Historical modelling of changes in Lake Erken thermal conditions – Supplementary material

Simone Moras, Ana Isabel Ayala-Zamora, Don Pierson

Department of Ecology and Genetics, Uppsala university, Uppsala, 75236, Sweden

Correspondence to: Simone Moras (simone.moras@ebc.uu.se)

1. Data sources of driving parameters

Since October 12th 1988, the Erken meteorological station on Malma island has collected hourly data of wind speed (WS), air temperature (Air T) and and downwelling short wave radiation (SWR) in digital form using a data logging system. Prior to this, old paper records of hourly WS data from the Erken station were also available for the years 1964, 1966-1969 and 1972-1973. Old paper records of total daily SWR from Lake Erken were also available at SMHI archive (Swedish Meteorological and Hydrological Institute) measured in the periods 1961-1965 and 1969-1983. Hourly CC data were retrieved from Svenska Hogarna station (59.4445 N, 19.5059 E) for the entire period of study (1961-2016). These data were downloaded from SMHI website (http://opendata-download-metobs.smhi.se/explore/). Daily precipitation (DP) data were retrieved from Erken station paper records (1961-1965) and from several meteorological stations in the vicinity of Lake Erken: Norrtälje (59.7506 N, 18.7091 E, period 1966-1984), Norrveda (59.8298 N, 18.9524 E, periods Jun-Jul. 1978 and Feb.-Jun. 1979), Rimbo (59.7487 N, 18.3535 E, period 1984-1988), Norrtälje-Vasby (59.8524 N, 18.7296 E, period 1989-1994) and Svanberga (59.8321 N, 18.6348 E, period 1995-2016). DP data of these stations were downloaded from SMHI website. For the period 1995-2016, air pressure (Air P) and relative humidity (RH) data were provided by Svanberga weather station (59.8321 N, 18.6348 E), the closest SMHI station to the lake (about 800 m far from the Erken Labortory field station).

2. Wind speed prediction

Missing WS data were predicted using the Erken Malma island station wind speed data as target and wind speed datasets from different SMHI stations (downloaded from SMHI database) as input data (table 1). In Matlab version R2017b (MathWorks Inc. Natick, Massachusetts), five different ANN function fitting analyses (*nftool*) were developed to predict Lake Erken wind speed, using the maximum number of input datasets available for each time period. Once the input-target best fit function was calculated, WS was predicted for the entire time interval of the input datasets for each ANN analysis. WS data above 25 m/s were removed from the analysis. Figure 1 shows the regression plots of the 5 different ANN function fitting analyses developed for the overlapped time periods between input and target datasets and the error histogram associated with every analysis. If different analyses predicted wind speed at the same time point, the prediction with the highest correlation and smallest error between output and target was preferred. If none of the 5 fitting function analyses predicted wind speed data for a certain time point (i.e. none of the input datasets had a wind speed value at that time point) a linear interpolation was used to fill missing

data in the final wind speed dataset used in the GOTM model. In this case, 28700 missing data were replaced with a linear interpolation (total N = 499656). Figure 1 shows the graphical results of ANN *nftool* analysis for predicting wind speed.

Table 1: Schematic description of wind speed prediction using ANN analysis

WS ANN analysis no.	WS input data (SMHI stations) *	WS target data (station)	Compared years between inputs and target	Time interval of predicted data	No. of missing data replaced by predicted WS (% of total data)
1	UF+SB	Erken (Norr Malma)	1964, 1966-1969, 1972-1973, 1988-2016	1961-2017	6900 (1.4 %)
2	UF+SB+Ar	Erken (Norr Malma)	1964, 1966-1969, 1972-1973, 1988-1997	1962-1997	134199 (26.9 %)
3	UF+SB+Ar+Sk	Erken (Norr Malma)	1988-1997	1976-1997	20524 (4.1 %)
4	UF+SB+Ar+Sk+AUT	Erken (Norr Malma)	1985-1997	1985-1997	22526 (4.5 %)
5	UF+SB+AUT+Sv	Erken (Norr Malma)	1995-2016	1995-2017	4418 (0.9 %)
					Total 188567 (38.1%)

*UF = Uppsala Flygplats (59.8953 N, 17.5935 E), SB = Stockholm-Bromma (59.3537 N, 17.9513 E), Ar = Arlanda (59.6557 N, 17.9462 E), Sk = Skarpö A (59.3455 N, 18.7406 E), AUT = Uppsala Aut (59.8586 N, 17.6253 E), Sv = Svanberga (59.8321 N, 18.6348 E)

(a)

WS ANN 1

(b)

WS ANN 1

(c)

WS ANN 2

(d)

WS ANN 2

(e)

