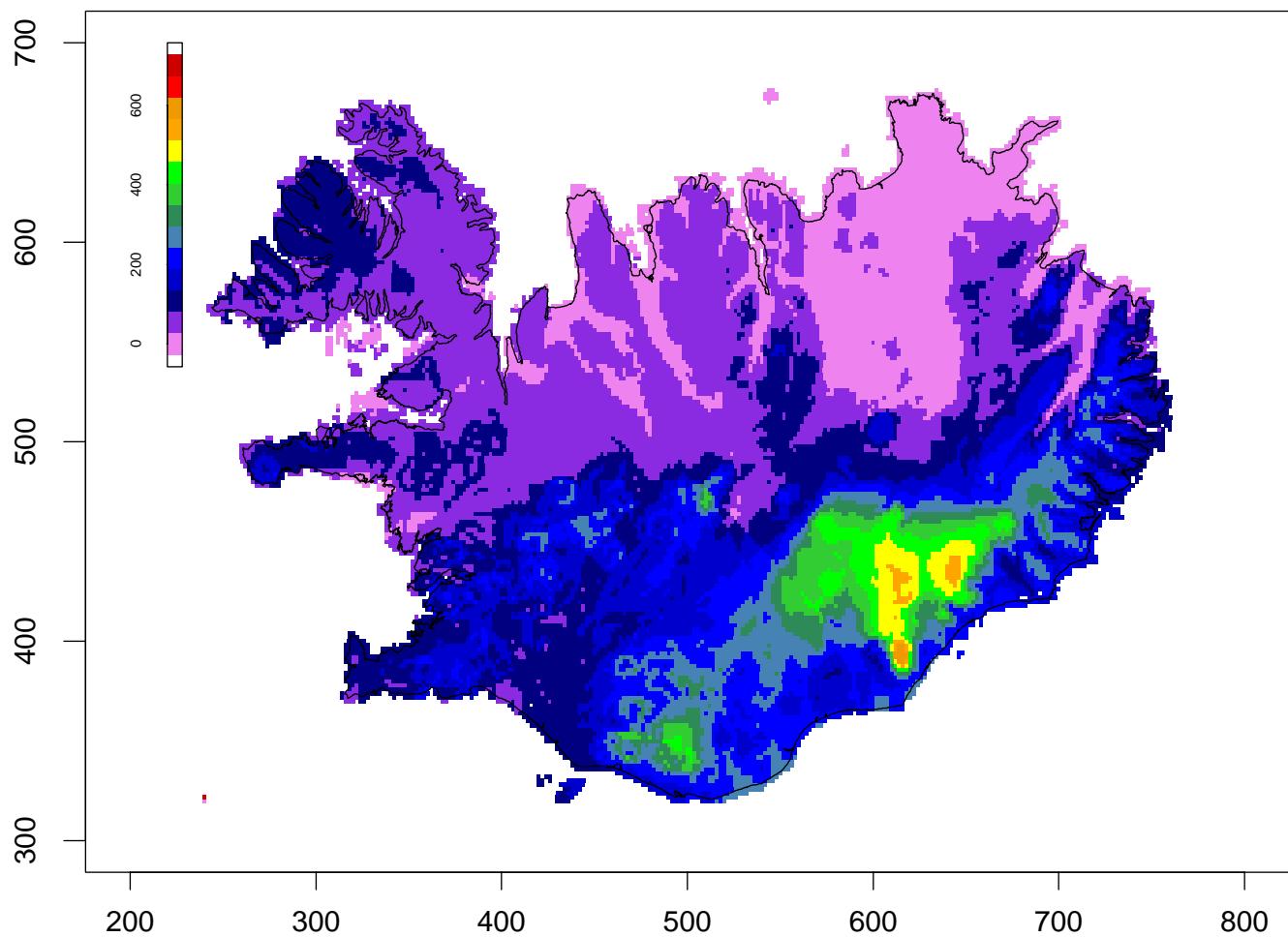
Appendix 8

Monthly precipitation