RECONSTRUCTION OF REGIONAL MEAN SEA LEVEL ANOMALIES FROM TIDE GAUGES USING THE NEURAL NETWORK APPROACH

M. Wenzel and J. Schröter

Alfred-Wegener-Institute for Polar and Marine Research, Bremerhaven, Germany

The Neural Network

Regional mean sea level anomalies (RMSLA) are estimated from tide gauge values directly using the neural network approach. A neural network is an artificial neural system, a computational model inspired by the notion of neurophysical processes. It consists of several processing elements called neurons, which are interconnected with each other exchanging information (see e.g. Freeman and Skapura, 1991) . In this presentation a *backpropagation network* (BPN) is used. In this type of network the neurons are ordered into layers: an input layer on the top, one or more hidden layers below and an output layer at the bottom.

Figure 1 shows a BPN that is enriched by direct connections between the input and the output layer. Its output $\vec{y}$ in dependence to the input $\vec{x}$ can be desribed by the equation:

$$\vec{y} = O(\vec{b}_o + \mathbf{WIO} \cdot \vec{x} + \mathbf{WHO} \cdot \mathcal{H}(\vec{b}_h + \mathbf{WIH} \cdot \vec{x}))$$

where $\mathcal{H}()$ and $O()$ desribe the transfer functions of the hidden and the output neurons, respectively. The matrices of the connection strength between the neurons from the different layers (**WIO**, **WIH** and **WHO**) as well as the bias terms $\vec{b}_h$ and $\vec{b}_o$ are estimated in a training phase, i.e. the BNP learns from given examples. This leads to a costfunction that is minimized by gradient descent.
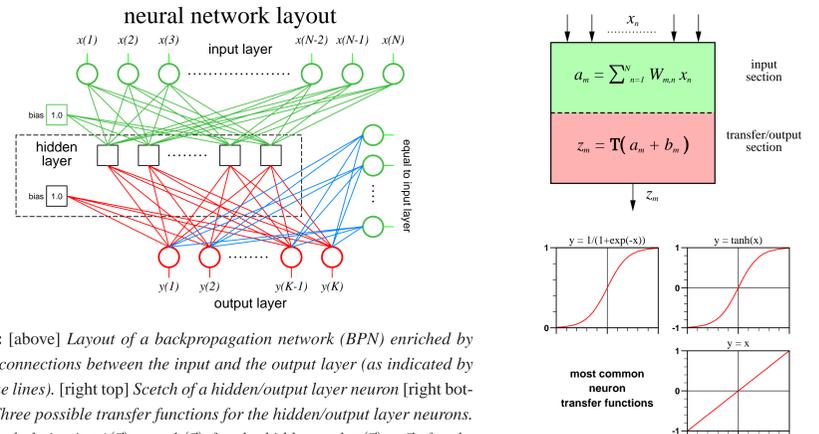
Fig. 1: [above] *Layout of a backpropagation network (BPN) enriched by direct connections between the input and the output layer (as indicated by the blue lines).* [right top] *Sketch of a hidden/output layer neuron* [right bottom] *Three possible transfer functions for the hidden/output layer neurons. The final choice is $\mathcal{H}(\vec{x}) = \tanh(\vec{x})$ for the hidden and $O(\vec{x}) = \vec{x}$ for the output layer.*

The Data

For our purpose 56 tide gauges are selected from the PSMSL monthly data that comply with the following conditions:

- there are more than 11 annual mean values given in [1993,2005]
- more than 50 annual mean values are given in [1900,2007] and
- the tide gauge is neighboured by at least one ocean point on a 1° x 1°grid.

The selected tide gauges are GIA corrected using the Peltier `ICE5G_VM4_L90` dataset also available on the PSMSL web side.

A single BPN is trained to compute all regional mean SLA's (trop. Indian Ocean, ... South Atlantic Ocean to Global Ocean) at once from the tide gauge values. To avoid possible problems with the different local reference frames all computations are done in the space of temporal derivatives. Beyond that, this makes the data more suitable for the BPN because it better limits the possible range of the numerical values.