WS ANN 3

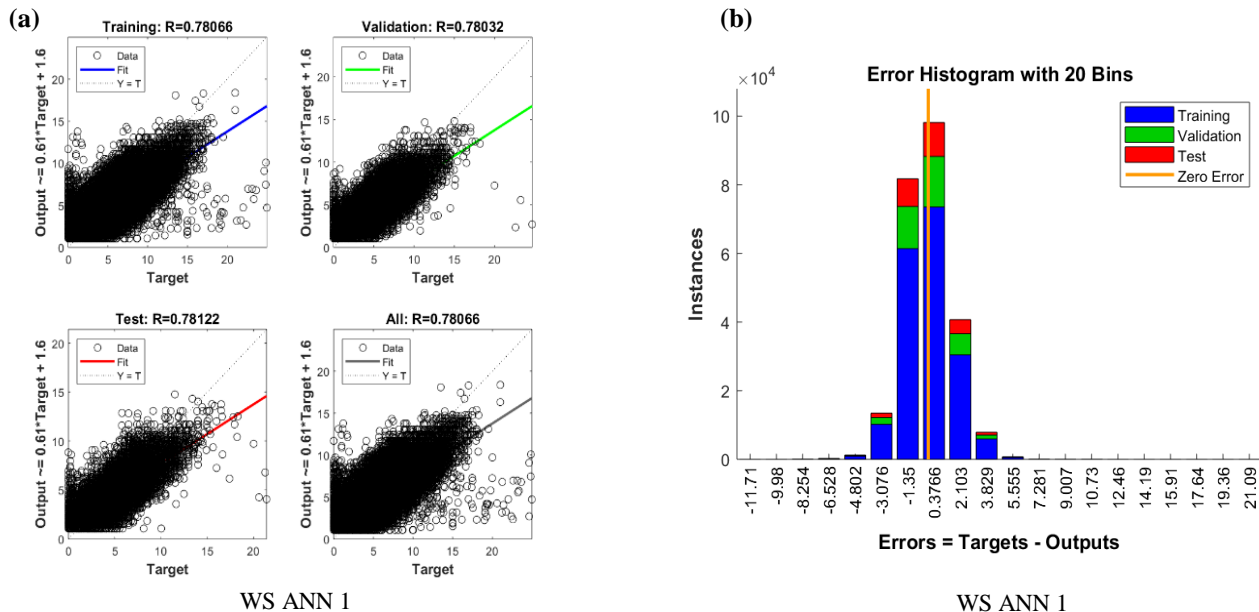
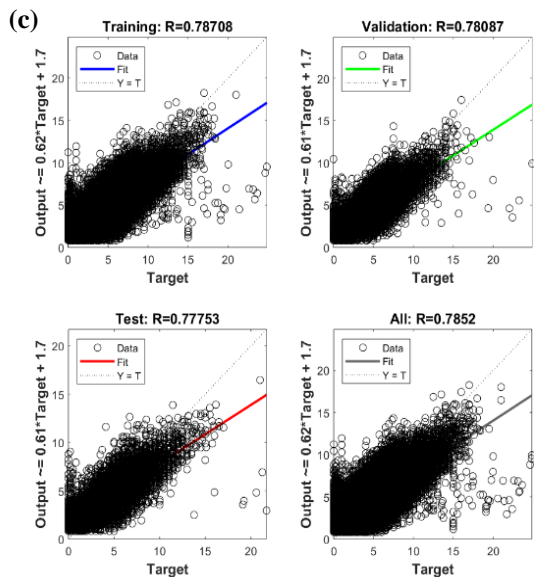
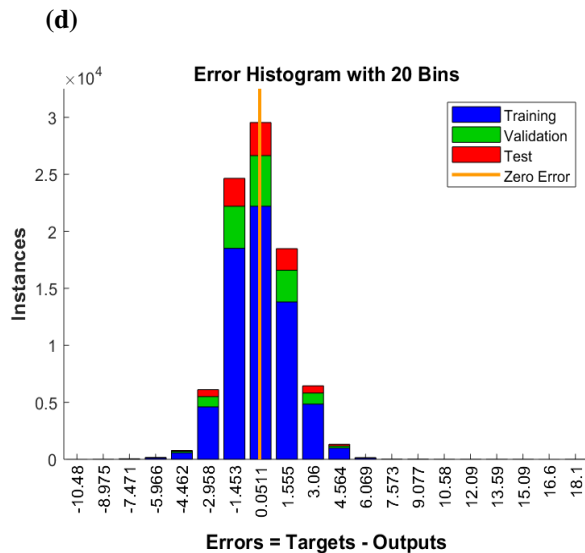
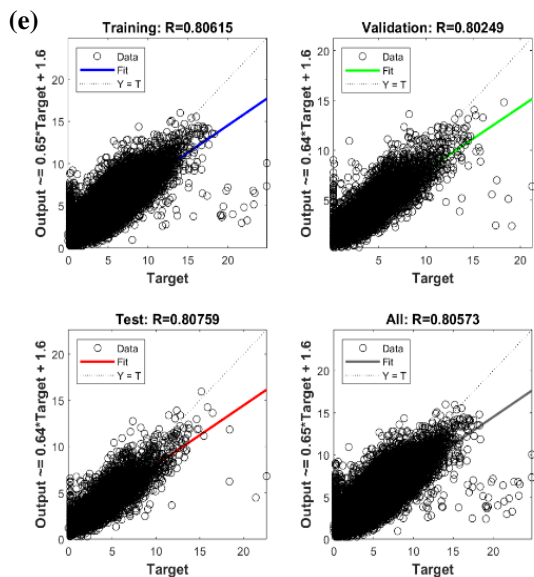
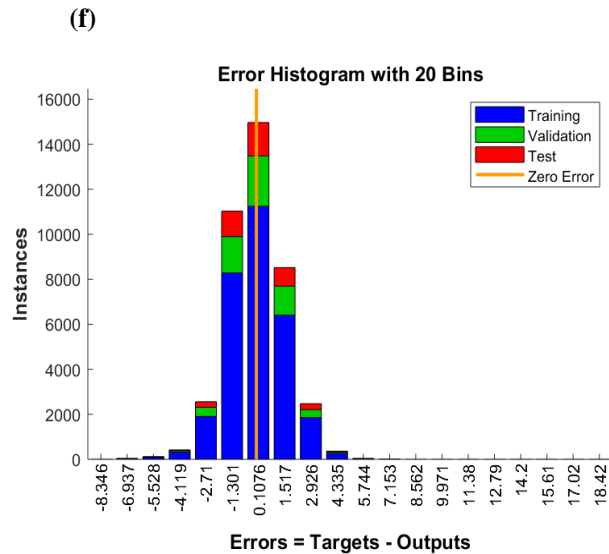
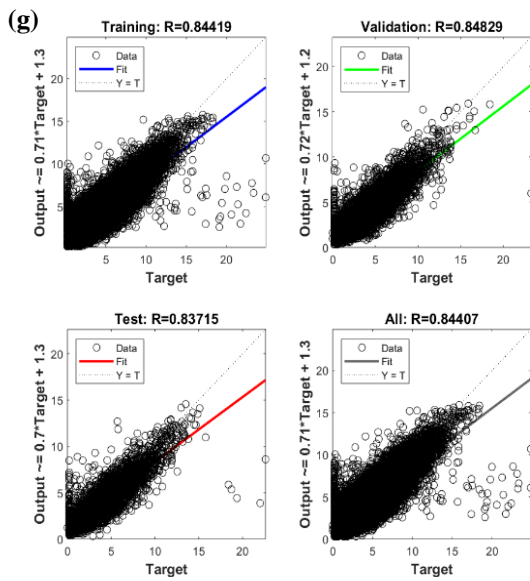
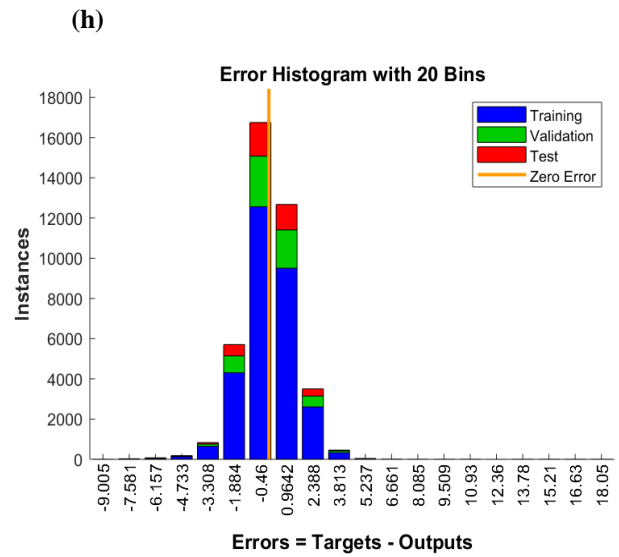
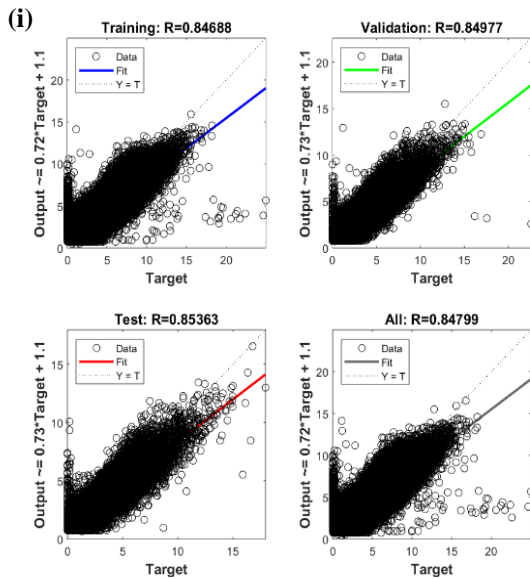
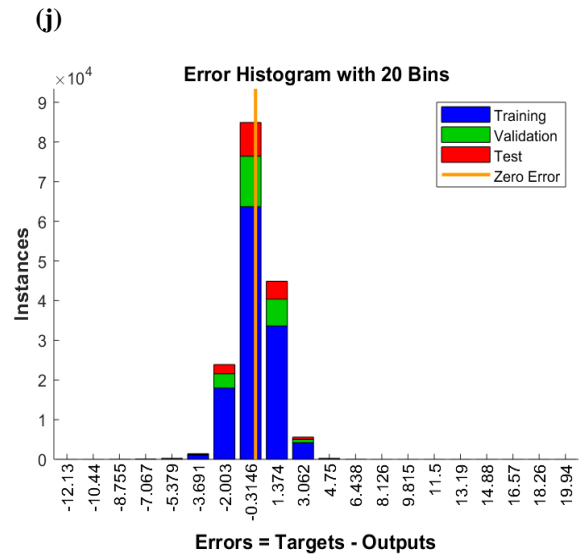
(f)

WS ANN 3

Figure 1: Graphical results of WS ANN 1-5. Panels (a), (c), (e), (g), (i) show the regression plots between target (Erken WS) and output for training, validation, test and overall regression; panels (b), (d), (f), (j) show the error histogram plots of the different WS ANN analyses performed to predict wind speed.

