Session 2:
Data challenges & Lessons learned

Training, Dakar, Senegal, July 2016
Trainer: Harald Kling
Pöyry, Hydro Consulting, Hydropower, Austria
Data challenges & Lessons learned

Overview

- Geo-referencing in GIS
  Gauges, Dams

- Data availability in different periods
  Observed discharge, rainfall data sets

- Data processing
  Software issue

- Lessons learned

- Group discussion
Discharge measurement

Some gauges in Nigeria

- Work of gauge readers is extremely important!
- Water level measured once or thrice a day
- Often manual readings with hand-written records
- Rating curve required to convert water level (m) to discharge (m$^3$/s)
Discharge measurement in West Africa

410 discharge gauges available for this study

- 360 gauges obtained from GRDC
- 50 gauges obtained from
  - River Basin Organizations
  - National Hydrological Services
  - JICA

- Hardly any gauges located at small rivers that are suitable for small-scale hydropower.

- Gauge data cover different observation periods.

- Inaccurate geo-referencing major problem before data can be used for modelling study
GRDC discharge gauges

Manual geo-referencing required in GIS

- Requirement for further GIS work
  - Gauges must be located at river network

- Information used
  - River name
  - Gauge name (Where is this village?)
  - Satellite image (Where is nearest bridge or river access?)
  - Area reported vs. area computed.
  - Country
  - Sierem data base (inaccurate!)
  - Reports (Google search)

- Typical errors
  - Insufficient decimal places for latitude & longitude, e.g. lat = 7.5°
  - Inaccurate coordinates
  - Typing error
e.g. lat = 7.531 -> lat = 8.531
GRDC discharge gauges

Manual geo-referencing required in GIS
Existing hydropower plants layer

Geo-referencing example: Bui HPP in Ghana (recently constructed)
Temporal data availability

GRDC discharge data (daily)

Notes:
Only 250 out of 361 gauges displayed. Missing data not visualized (data gaps!)
Data quality