maps in 2001
(2 km resolution)

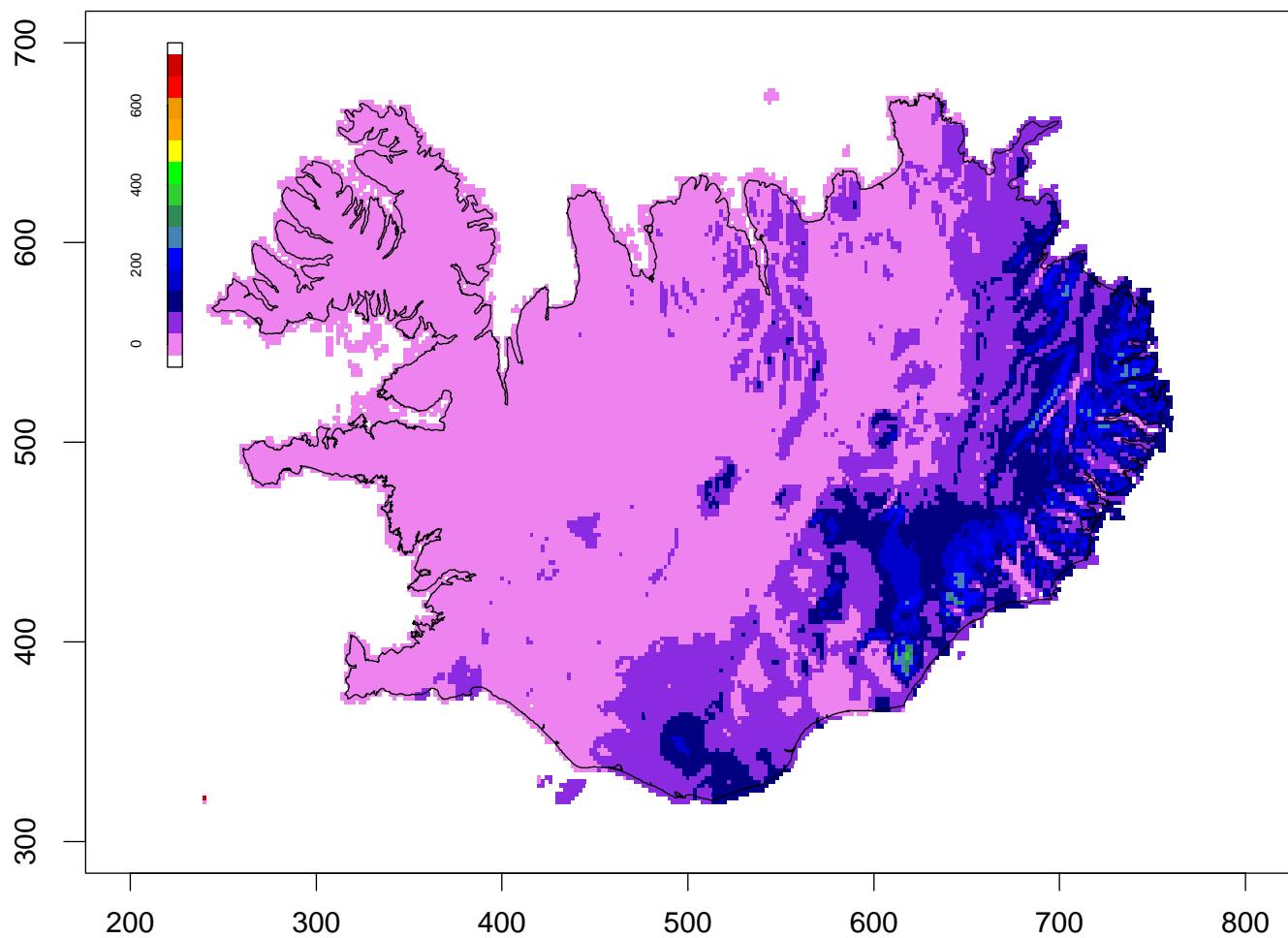