3. Air temperature prediction

Hourly air temperature (Air T) have been recorded by the Erken Malma Island weather station starting from October 12th, 1988 and continues today. The number of hourly Air T data recorded by the Erken station in the period Oct 1988-Dec 2017 is N=235257. The number of missing values in the same time interval is NA = 14696. To complete air temperature dataset for the period 1961-1988, Erken Air T was estimated from air temperature of Stockholm-Bromma (SB) and Uppsala Flygplats (UF) stations using ANN function fitting analysis (*nftool*) in Matlab version R2017b (MathWorks Inc. Natick, Massachusetts). To replace missing data in the time interval 1988-2016, Erken Air T was predicted from SB, UF, Arlanda (Ar) and Uppsala AUT for the period 1988-1997. For the period 1998-2000 Erken Air T missing values were predicted from SB, UF, AUT and Söderarm (S). Missing values for the period 2001-2016 were predicted from air temperature of SB, UF, AUT, S and Films Kirkby (FK) stations (table 2). The remaining missing data that were not possible to predict using air temperature data from neighbouring stations were replaced by linear interpolation (missing data replaced by linear interpolation over the period 1961-2016: NA=29417). Results of Air T ANN function fitting analyses are shown in figure 2.

Table 2: Schematic description of air temperature prediction using ANN analysis

Air T ANN analysis no.	Air T input data (SMHI stations) *	Air T target data (station)	Compared years between inputs and target	Time interval of predicted data	No. of missing data replaced by predicted WS (% of total data)
1	UF+SB	Erken (Norr Malma)	1988-2016	1961-1988	224228 (45.2 %)
2	UF+SB+Ar+AUT	Erken (Norr Malma)	1988-1997	1988-1997	6111 (1.2 %)
3	UF+SB+AUT+S	Erken (Norr Malma)	1998-2000	1998-2000	66 (0.1 %)
4	UF+SB+AUT+S+FK	Erken (Norr Malma)	2001-2016	2001-2016	4577 (0.9%)
				Total	**188567 (38.1%)**

*UF = Uppsala Flygplats (59.8953 N, 17.5935 E), SB = Stockholm-Bromma (59.3537 N, 17.9513 E), Ar = Arlanda (59.6557 N, 17.9462 E), AUT = Uppsala Aut (59.8586 N, 17.6253 E), S = Söderarm (59.7538 N, 19.4089 E), FK = Films Kirkby (60.2363 N, 17.9078 E)

(a)

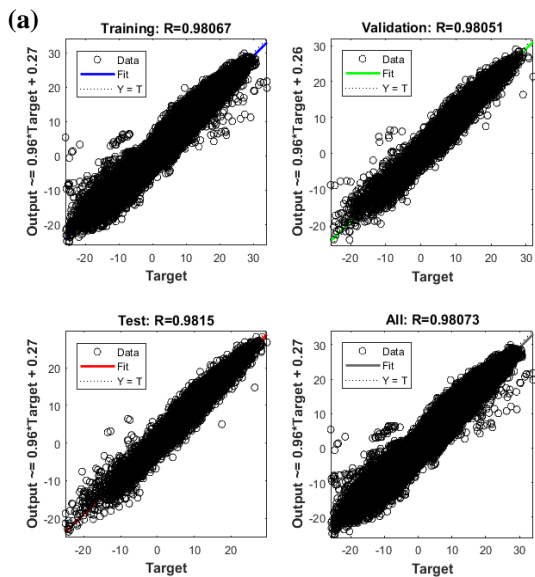
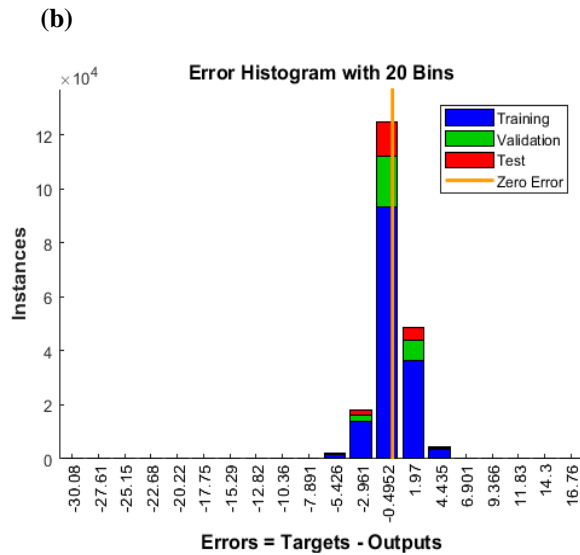
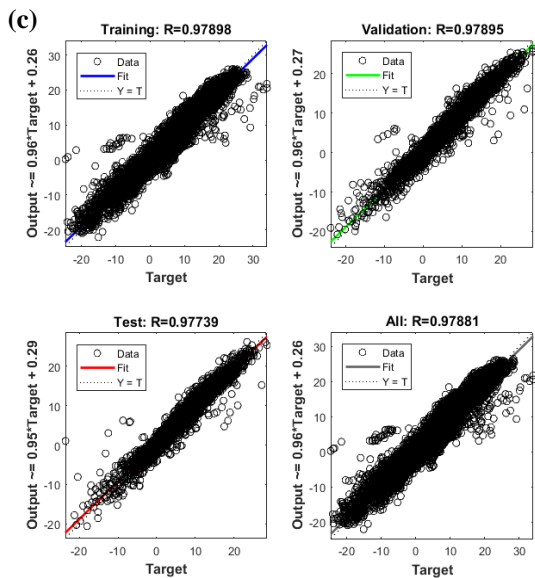
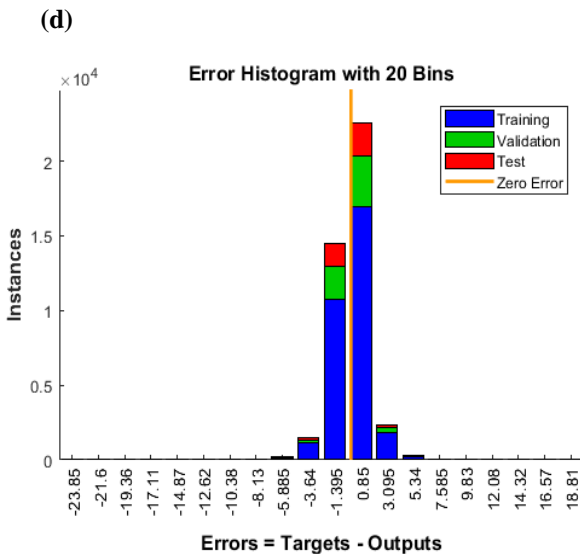
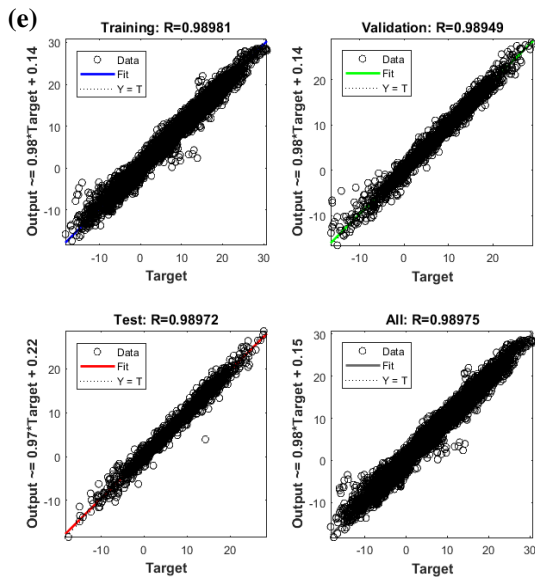
Air T ANN 1