Questionable discharge data at some gauges

<table>
<thead>
<tr>
<th>Année</th>
<th>Jan</th>
<th>Fév</th>
<th>Mar</th>
<th>Avr</th>
<th>Mai</th>
<th>Juin</th>
<th>Juil</th>
<th>Aou</th>
<th>Août</th>
<th>Sep</th>
<th>Oct</th>
<th>Nov</th>
<th>Déc</th>
<th>Annuel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1964</td>
<td></td>
<td></td>
<td>19.75</td>
<td>63.34</td>
<td>(35.95)</td>
<td>196.08</td>
<td>233.81</td>
<td>(271.37)</td>
<td>(291.51)</td>
<td>(114.48)</td>
<td>52.98</td>
<td>(148.52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1965</td>
<td>26.18</td>
<td>59.65</td>
<td>28.03</td>
<td>47.91</td>
<td>50.67</td>
<td>137.24</td>
<td>258.92</td>
<td>258.6</td>
<td>290.8</td>
<td>212.9</td>
<td>85.62</td>
<td>31.03</td>
<td>126.02</td>
<td></td>
</tr>
<tr>
<td>1967</td>
<td>21.92</td>
<td>13.55</td>
<td>13.24</td>
<td>48.57</td>
<td>102.94</td>
<td>100.86</td>
<td>256.55</td>
<td>273.39</td>
<td>(342.73)</td>
<td>256.73</td>
<td>42.23</td>
<td>(135.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1968</td>
<td>23.14</td>
<td>23.9</td>
<td>24.66</td>
<td>25.43</td>
<td>26.2</td>
<td>26.98</td>
<td>27.75</td>
<td>29.53</td>
<td>29.3</td>
<td>30.08</td>
<td>30.65</td>
<td>31.62</td>
<td>27.36</td>
<td></td>
</tr>
<tr>
<td>1969</td>
<td>(18.8)</td>
<td>13.55</td>
<td>(37.53)</td>
<td>10.79</td>
<td>258.63</td>
<td>269.55</td>
<td>275.68</td>
<td>235.5</td>
<td>92.76</td>
<td>(47.58)</td>
<td>(152.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1970</td>
<td>(23.99)</td>
<td>(18.68)</td>
<td>(21.07)</td>
<td>(26.82)</td>
<td>58.28</td>
<td>98.77</td>
<td>149.11</td>
<td>197.4</td>
<td>201.88</td>
<td>160.44</td>
<td>101.42</td>
<td>36.37</td>
<td>(103.02)</td>
<td></td>
</tr>
<tr>
<td>1971</td>
<td>(21.54)</td>
<td>(25.22)</td>
<td>30.06</td>
<td>40.47</td>
<td>43</td>
<td>79.04</td>
<td>207.37</td>
<td>185.38</td>
<td>(186.87)</td>
<td>65.78</td>
<td>33.21</td>
<td>(93.66)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1972</td>
<td>(22.12)</td>
<td>(17.91)</td>
<td>27.58</td>
<td>55.71</td>
<td>112.89</td>
<td>161.68</td>
<td>200.87</td>
<td>212.1</td>
<td>165.94</td>
<td>74.97</td>
<td>30.45</td>
<td>(108.24)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1973</td>
<td>(27.07)</td>
<td>(13.2)</td>
<td>15.57</td>
<td>25.7</td>
<td>48.86</td>
<td>62.77</td>
<td>98.47</td>
<td>143.35</td>
<td>199.42</td>
<td>125.84</td>
<td>75.52</td>
<td>(84.19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1974</td>
<td>16.01</td>
<td>11.92</td>
<td>12.56</td>
<td>42.74</td>
<td>42.72</td>
<td>102.94</td>
<td>180.27</td>
<td>196.15</td>
<td>233.16</td>
<td>214.67</td>
<td>123.35</td>
<td>46.13</td>
<td>102.54</td>
<td></td>
</tr>
<tr>
<td>1975</td>
<td>23.86</td>
<td>(21.39)</td>
<td>(29.73)</td>
<td>45.45</td>
<td>51.01</td>
<td>83.42</td>
<td>142.56</td>
<td>155.13</td>
<td>233.41</td>
<td>239.34</td>
<td>95.34</td>
<td>47.58</td>
<td>(104.4)</td>
<td></td>
</tr>
<tr>
<td>1976</td>
<td>24.32</td>
<td>24.75</td>
<td>35.65</td>
<td>52.59</td>
<td>50.47</td>
<td>81.39</td>
<td>187.99</td>
<td>227.34</td>
<td>238.23</td>
<td>221.12</td>
<td>129.98</td>
<td>48.97</td>
<td>110.48</td>
<td></td>
</tr>
<tr>
<td>1977</td>
<td>27.05</td>
<td>(18.53)</td>
<td>(9.25)</td>
<td>16.43</td>
<td>41.26</td>
<td>98.92</td>
<td>208.15</td>
<td>186.11</td>
<td>251.36</td>
<td>197.38</td>
<td>69</td>
<td>31.02</td>
<td>(107.24)</td>
<td></td>
</tr>
<tr>
<td>1978</td>
<td>18.49</td>
<td>12.37</td>
<td>20.1</td>
<td>47</td>
<td>44.08</td>
<td>123.15</td>
<td>177.52</td>
<td>207.9</td>
<td>237.07</td>
<td>206.31</td>
<td>94.06</td>
<td>38.4</td>
<td>102.69</td>
<td></td>
</tr>
<tr>
<td>1979</td>
<td>20.58</td>
<td>15.86</td>
<td>18.9</td>
<td>28.49</td>
<td>67.17</td>
<td>115.4</td>
<td>192.46</td>
<td>263.85</td>
<td>194.78</td>
<td>172.8</td>
<td>(147.18)</td>
<td>(120.47)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td>20.19</td>
<td>14.53</td>
<td>46.32</td>
<td>32.94</td>
<td>29.91</td>
<td>110.59</td>
<td>126.65</td>
<td>124.3</td>
<td>239.45</td>
<td>186.39</td>
<td>100.59</td>
<td>42.2</td>
<td>(103.68)</td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td>30.87</td>
<td>30.78</td>
<td>30.7</td>
<td>30.52</td>
<td>30.54</td>
<td>30.45</td>
<td>30.37</td>
<td>30.28</td>
<td>30.2</td>
<td>30.12</td>
<td>30.03</td>
<td>29.95</td>
<td>30.41</td>
<td></td>
</tr>
<tr>
<td>1982</td>
<td>29.87</td>
<td>29.78</td>
<td>29.7</td>
<td>29.52</td>
<td>29.64</td>
<td>29.45</td>
<td>29.37</td>
<td>29.28</td>
<td>29.2</td>
<td>29.12</td>
<td>29.03</td>
<td>28.95</td>
<td>29.41</td>
<td></td>
</tr>
<tr>
<td>1984</td>
<td>27.87</td>
<td>27.78</td>
<td>27.7</td>
<td>27.52</td>
<td>27.53</td>
<td>27.45</td>
<td>27.37</td>
<td>27.28</td>
<td>27.2</td>
<td>27.12</td>
<td>27.03</td>
<td>26.95</td>
<td>27.41</td>
<td></td>
</tr>
<tr>
<td>1988</td>
<td>23.86</td>
<td>23.78</td>
<td>23.7</td>
<td>23.62</td>
<td>23.53</td>
<td>23.45</td>
<td>23.37</td>
<td>23.28</td>
<td>23.2</td>
<td>23.11</td>
<td>23.03</td>
<td>22.95</td>
<td>23.41</td>
<td></td>
</tr>
</tbody>
</table>