200101



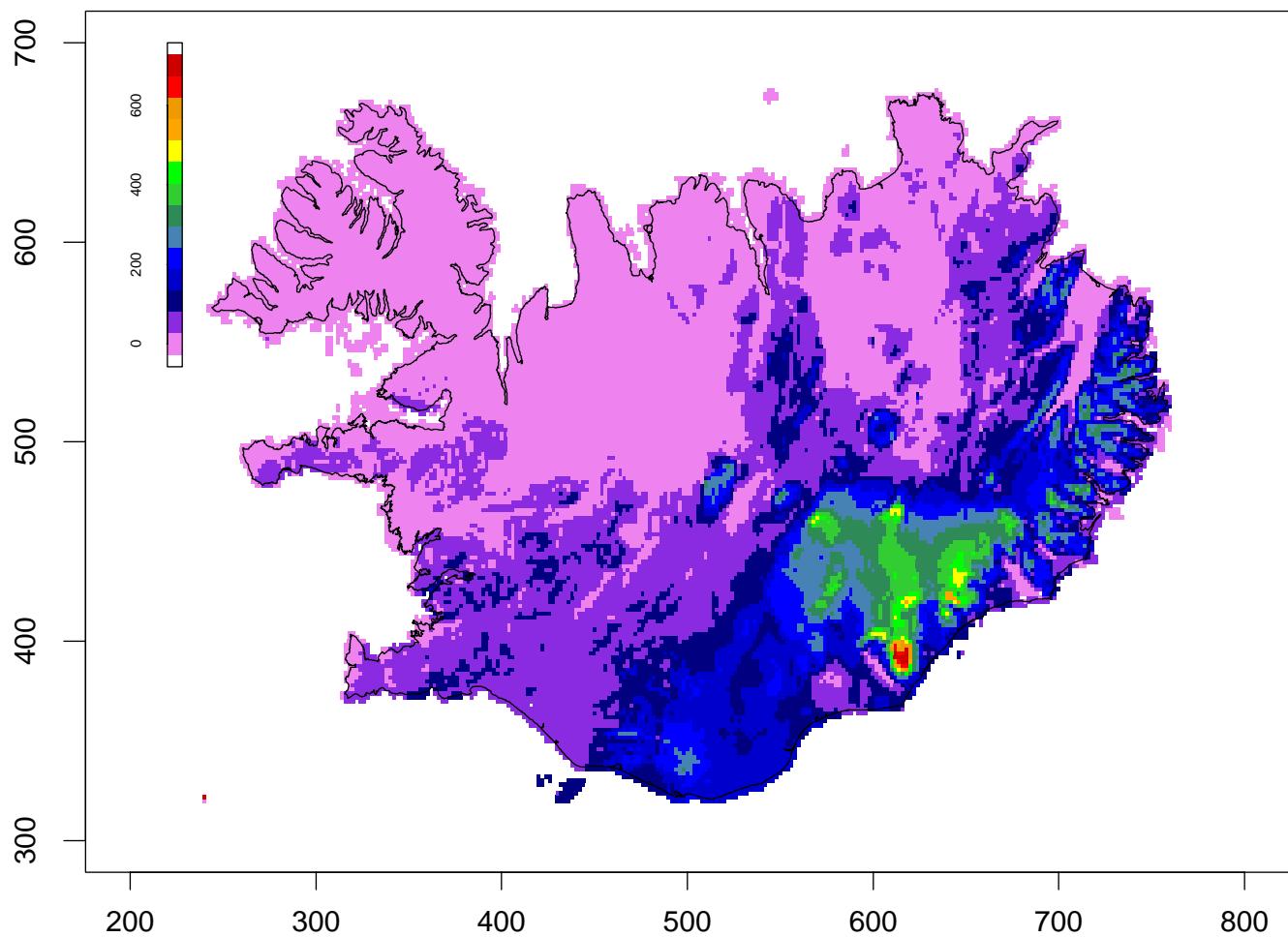
200102