(b)

Air T ANN 1

(c)

Air T ANN 2

(d)

Air T ANN 2

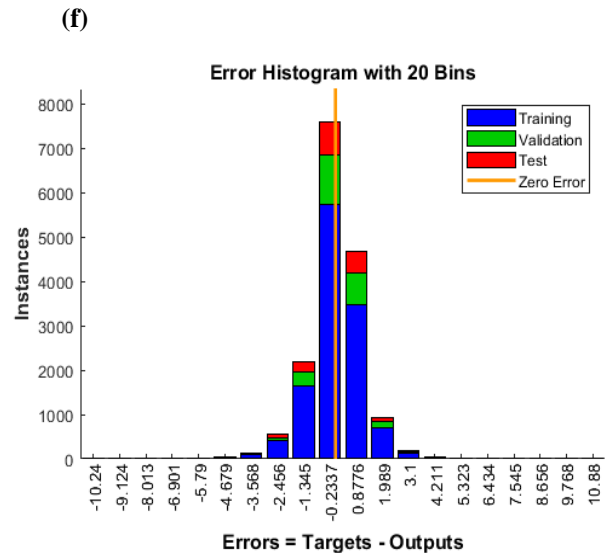
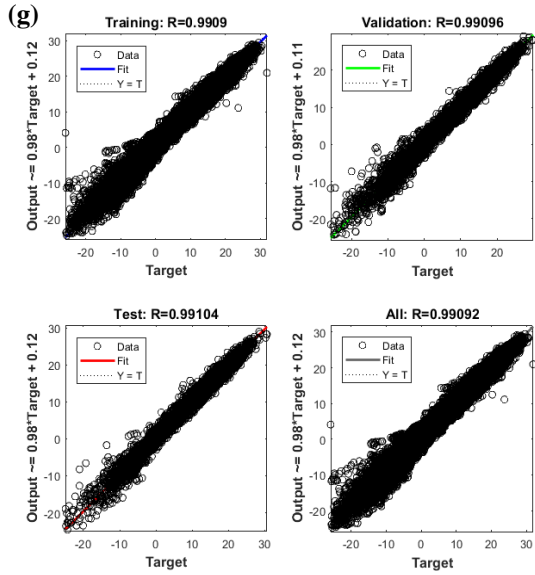
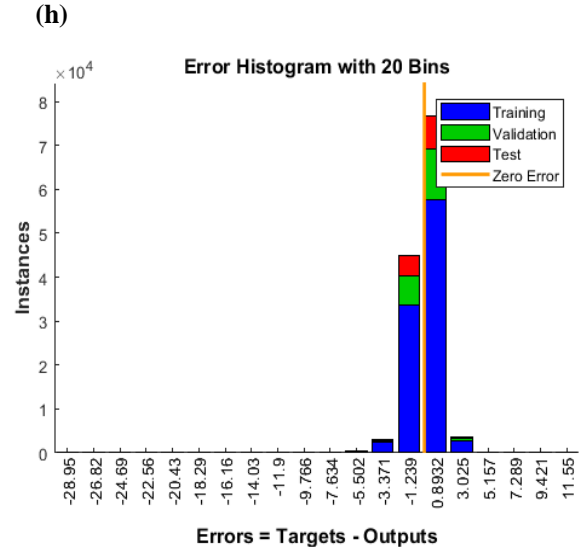
Air T ANN 3

Air T ANN 3

Air T ANN 4

Air T ANN 4

Figure 2: Graphical results of AirT ANN 1-4. Panels (a), (c), (e), (g) show the regression plots between target (Erken AirT) and output for training, validation, test and overall regression; panels (b), (d), (f), (h) show the error histogram plots of the different AirT ANN analyses performed to predict Air Temperature.

4. Relative humidity prediction

Hourly relative humidity (RH) from the Svanberga station was used as the target dataset in ANN analysis, since no humidity data have been collected from Erken Malma Island station. Svanberga relative humidity records start on August 1st 1995 and continue to present days. The number of RH data collected from Svanberga station during the period August 1995-December 2017 is N=191678. To predict relative humidity for the period 1961-1995, relative humidity from UF was used as input data in ANN best fit function analysis. Missing data between 1995-2017 have been predicted from UF and AUT relative humidity (Table 3). These humidity data were retrieved from SMHI database (Thomas Carlund, personal communication). Output-target correlation and related error of ANN analyses are shown in figure 3. Missing data that could not be predicted from ANN function fitting analyses have been replaced using linear interpolation (number of missing data replaced by linear interpolation over the period 1961-2017: NA = 13886).

Table 3: Schematic description of relative humidity prediction using ANN analysis

RH ANN analysis no.	RH input data (stations)*	RH target data (station)	Compared years between inputs and target	Time interval of predicted data	No. of missing data replaced by predicted RH (% of total data)
1	UF	Svanberga	1995-2017	1961-2017	228757 (45.8%)
2	UF+AUT	Svanberga	1995-2017	1986-2017	65335(13.1%)
					Total 294092 (58.9%)

*UF = Uppsala Flygplats (59.8953 N, 17.5935 E), AUT = Uppsala Aut (59.8586 N, 17.6253 E)

(a)

RH ANN 1

(b)