data source: NBA
Data quality

Questionable discharge data at some gauges
Data quality

Example for data gaps

“Observed” discharge data have to be treated with caution!
Pre-processing of observed discharge data

**Gap-filling** (yellow shading)

- Discharge data of all 410 gauges were manually checked. Implausible data removed.
- Manual gap-filling of monthly data to enable computation of annual means.
- At gauges in semi-arid regions often missing records in dry season. Staff gauge readings only during wet season.
Observed discharge data

Availability of annual data at 410 gauges after pre-processing

- Implausible data removed
- Filling of short data gaps
Precipitation data in Africa

Data sources

- Individual station measurements
  - Data collection, gap filling, spatial mapping would require huge work effort
  - Not considered in this study

- Gridded station based data
  - GPCC: Global Precipitation Climatology Centre

- Satellite based data
  - TRMM: Tropical Rainfall Measuring Mission
    - TRMM 3B42: “High” quality product, “corrected” with ground measurements
    - TRMM 3B42RT: Real-time product, no ground measurements
  - RFE Africa: Rainfall Estimator (FEWS-NET,)

- Various other products not considered
  - GTS CPC
  - RFE ARC
Precipitation data
Spatial and temporal resolution

<table>
<thead>
<tr>
<th>Product</th>
<th>Spatial resolution</th>
<th>Temporal resolution</th>
<th>Availability</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPCC</td>
<td>0.5 x 0.5 °</td>
<td>monthly</td>
<td>1901 – 2009</td>
<td>Coarse resolution, best reliability (especially 1950-1990)</td>
</tr>
<tr>
<td>TRMM 3B42</td>
<td>0.25 x 0.25 °</td>
<td>daily</td>
<td>1998 – 2014</td>
<td></td>
</tr>
<tr>
<td>TRMM 3B42RT</td>
<td>0.25 x 0.25 °</td>
<td>daily (3h)</td>
<td>2000 – now</td>
<td>Real-time product.</td>
</tr>
<tr>
<td>RFE Africa</td>
<td>0.1 x 0.1 °</td>
<td>daily</td>
<td>2001 – now</td>
<td>Finest resolution, quality problems in some regions</td>
</tr>
</tbody>
</table>
Availability of precipitation data

Period coverage

GPCC stations in Niger basin

Satellite data

TRMM 3B42RT
TRMM 3B42
RFE
Annual precipitation

Long-term average of different products

GPCC 1961-1990

TRMM3B42 1998-2013

RFE 2001-2013

Precip (mm/y)
- 0.1 - 250
- 250.1 - 500
- 500.1 - 750
- 750.1 - 1,000
- 1,000.1 - 1,250
- 1,250.1 - 1,500
- 1,500.1 - 2,000
- 2,000.1 - 2,500
- 2,500.1 - 3,000
- 3,000.1 - 4,000
- 4,000.1 - 12,000