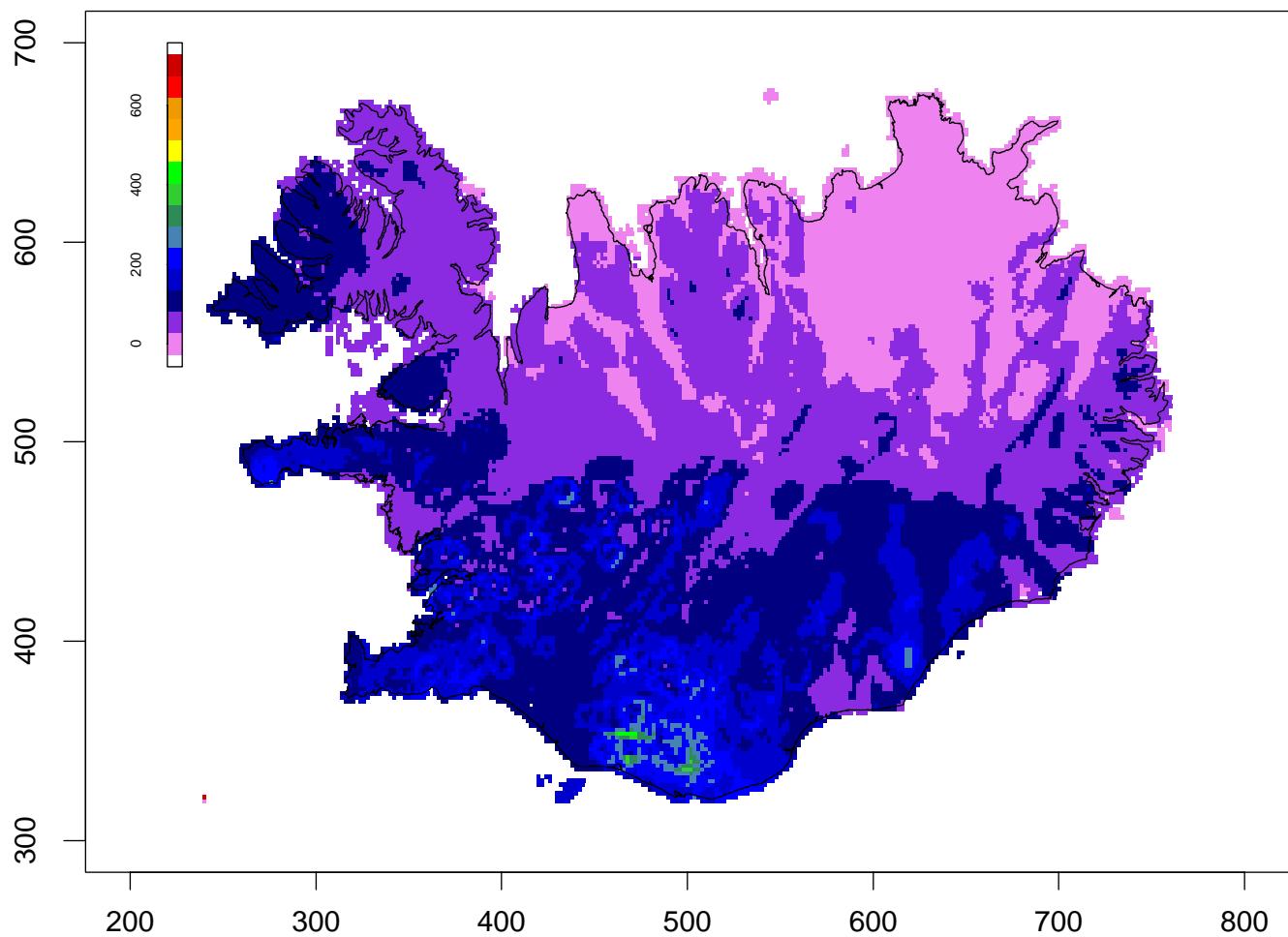
200103



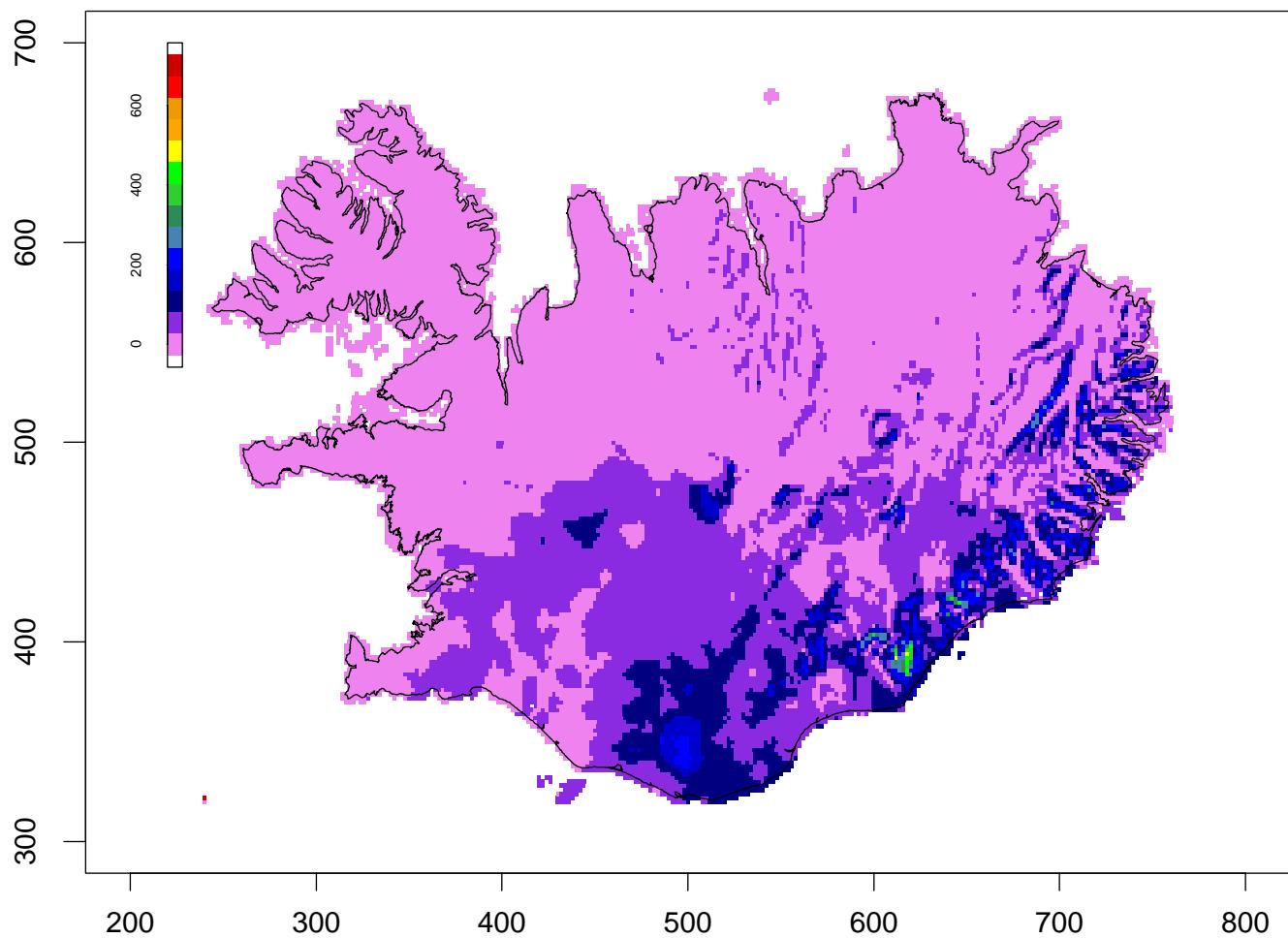