RH ANN 1

RH ANN 2   RH ANN 2

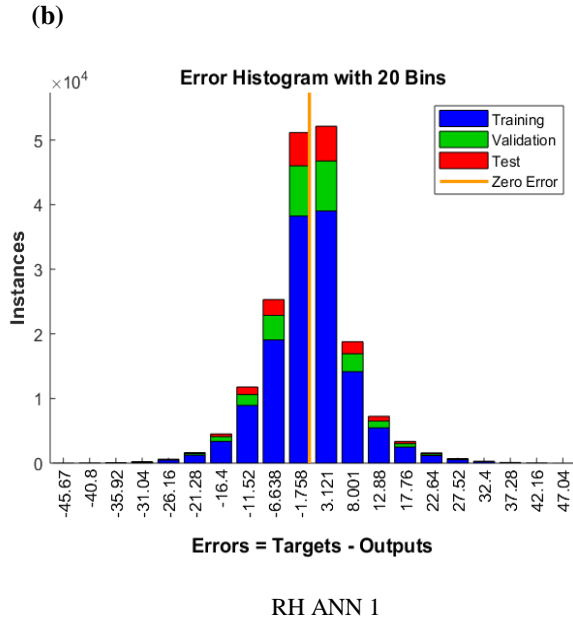
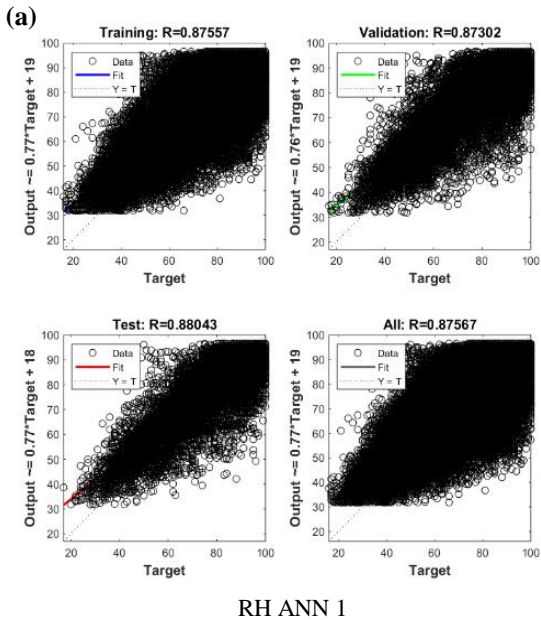
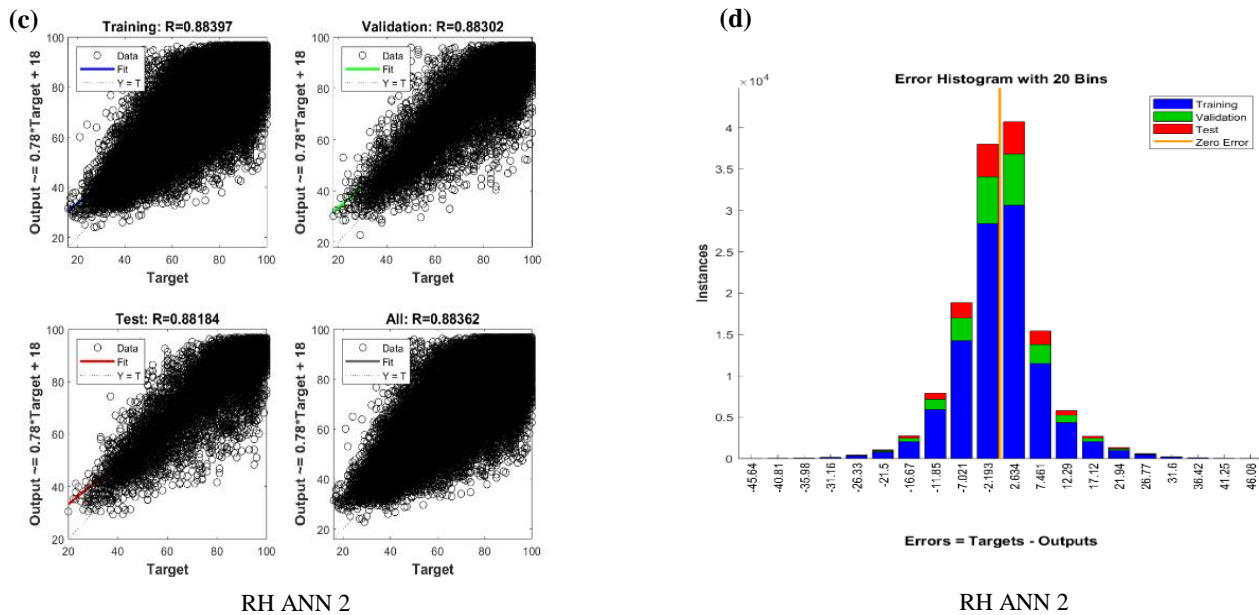



Figure 3: Graphical results of RH ANN 1-2. Panels (a), (c) show the regression plots between target (Svanberga RH) and output for training, validation, test and overall regression; panels (b), (d) show the error histogram plots of the different RH ANN analyses performed to predict relative humidity.

5. Air pressure prediction

Since air pressure (Air P) data from Malma island were available only between Nov 2016-Dec 2017, hourly air pressure prediction with ANN was performed using Svanberga Air P dataset (N=185026) as target and air pressure datasets from UF and SB as inputs data for predicting Air P missing data over the period 1961-2016 (table 4). Missing data that could not be predicted by ANN function fitting analysis were replaced by linear interpolation (NAs replaced by linear interpolation: 45212). Output-target regression plots and errors are shown in figure 4.

Table 4: Schematic description of air pressure prediction using ANN analysis

Air P ANN analysis no.	Air P input data (station)*	Air P target data (station)	Compared years between inputs and target	Time interval of predicted data	No. of missing data replaced by predicted WS (% of total data)
1	UF+SB	Svanberga	1995-2016	1961-2016	259563
					Total 259563 (53.03 %)

*UF = Uppsala Flygplats (59.8953 N, 17.5935 E), SB = Stockholm-Bromma (59.3537 N, 17.9513 E)

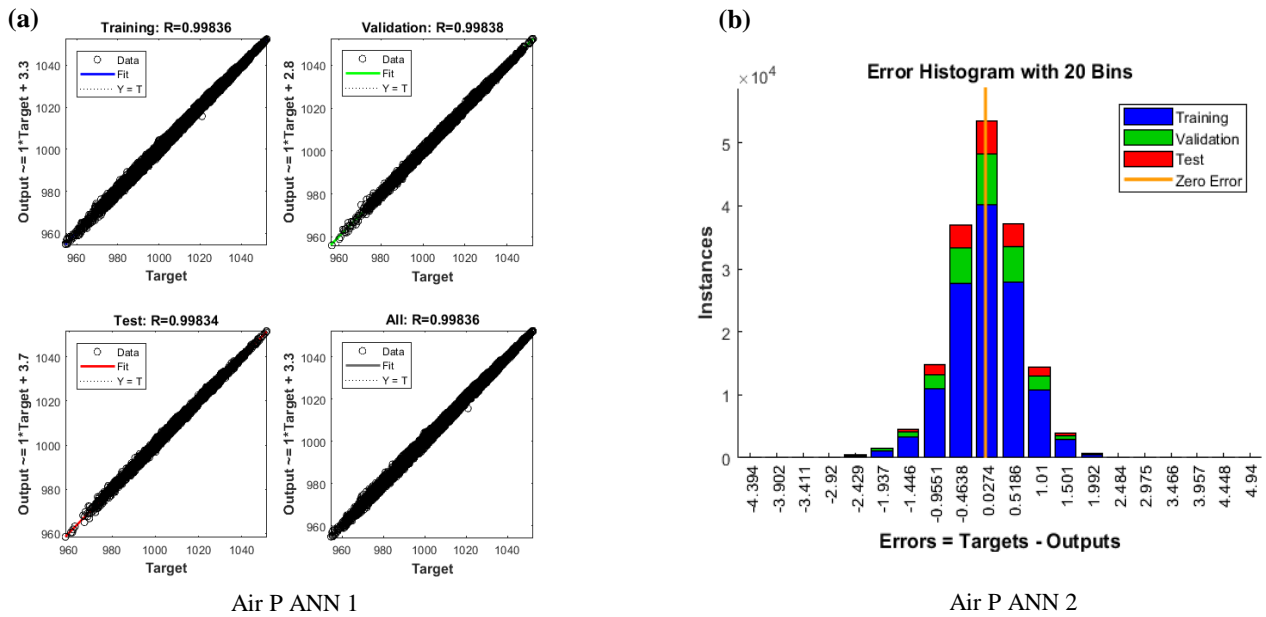
(a)   (b)

Air P ANN 1   Air P ANN 2

Figure 4: Graphical results of Air P ANN. Panel (a) shows the regression plots between target (Svanberga Air P) and output for training, validation, test and overall regression; panels (b) shows the error histogram plots of the Air P ANN analysis performed to predict Air P.

6. Shortwave radiation prediction

Daily shortwave radiation (SWR) data from the Swedish Agricultural University (SLU) Ultuna weather station were used to predict missing data of Erken SWR using ANN function fitting analysis (table 5). Daily SWR data from Ultuna station were available from 1973-1999. Early records of SWR measured on Malma island were also available from SMHI. Since the SWR data from both SLU and SMHI were only available as total daily SWR, the total daily SWR was predicted for Erken using ANN function fitting analysis instead of hourly SWR. Missing data for which SWR could not be predicted by ANN function fitting analysis (258 points over the period 1961-2016) were replaced by averaging SWR daily values measured at the same date of the missing value but in different years. To estimate hourly SWR values from total daily SWR, we used the method described by Martin and McCutcheon (1999, see section 6.1). The measured daily total solar inputs were distributed over a daily cycle based on the idealized hourly variations in solar radiation. Finally, as a quality check the obtained values of hourly shortwave solar radiation were compared with the theoretical values. From this it became apparent that the measured shortwave radiation of the years 1970, 1971,1973,1974 were overestimated. To correct for this these data were reduced by 15.27 %, 28.38%, 15.82% and 10.19 % respectively, due to a probable overestimation of solar radiation throughout the entire year (see section 6.2).