To train the BPN known regional mean target values are needed. These values are derived either from the TOPEX/Poseidon data processed by GFZ Potsdam (T.Schöne, S.Esselborn pers. communication) and/or from the combined TOPEX/Poseidon and Jason-1 sea level fields available at the CSIRO sea level webpage (`www.cmar.csiro.au/sealevel/sl_data_cmar.html`). Differences between these datasets appear mainly in the tropical belt (15°N-15°S, e.g. the tropical Pacific, Fig. 3 left) and are also visible in the global mean (Fig. 3 right).
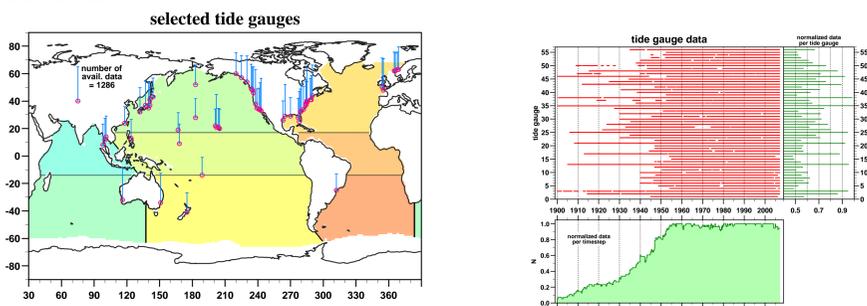
Fig. 2: *In the plate on the left the positions of the selected tide gauges and the corresponding amount of data are given by the red circles and the vertical bars, respectively. The data availability is demonstrated in more detail in the graphs on the right.*
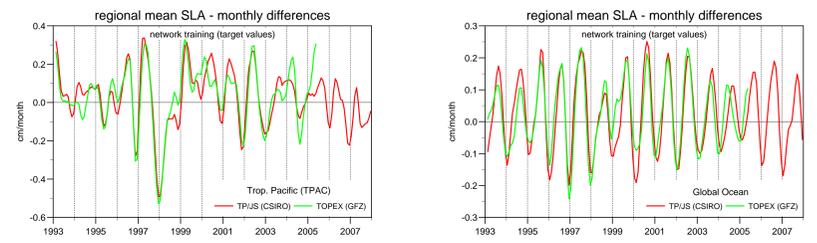
Fig. 3: *Comparing the RMSLA's (monthly differences) from the CSIRO and the GFZ dataset for the tropical Pacific (left) and the global ocean (right).*

Filling Tide Gauge Data Gaps

Although every tide gauge has more then 50 years of data, many values are missing, especially prior to 1950 (see Fig. 2). To fill these data gaps at the input layer of the BPN several alternatives (see Table on the right) are tested. This includes a reconstruction using an EOF basis estimated from all timesteps that have a complete tide gauge dataset. Furthermore a *forecast network* is build, that is trained to compute the values at all tide gauge positions for timestep (n+1) from all values at the steps (n) and (n-1). Additionally an equivalent *backcast network* is constructed that computes the values for step (n-1) from the steps (n) and (n+1). Each of these networks has the following dimension: 112 input, 224 hidden and 56 output neurons. The best estimate, i.e. with minimal error at known data points, is achieved at most timesteps by the backcast network with input gaps filled by EOF reconstruction (case 8, ∼40%) and by the forecast network with EOF filling (case 5, ∼34%).

acronym	method
1: mac	mean annual cycle (MAC)
2: eof	EOF reconstruction (EOFR)
3: fc/recurr	forecast network, recurrent, reset input to known values
4: fc/mac fill	forecast network, input gaps filled by MAC
5: fc/eof fill	forecast network, input gaps filled by EOFR
6: bc/recurr	backcast network, recurrent, reset input to known values
7: bc/mac fill	backcast network, input gaps filled by MAC
8: bc/eof fill	backcast network, input gaps filled by EOFR
9: fc/bc best	best of 3 to 8 (minimal fore-/backcast error at known values)
10: fc/bc mean	error weighted mean of 3 to 8
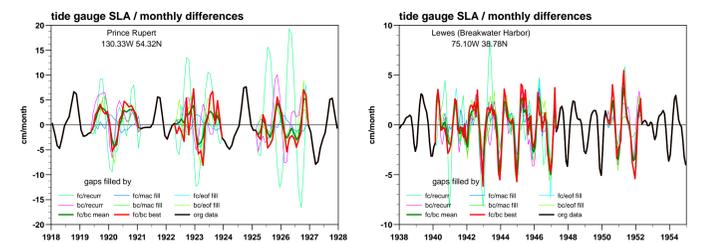
Fig. 4: *Data gap filling given for two example the tide gauges: Prince Rupert (left) and Lewes, Breakwater Harbor (right)*