200104



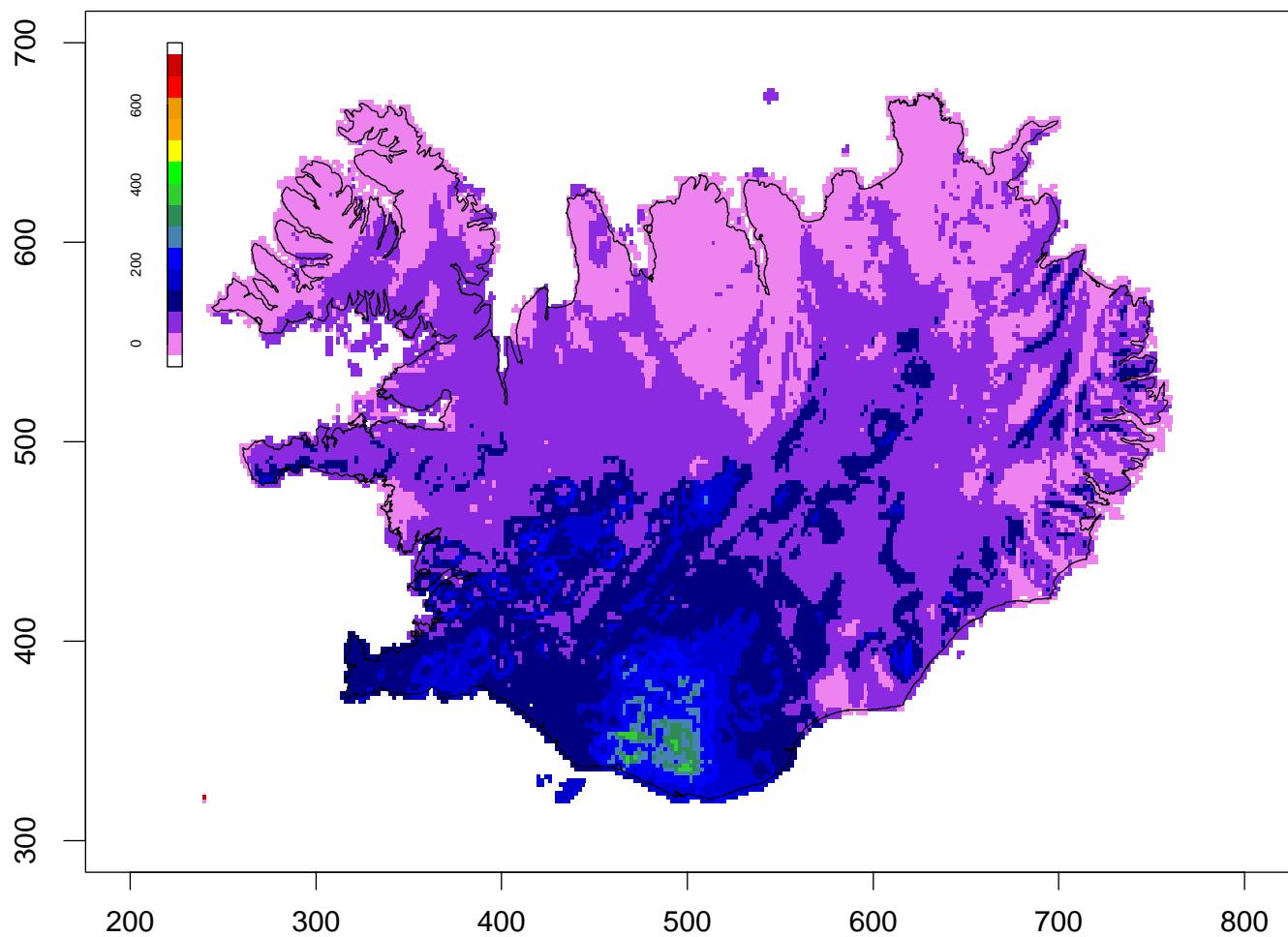
200105