Table 5: Schematic description of shortwave radiation prediction using ANN analysis

Daily SWR ANN analysis no.	Daily SWR input data (SLU station)*	Daily SWR target data (station)	Compared years	Predicted years	Missing data replaced by predicted daily SWR (% of total data)
1	Ultuna	Erken	1973-1999	1961-2016	3972
					Total 3972 (19.1 %)

*Ultuna (59.8175 N, 17.6536 E)

SWR ANN 1   Air P ANN 2

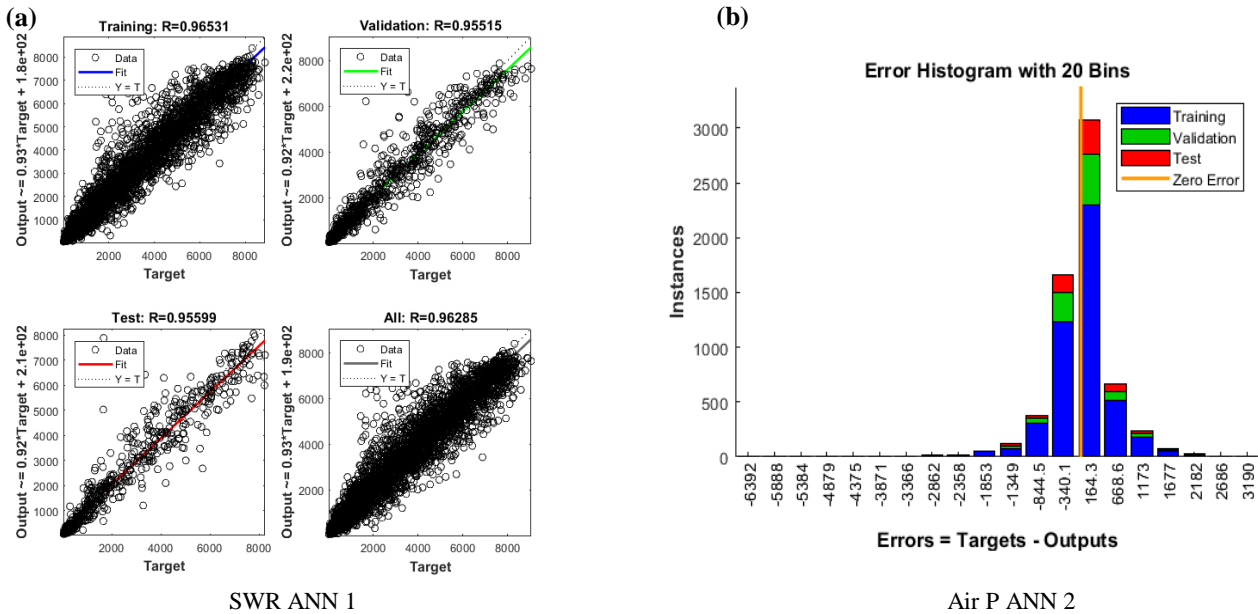



Figure 5: Graphical results of SWR ANN 1. Panel (a) shows the regression plots between target (Erken SWR) and output for training, validation, test and overall regression; panel (b) shows the error histogram plots of the SWR ANN analysis performed to predict short-wave solar radiation.

6.1 Calculation of theoretical shortwave radiation

The equations for calculating the theoretical shortwave radiation H_{sw} that reaches the Earth surface according to Martin and McCutcheon (1999) are here presented for Lake Erken. H_{sw} is defined in Eq. (1):

$$H_{sw} = H_0 * a_t * (1 - R_s) * C_a, \tag{1}$$

10   where:

H_{sw}: short-wave radiation (W/m^2)

H_0: radiation that reaches the outer atmosphere (W/m²)

a_t: atmospheric transmission term

R_s: reflection coefficient

C_a: fraction of solar radiation not absorbed by clouds.

5   Equation (2) shows how H_0 is calculated:

$$H_0 = \frac{H_{sc}}{r^2} * \left\{ \sin\left(\frac{\pi*\theta}{180}\right) * \sin(\delta) + \frac{12}{\pi} * \cos\left(\frac{\pi*\theta}{180}\right) * \cos(\delta) * [\sin(h_e) - \sin(h_b)] \right\} * \Gamma, \tag{2}$$

where:

H_{sc}: solar constant (1390 W/m²)

r: Earth-Sun relative distance

10   θ: latitude (degrees)

δ: declination

h_e: solar hour angle (radians) at the end of time interval in which H_0 is estimated

h_b solar hour angle (radians) at the beginning of time interval in which H_0 is estimated

Γ: correction factor for diurnal radiation exposure

15   Equation (3) describe how to estimate r:

$$r = 1 + 0.017 * \cos\left[\frac{2\pi}{365} * (186 - D_y)\right], \tag{3}$$

where:

D_y: day of the year (1-365,366)

δ is defined in Eq. (4):

20    $$\delta = \frac{23.45}{185} * \cos\left[\frac{2\pi}{365} * (172 - D_y)\right], \tag{4}$$

h_e and h_b are calculated according to Eq. (5) and Eq. (6) respectively:

$$h_e = \left[\frac{\pi}{12} * (h_r - \Delta t_s + a * \pi)\right] + b * (2\pi), \tag{5}$$

$$h_b = \left[\frac{\pi}{12} * ((h_r - 1) - \Delta t_s + a * \pi)\right] + b * (2\pi), \tag{6}$$

where:

25   h_r: hour of the day between 1-24

$a = 1$ if $h_r \leq 12$, $a = -1$ if $h_r > 12$

b: coefficient that varies according to the value within the square brackets [] of equation (5) and (6). $b = -1$ if [] $> 2\pi$, $b = 1$ if [] < 0 and $b = 0$ in all the other cases.

Δt_s: fraction of local meridian west of the standard meridian.

Δt_s is defined for Lake Erken in Eq. (7):

$$\Delta t_s = \frac{E_a}{15}(L_{sm} - L_{lm}),\qquad(7)$$

where:

5   L_{sm}: local standard meridian (15°)

L_{lm}: local meridian (18.63°)

$E_a = 1$ for east longitude.

Before the calculation of correction factor Γ that appears in Eq. (2), it is essential to set solar radiation to 0 before the sunrise and after the sunset. The standard time of sunset t_{ss} and the standard time of sunrise t_{su} are defined in Eq. (8) and Eq. (9)

10   respectively:

$$t_{ss} = \frac{12}{\pi} * cos^{-1}\left(-\frac{sin\left(\frac{\pi*\theta}{180}\right)*sin(\delta)}{cos\left(\frac{\pi*\theta}{180}\right)*cos(\delta)}\right) + \Delta t_s + 12,\qquad(8)$$

$$t_{su} = -t_{ss} + 2 * \Delta t_s + 24,\qquad(9)$$

Therefore, the correction factor Γ is equal to 1 when $t_{su} < h_r < t_{ss}$ and 0 otherwise.

Eeuation (10) defines the atmospheric transmission term a_t:

$$a_t = \frac{a_2 + 0.5*(1 - a_1 - c_d)}{1 - 0.5*R_g*(1 - a_1 - c_d)},\qquad(10)$$

where:

c_d: dust coefficient (0.06)

$R_g = 0.045$: ground reflection coefficient of catchment area (extended mixed forest)

a_1 and a_2: atmospheric transmission coefficients dependent on air moisture content and optical air mass.

