Computer-aided quality assurance of high-resolution digitized historic tide-gauge records

Dr. Hartmut Hein, Mai, S., C. Blasi, Barjenbruch, U.
Federal Institute of Hydrology, Koblenz, Germany, hein@bafg.de
Outline

- Why should we know historic observations?
- Test data set
- Outlier search
- Closing gaps
- Conclusion
KLIWAS Objectives

With diversified Approaches to a holistic Analysis

- start at future greenhouse gas emission scenarios, assess projections of their **impacts on regional climate for the next 100 years** and investigate the impacts of these changes on the waterways.
- focuses on all crucial climate change parameters of inland and coastal waterways and canals – to maintain navigation safe. For instance, changes on **currents, tides and sea states, concentrations of salt** and suspended loads, water quality and ecology.
- 17 research disciplines work – and network – in 30 projects.
- develop ways to adapt that are ecologically and economically state-of-the-art.
Climate is the characteristic frequency distribution of local conditions and processes for a sufficient period of time. This period reflects the probability density of states and processes of the typical conditions in the region.

Sea level is not level. So, sea level rise includes more than just rise of the mean sea level. The rise is overlaid by strong variability on all temporal and spatial scales. In the coastal ocean or estuaries the change of tidal characteristics means vulnerability. > Take care of the seas
Long Term Processes & Climate Change

Climate change
Rise (Mean) Sea Level

Coastal long term processes & tidal characteristics

Uncertainties Scales Probabilities Processes System Climate

Computer-aided quality assurance of high-resolution digitized historic tide-gauge records
hein@bafg.de, Hydro2012
Historic secular changes of the MSL (since 1900)

Global rise (Church & White, 2006) vs. acceleration

Trend in M2 per decade

"Changes of tidal amplitude and/or phase have taken place over large scales"

Trend in S1 per decade

What can we do to estimate regional changes?

1. Numerical modeling
   Model chain (another talk)

2. Use tide gauge data, if we have some.

However, not all tide gauge data are in digital form:

GLOSS Group of Experts (7-11 November, 2011, Paris):
Rescue of tide gauge data which currently stored in non-computer forms (sheets, tabulations, etc.)
(also: Circular Letter, IHB File No S3/2705).

- Digitalization of tide gauge data in paper form
- The crucial challenge is situated in the quality control of the data
- Generally, these data are extensive
- Automatic methods, Computer-aided quality assurance (CAQ)
- Identify failed digitization, data gaps or distortions of water levels
- Close Gaps
Test data set:

O(100) years
>3000 paper sheets

Resolution: 10 minutes
⇒ > 5 million data points

Still 10000 (?) years to digitize

<table>
<thead>
<tr>
<th>Tide gauge</th>
<th>Location (DHDN Bessel 1841)</th>
<th>Digitalized Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Borkum Südstrand</td>
<td>5.938.584,00; 2.543.850,72</td>
<td>1949 - 1985</td>
</tr>
<tr>
<td>Neue Seeschleuse</td>
<td>5.912.318,32; 2.579.065,03</td>
<td>1949 - 1982</td>
</tr>
<tr>
<td>Mellumplate</td>
<td>5.960.449,00; 3.440.237,00</td>
<td>1964 - 1990</td>
</tr>
<tr>
<td>Alter Leuchtturm</td>
<td>5.934.916,00; 3.471.446,00</td>
<td>1965 - 1972</td>
</tr>
</tbody>
</table>
Computer-aided quality assurance of high-resolution digitized historic tide-gauge records

KLIWAS
Impacts of Climate Change on Waterways and Navigation in Germany

hein@bafg.de, Hydro2012
Computer-aided quality assurance of high-resolution digitized historic tide-gauge records

hein@bafg.de, Hydro2012
Outlier test

- Running standard deviation
- Modification of Thomson Rule
- Strategy: Find outlier -> remove, iterate
- Philosophy: Better to have more gabs, than uncertain data
Gab closing methods

1. Single data set:
   Use the periodicity inside the signal

2. Network of tide gauges
   Test the information content between signals
   If ok, stochastic reconstruction of gabs
Lomb-Scargle Periodogram

\[ P(f) = \frac{1}{2s^2} \left( \frac{\sum_i \left(y_i - \bar{y}\right) \cos 2\pi f (t_i - \tau)}{\sum_i \cos^2 2\pi f (t_i - \tau)} \right)^2 \left( \frac{\sum_i \left(y_i - \bar{y}\right) \sin 2\pi f (t_i - \tau)}{\sum_i \sin^2 2\pi f (t_i - \tau)} \right)^2 \]

Time constant of the frequency:

\[ \tan(4\pi \tau) = \frac{\sum \sin(4\pi f t_i)}{\sum \cos(4\pi f t_i)} \]
Historic insertion by tide gauge operator
Fuzzy Logic

Fuzzy inference system simulates the behavior of the sea level system by means of "if-then" rules of correlations in the different gauge data.

From actual observations we know: More stations, which are (non-linear) correlated, yield better reconstructions.
Uncertainty of the reconstruction: 2 - 8 cm
Conclusions

- High resolution long term data sets are necessary to understand the rise of the sea levels and regional processes.
- Modelling or use historic tide gauge data
- Historic tide gauge data may not be digital
- Digitalisation is possible, quality assurance is the challenge
- Outlier search
- For non redundant data Lomb-Scargle works quite well
- For digitized networks of tide gauges: Fuzzy logic
- Future research: (Semi-)automatic digitalisation

Optimisation of Fuzzy networks
Thank you for your attention.

Go out and digitize your archives!

Hartmut Hein
Christoph Blasi
Stephan Mai
Ulrich Barjenbruch
Department M1
German Federal Institute
of Hydrology (BfG)
Am Mainzer Tor 1
56068 Koblenz, Germany
+49-261-1306-5326 e-mail
hein@bafg.de
www.bafg.de

Computer-aided quality assurance of high-resolution digitized historic tide-gauge records