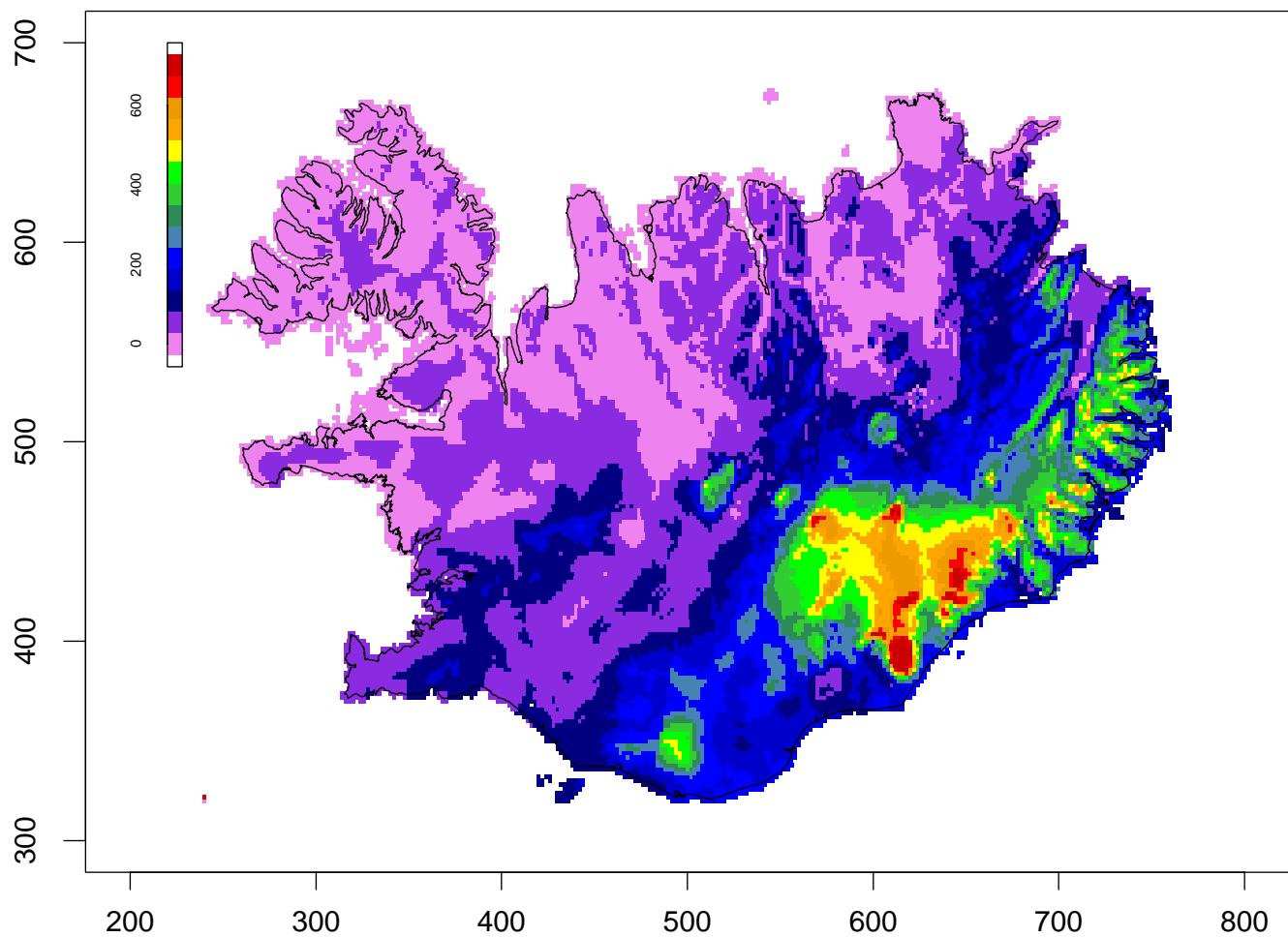
200106



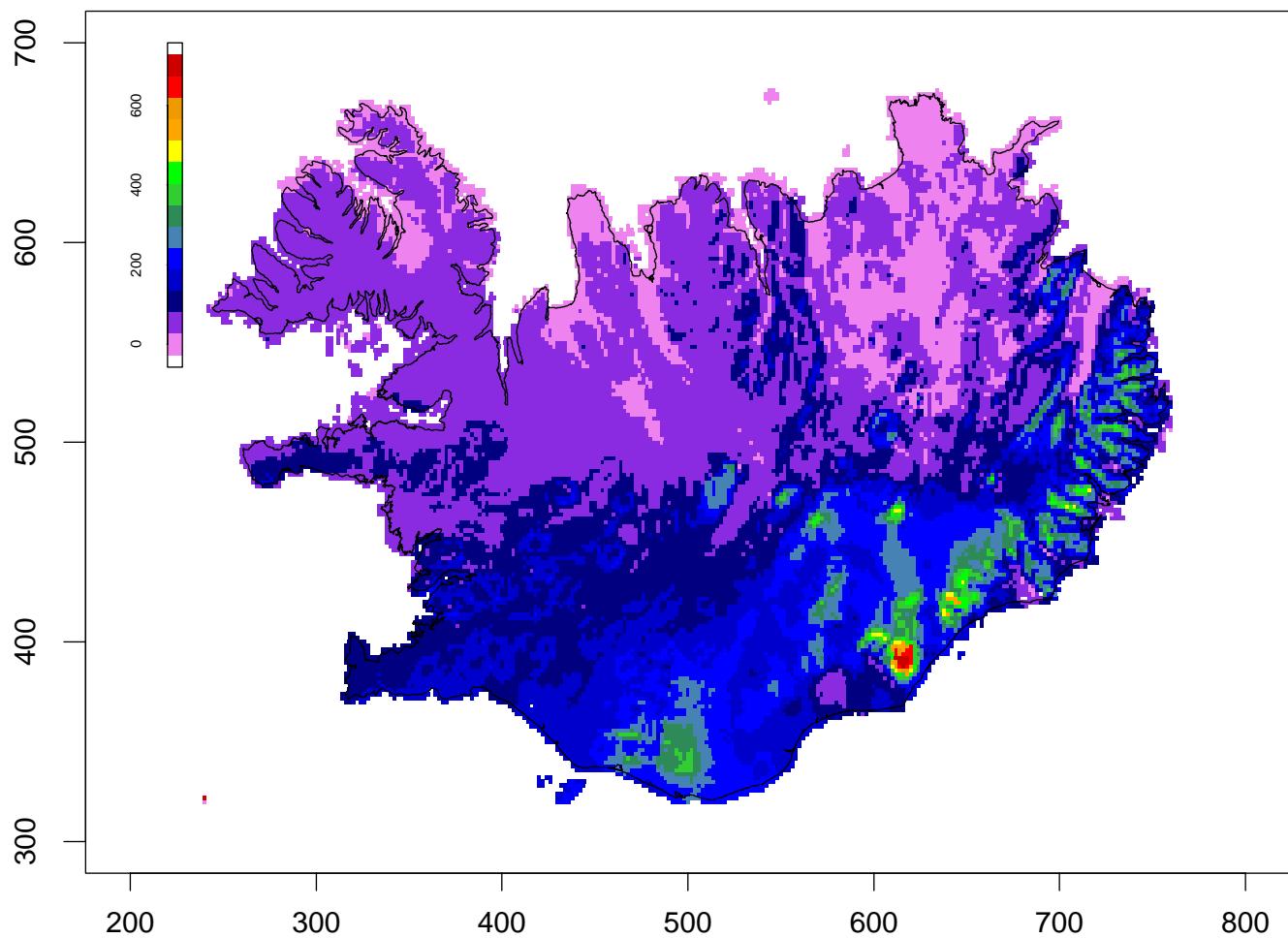