20   a_1 and a_2 are defined according to Eq. (11) and Eq. (12):

$$a_1 = exp[-(0.465 + 0.134 * P_{wc}) * (0.129 + 0.171 * \exp(-0.88 * \vartheta_{am}))] * \vartheta_{am},\qquad(11)$$

$$a_2 = exp[-(0.465 + 0.134 * P_{wc}) * (0.179 + 0.421 * \exp(-0.721 * \vartheta_{am}))] * \vartheta_{am},\qquad(12)$$

where:

P_{wc}: mean daily precipitable water content

25   ϑ_{am}: optical mass

P_{wc} is defined in Eq. (13):

$$P_{wc} = 0.85 * exp(0.11 + 0.0614 * T_d), \tag{13}$$

where:

T_d: dewpoint temperature (°C)

The dewpoint is calculated in Eq. (14) according to Lawrence (2005) using the relative humidity (RH) and air temperature

5 (AirT) used to drive the GOTM model:

$$T_d = \frac{243.04 * \left[ln\left(\frac{RH}{100}\right) + \frac{17.625 * AirT}{243.04 + AirT} \right]}{17.625 - ln\left(\frac{RH}{100}\right) - \frac{17.625 * AirT}{243.044 + AirT}}, \tag{14}$$

The optical mass ϑ_{am} is defined in Eq. (15):

$$\vartheta_{am} = \frac{\left(\frac{288 - 0.0065 * Z}{288}\right)^{5.256}}{sin(\alpha) + 0.15 * \left(\frac{\alpha * 180}{\pi} + 3.855\right)^{-1.253}}, \tag{15}$$

where:

10 $Z = 11$: altitude of Lake Erken (m)

α: solar altitude (radians)

The solar altitude is defined in Eq. (16):

$$\alpha = tan^{-1}\left(\frac{\alpha_1}{\sqrt{1 - \alpha_1^2}}\right), \tag{16}$$

where α_1 is defined in Eq. (17):

15 $$\alpha_1 = \left| sin\left(\frac{\pi * \theta}{180}\right) * sin(\delta) + cos\left(\frac{\pi * \theta}{180}\right) * cos(\delta) * cos\left(\frac{h_e + h_b}{2}\right) \right|, \tag{17}$$

The reflection factor, or albedo, (R_s) is defined in Eq. (18):

$$R_s = a * \left(\frac{180}{\pi} * \alpha\right)^b, \tag{18}$$

where:

α: solar altitude (radians)

20 a and b: coefficients dependent on cloud cover (CC). In particular:

If $CC > 0.9, a = 0.33, b = -0.45$

If $0.5 < CC < 0.9, a = 0.95, b = -0.75$

If $0.1 < CC < 0.5, a = 2.20, b = -0.97$

If $CC < 0.1, a = 1.18, b = -0.77$

25

The fraction of solar radiation not absorbed by clouds (c_a) is defined in Eq. (19):

$$C_a = 1 - 0.65 * CC^2, \tag{19}$$

where:

CC: cloud cover.

6.2 Shortwave radiation reduction

5   The comparison graphs (fig. 6-9) between measured and theoretical SWR are shown for the years 1970, 1971, 1973 and 1974. A simple linear regression has been performed to compare measured and theoretical SWR. The slope of the regression line has been used to calculate the percentage of reduction of the overestimated measured SWR for the years 1970, 1971, 1973 and 1974.

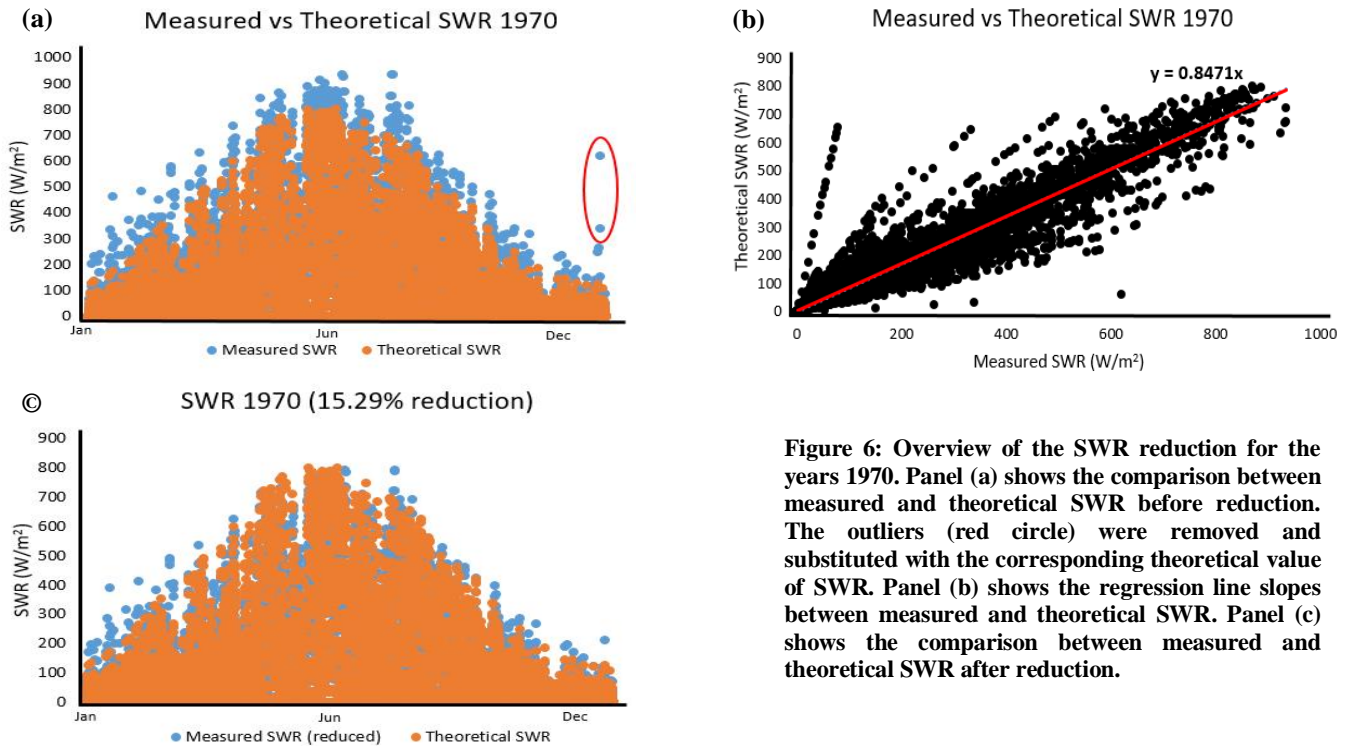



Figure 6: Overview of the SWR reduction for the years 1970. Panel (a) shows the comparison between measured and theoretical SWR before reduction. The outliers (red circle) were removed and substituted with the corresponding theoretical value of SWR. Panel (b) shows the regression line slopes between measured and theoretical SWR. Panel (c) shows the comparison between measured and theoretical SWR after reduction.