COPYRIGHT@PÖYRY

GIS Hydropower Resources Mapping for ECOWAS Region
July 2016
Potential evapotranspiration

Data sources

- CRU
  - Climate Research Unit: University of East Anglia (UK)
  - Monthly global grids 1901-2009
  - Penman-Monteith method
  - Air temperature also available

- CROPWAT & CLIMWAT
  - provided by FAO
  - Station based
  - Long-term monthly averages
  - Penman-Monteith method
Definition of common reference period

Data sources & availability

Wet & dry periods

Good data availability

Recent years

Niger River discharge

Availability GRDC discharge data

Availability precipitation data

Wet & dry periods

Availability GRDC discharge data

Availability precipitation data

Good data availability

Recent years
Definition of common reference period

- General considerations
  - Should be long enough to smooth out variability of individual years.
  - Should be well accepted by stakeholders.
  - Should have good data availability / reliability.

- 1961-1990
  - Calibration of water balance model.
  - Good availability of observed discharge data.
  - High number of stations available for GPCC precipitation data.
  - Includes prolonged drought of the 1980s.
  - 1990 was 26 years ago, acceptance by stakeholders?

- 1998-2014
  - Adopted reference period for final results.
  - Poor data availability for observed discharge data.
  - GPCC precipitation data not reliable / available.
  - Satellite precipitation data available.
  - Since 1998 relatively stable meteorological conditions (moderately wet compared to last 100 years).
Data processing

Software issues

- **Standard GIS software** ArcGIS 9.2, 10.0, QGIS
  
  Used extensively in this study.
  
  Frequently crashed during processing (overflow of data).
  
  Ancient ArcView 3.1 more stable for some tasks.

- **Advanced data analysis / modelling**
  
  Higher performance (faster, no crashes) with tools outside GIS.
  
  gdal, shell scripts, python (slow), Fortran (fast), cdo
  
  Good programming skills required.

- **Meteorological data**
  
  GPCC, TRMM, RFE, climate model data in specific formats (ASCII, binary, NetCDF)
  
  Processing of time-series in GIS is not feasible.
  
  Instead use Fortran, cdo, etc.
Lessons learned

- Correct geo-referencing is highly time consuming due to lack of accurate information.
  - 410 discharge gauges
  - 91 existing hydropower plants

- Observed discharge data:
  - Several gauges appear to be affected by severe bias, especially after 1990. Outdated rating curve?
  - Gap filling is highly time consuming, but required to enable computation of annual means.

- There are large differences in meteorological data sets
  - Precipitation: GPCC & TRMM vs. RFE
  - Potential evapotranspiration: CRU, E2O, Climwat

- The period 1961-1990 has best data availability, but includes drought of 1980s. 1998-2014 is a better reference period for assessing the “current” hydropower potential.

- Implementation of water balance model in GIS failed, due to too slow computation time. Alternative Fortran model enabled fast execution required for:
  - Time-series simulation
  - Model calibration (many repeated model runs)
  - Climate change simulations (60 model runs)
Lessons learned

Very valuable data

- **Hydrosheds**
  - Flow direction grids
  - Digital elevation model (unconditioned)

- **Rainfall data**
  - Tropical Rainfall Measuring Mission (TRMM)
  - Global Precipitation Climatology Center (GPCC)

- **Discharge data**
  - River Basin Organizations
  - National Hydrological Services
  - Global Runoff Data Center (GRDC)

- **Gauge readers**
  - Without them we would not have field information!

Many datasets available. Use them!
Group discussion

Data challenges in your country

• What are the key challenges for hydro-meteorological data in your country?
  – Sufficient funding for continuous field measurements?
  – Institutional challenges?
  – Personal experience?

• Data sharing policy?
  – Whom to contact to obtain observed discharge data?
  – Are data free or is a service charge required?
  – How fast are the data delivered?
  – Online data repositories?

• Reliability of data?
  – Sufficient number of rainfall stations?
  – Sufficient number of streamflow gauges?
  – How often are streamflow rating curves updated?

• Are global datasets used in your country?
  – Rainfall: GPCC, TRMM, RFE
  – Discharge: GRDC