200107



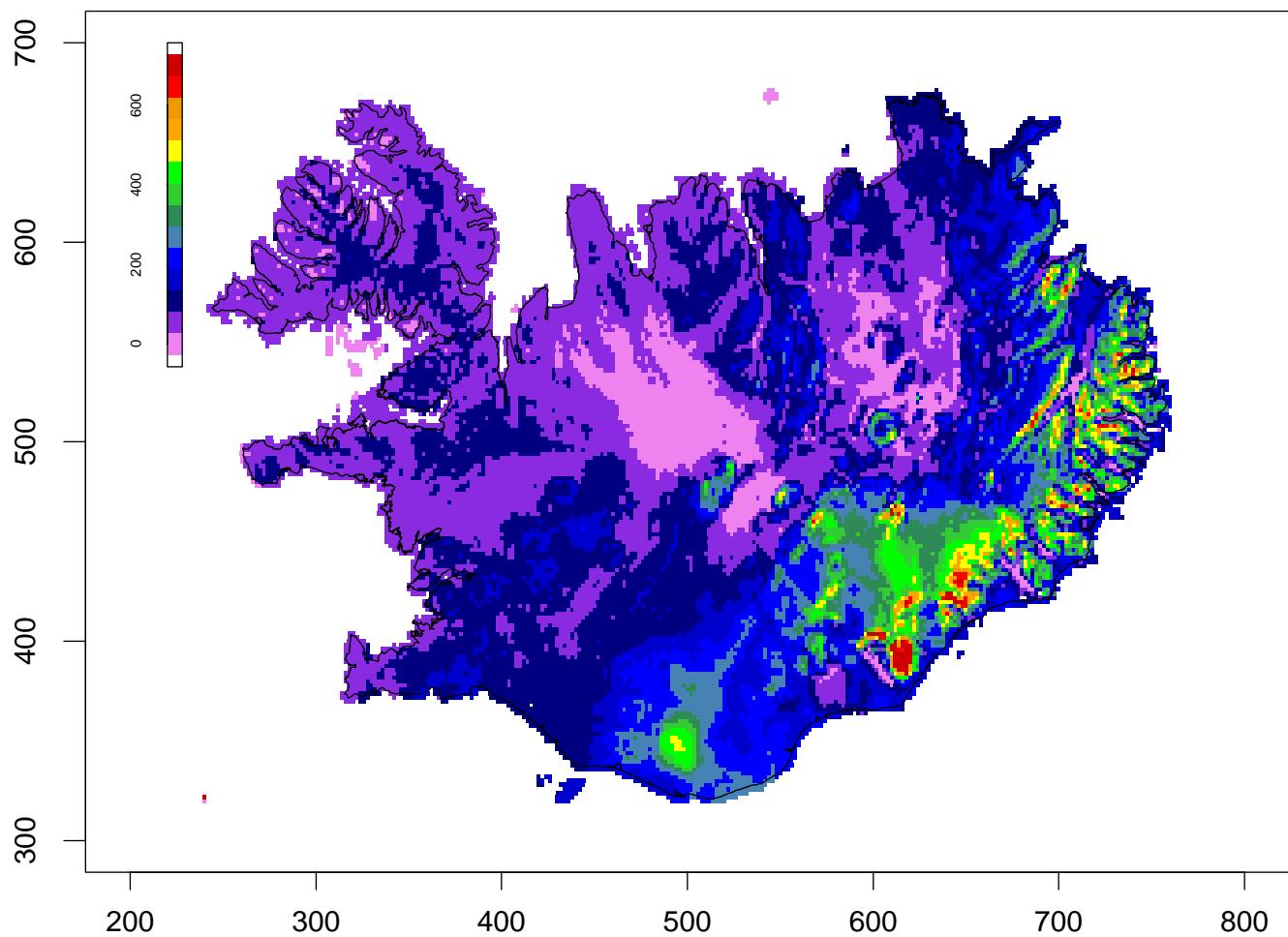
200108