(a)

(c)

(b)

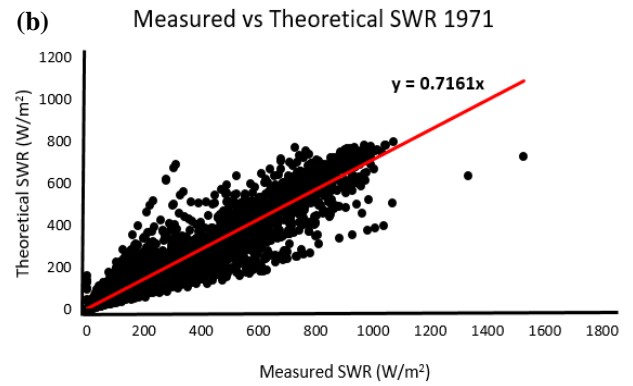
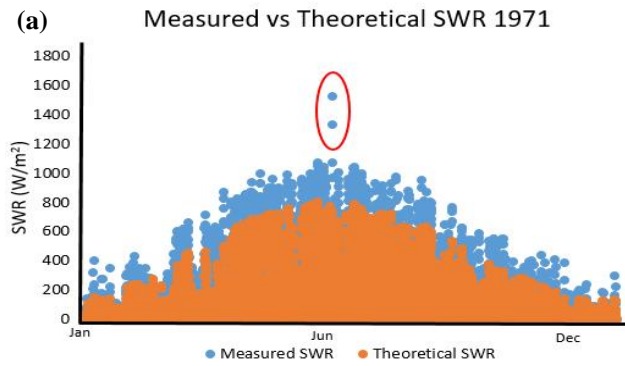



Figure 7: Overview of the SWR reduction for the years 1971. Panel (a) shows the comparison between measured and theoretical SWR before reduction. The outliers (red circle) were removed and substituted with the corresponding theoretical value of SWR. Panel (b) shows the regression line slopes between measured and theoretical SWR. Panel (c) shows the comparison between measured and theoretical SWR after reduction.

(a)

(c)

(b)

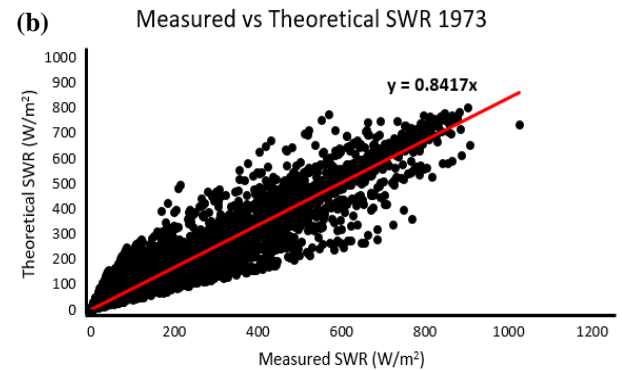
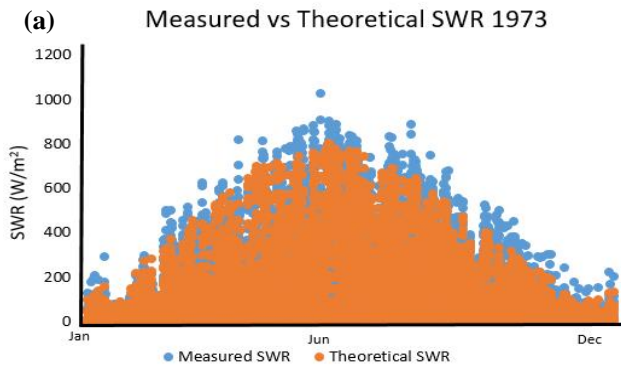
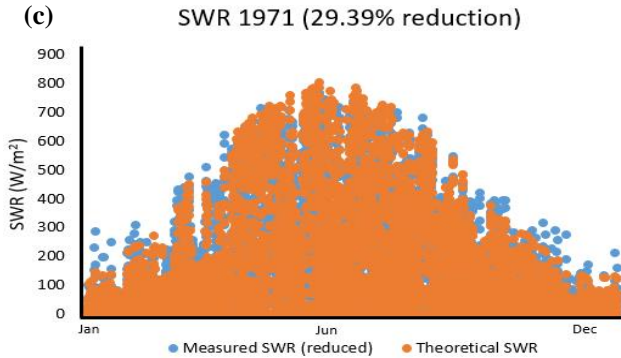



Figure 8: Overview of the SWR reduction for the years 1973. Panel (a) shows the comparison between measured and theoretical SWR before reduction. Panel (b) shows the regression line slopes between measured and theoretical SWR. Panel (c) shows the comparison between measured and theoretical SWR after reduction.

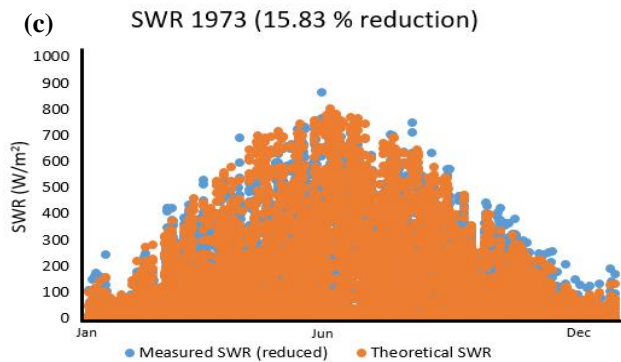
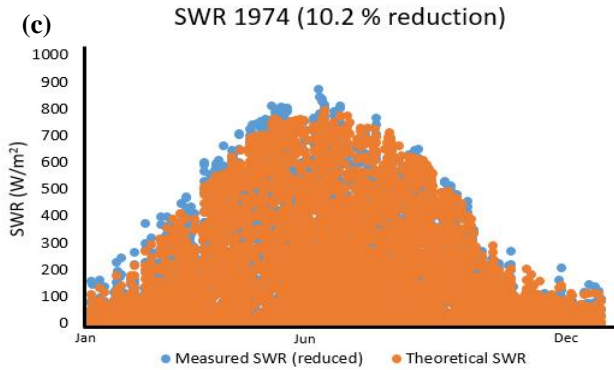
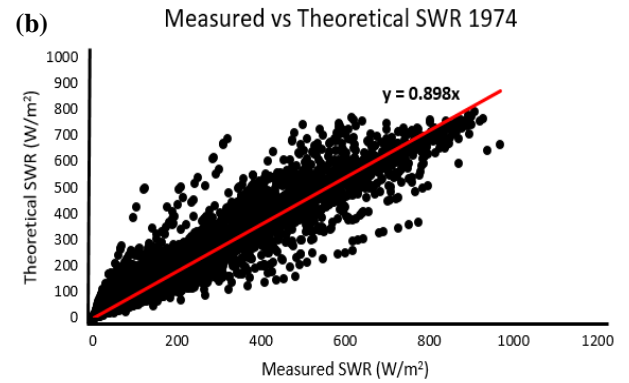
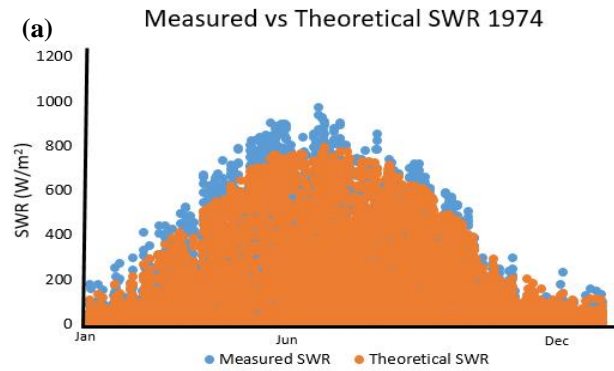



Figure 9: Overview of the SWR reduction for the years 1974. Panel (a) shows the comparison between measured and theoretical SWR before reduction. Panel (b) shows the regression line slopes between measured and theoretical SWR. Panel (c) shows the comparison between measured and theoretical SWR after reduction.

References

5   Lawrence, M. G.: The relationship between relative humidity and the dewpoint temperature in moist air, Bull. Am. Meteorol. Soc., 86, 225–233, doi:10.1175/bams-86-2-225, 2005.

Martin, J. L. and McCutcheon, S. C.: Hydrodynamics and transports for water quality modelling, CRC Press, 1998.

MATLAB. 1992-2017. Neural Network Toolbox. The MathWorks Inc., Natick, Massachusetts, United States.