Regional Sea Level Anomaly
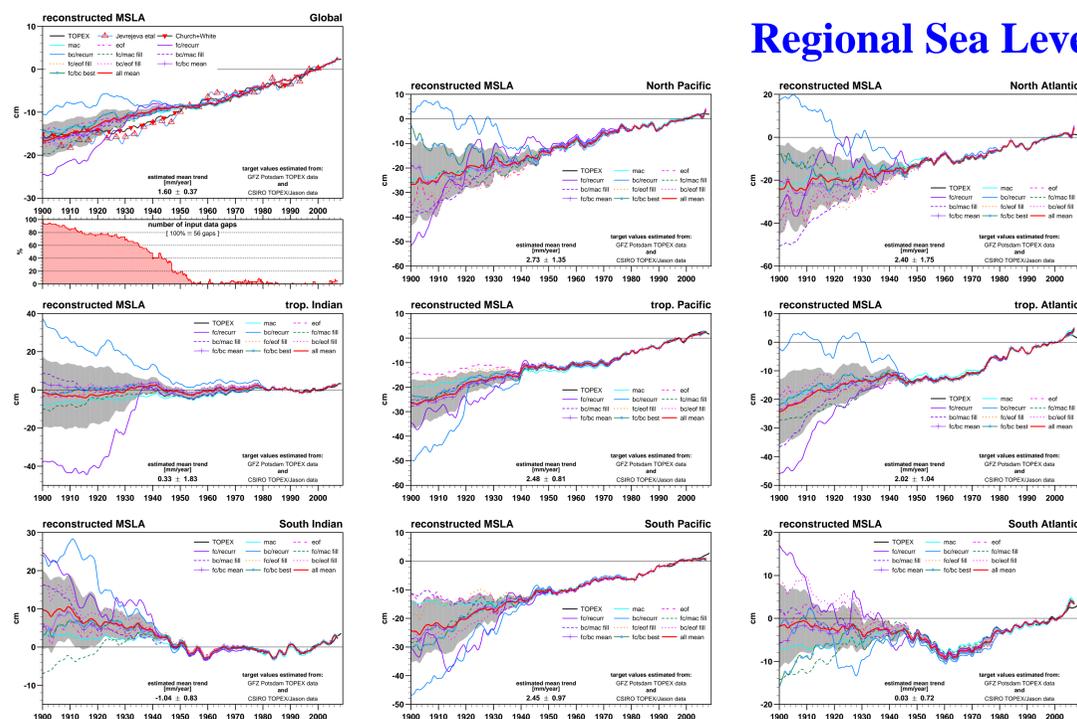
⇐ **Fig. 5:** *RMSLA for the different ocean regions (color shaded areas in Fig. 2,left). The solutions from all filling cases (see Table above) are shown. The thick red curves (all mean) give the ensemble mean and the grey shading the corresponding standard deviation. For the global ocean (top of left column) the results from Church and White (2006) and from Jevrejeva et al (2006) are included.*

NOTE: *All curves are smoothed before plotting to eliminate the annual cycle!*

A single BPN (56 input, 112 hidden and 8 output neurons) is trained to compute the monthly differences for the eight RMSLA's from the tide gauge values. The training includes the constraint that the area weighted sum of the regional must coincide with the given global. The corresponding target values for the training are derived from the CSIRO or/and the GFZ data.

Finally a recall is done for all data gap filling alternatives (see above). The resulting RMSLA's for the "CSIRO+GFZ" training are displayed in Fig. 5. The results for the single ensemble members are relatively insensitive to what is filled into the tide gauge data gaps as long as the number of gaps does not exceed 20%, i.e. beyond 1950. Before this date the sensitivity is higher for the regional than for the global mean sea level.

The linear trends are summarized in the table (ensemble mean and standard deviation). For the regional trends we find a strong dependence to the target data chosen for training, while the global trend shows less dependence and fits well to the estimates given by Church and White (2006) or Jevrejeva et al (2006): 1.7±0.3 mm/year and 1.8 mm/year, respectively.

regional mean sea level trend [mm/year]

region	CSIRO only	GFZ only	CSIRO+GFZ
	target values for training taken from		
TIND	1.01±1.36	0.58±1.61	0.33±1.83
SIND	-0.12±0.51	-0.17±0.38	-1.04±0.83
NPAC	1.27±1.24	2.07±1.49	2.73±1.35
TPAC	2.56±0.99	1.79±0.59	2.48±0.81
SPAC	1.95±0.67	2.64±0.97	2.45±0.97
NATL	1.57±1.54	1.53±1.73	2.40±1.75
TATL	1.59±0.77	0.34±0.93	2.02±1.04
SATL	0.57±0.80	-0.75±1.12	0.03±0.72
global= Σregion	1.42±0.39	1.29±0.31	1.60±0.37

Corresponding e-mail adresses:

Manfred.Wenzel@awi.de   Jens.Schroeter@awi.de

References:
Freeman and Skapura (1991) Neural Networks – Algorithms, Applications and Programming Techniques, *Addison-Wesley Publishing Company*, Reading, MA
Church et al (2004) Estimates of the Regional Distribution of Sea Level Rise over the 1950 to 2000 Period. *Journal of Climate*, 17, 2609-2625.
Church and White (2006) A 20th century acceleration in global sea-level rise, *Geophys. Res. Lett.*, 33, L01602, doi: 10.1029/2005GL024826
Jevrejeva et al (2006) Nonlinear trends and multiyear cycles in sea level records, *J. Geophys. Res.*, 111, C09012, doi:10.1029/2005JC003229