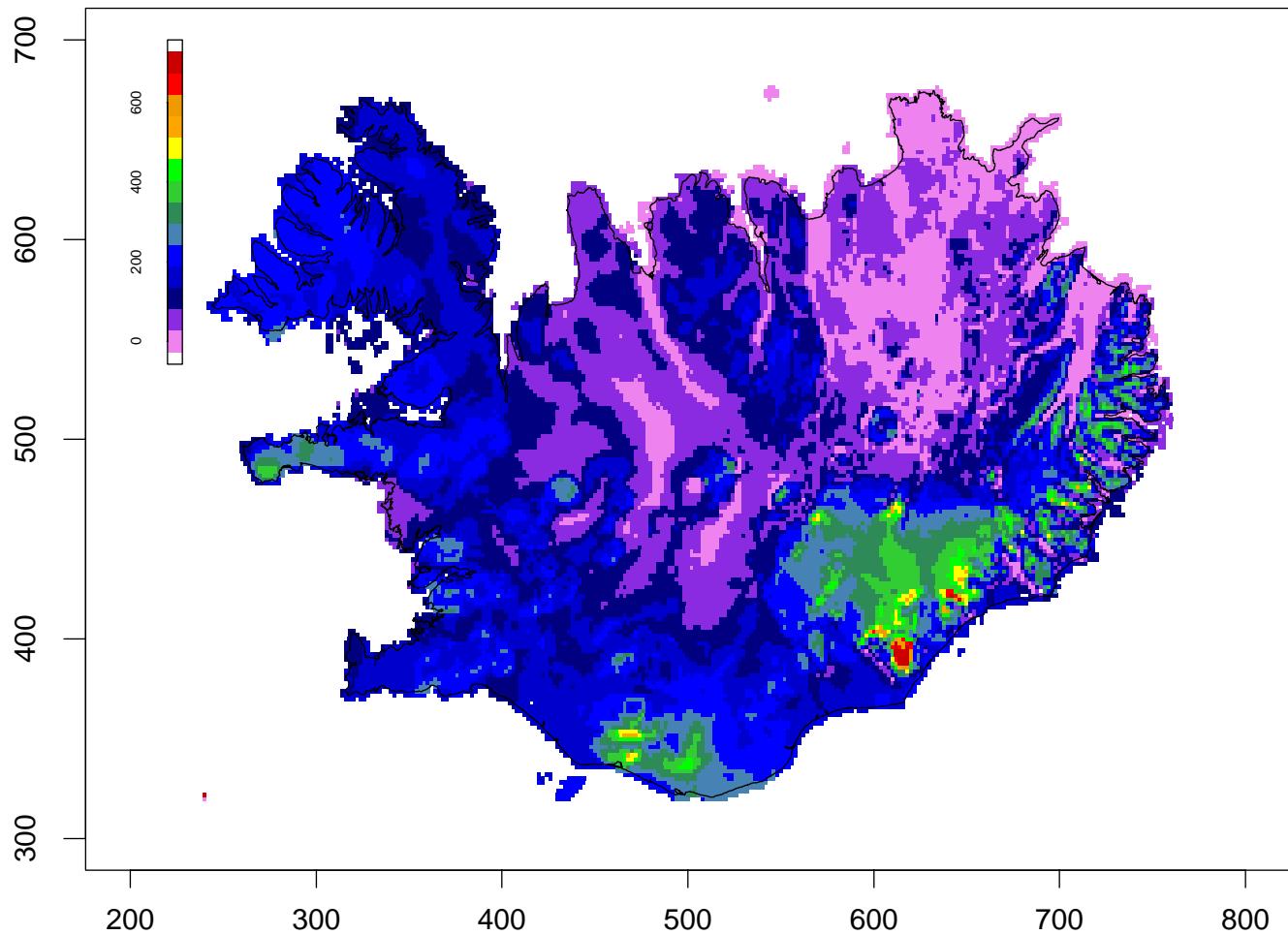
200109



200110



200111



200112

