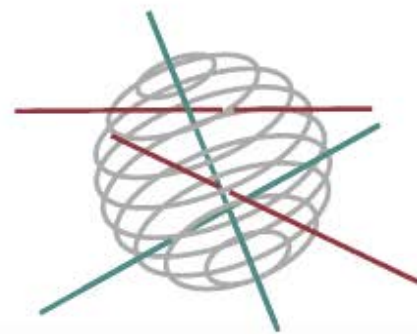# SSD

SCIENCE FOR A SUSTAINABLE DEVELOPMENT

CHOLERA OUTBREAKS AT LAKE TANGANYIKA INDUCED BY CLIMATE CHANGE?

«CHOLTIC»

Plisnier, P.-D., N. Poncelet, C. Cocquyt, H. De Boeck, D. Bompangue, J. Naithani, J. Jacobs, R. Piarroux, S. Moore, P. Giraudoux, D. Batumbo, D. Mushagalusa, L. Makasa, E. Deleersnijder, I. Tomazic, Y. Cornet

ENERGY

TRANSPORT AND MOBILITY

AGRO-FOOD
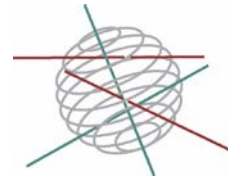
HEALTH AND ENVIRONMENT

CLIMATE

BIODIVERSITY

ATMOSPHERE AND TERRESTRIAL AND MARINE ECOSYSTEMS

TRANSVERSAL ACTIONS

.be

SCIENCE FOR A SUSTAINABLE DEVELOPMENT (SSD)

Climate - Biodiversity

FINAL REPORT

Cholera outbreaks at Lake Tanganyika induced by Climate Change?

"CHOLTIC"

SD/AR/04A

Promotors

Royal Museum for Central Africa
> Pierre-Denis Plisnier, Earth Sciences Department
> Leuvensesteenweg 13 - 3080 Tervuren

University of Liège
> Yves Cornet, Geography Department
> allée du 6 Août 17, B5 - 4000 Liège

Botanic Garden Meise
> Christine Cocquyt, Bryophyta and Thallophyta Department
> Nieuwelaan 38 - 1860 Meise

Institute of Tropical Medicine
> Jan Jacobs, Tropical Laboratory Medicine
> Nationalestraat 155 - 2000 Antwerpen

Catholic University of Louvain
> Eric Deleersnijder, Institute of Mechanics, Materials and Civil Engineering
> Avenue G. Lemaître 4 - 1348 Louvain-la-Neuve

Authors

Pierre-Denis Plisnier[1], Nadia Poncelet[2], Christine Cocquyt[3], Hilde De Boeck[4], Didier Bompangue[5], Jaya Naithani[6], Jan Jacobs[4], Renaud Piarroux[7], Sandy Moore[7], Patrick Giraudoux[8], Doudou Batumbo[5], Déo Mushagalusa[9], Lawrence Makasa[10], Eric Deleersnijder[6], Igor Tomazic[2] and Yves Cornet[2].

1. Royal Museum for Central Africa, 2. Université de Liège, 3. Botanic Garden Meise, 4. Institute of Tropical Medicine, 5. Ministère de la Santé Publique (DRC), 6. Université Catholique de Louvain, 7. Université de Marseille (France), 8. Université de Franche-Comté (France), 9. Centre de Recherches en Hydrobiologie (DRC), 8. Department of Fisheries (Zambia)

Plisnier, P.-D., Poncelet, N., Cocquyt, C., De Boeck, H. , Bompangue, D. , Naithani, J. , Jacobs, J., Piarroux, R., Moore, S., Giraudoux P., Batumbo, D. , Mushagalusa, D., Makasa, L., Deleersnijder, E., Tomazic, I., Cornet, Y. 2015. ***Cholera outbreaks at Lake Tanganyika induced by Climate Change? - "CHOLTIC".*** Final Report. Brussels: Belgian Science Policy 2015 - 117 p. (Research Programme Science for a Sustainable Development).

TABLE OF CONTENT

SUMMARY

A. Context

Cholera is one of the deadliest diseases in Africa. It reappeared in the area of the African Rift in the late 70's. The African Rift has been highlighted as major area of cholera propagation (Bompangue et al., 2008a). A link between cholera, phytoplankton blooms and copepod zooplankton has been demonstrated in Asia (Colwell et al., 1996). The African Great Lakes have been suspected to play a role as a reservoir of the bacteria *V. cholerae*, while human infection and movement are probably involved in the propagation of the disease inland.

B. Objectives

The objectives of the CHOLTIC project are to better understand the environmental conditions that trigger cholera outbreaks in the Lake Tanganyika region via an interdisciplinary study including the following aspects:

(1) *In situ* monitoring of meteorology, limnology, phytoplankton, zooplankton, fish abundance and bacteriology during a period of three years and collaboration with DRC health authorities and epidemiology researchers.

(2) Remote sensing to produce time series of daily images of Chl-a and lake-surface temperatures for the period 2000-2014.

(3) Eco-hydrological modeling to investigate links between climate, nutrient mixing and variable abundance of different planktonic groups.

(4) Microbiological monitoring and confirmation.

(5) Genetic characterization by mass spectra identification of cholera strains.

(6) Data analysis of spatio-temporal relationships between environmental factors and health data.

C. Conclusions

Meteorological monitoring at Uvira and Mpulungu recorded air temperature, atmospheric pressure, rainfall levels, relative humidity, wind speed, wind direction and solar radiation.

Limnological monitoring was implemented from 2011 to 2014 at five sites (three coastal and two pelagic). Parameters included T°, pH, dissolved oxygen, conductivity, turbidity, transparency and chl-a concentrations. Marked short-term variability was observed in relation to internal waves and pulse of hypolimnion nutrient. Isotherms in deep waters were found to be good indicators of pulsed fluctuations of parameters.

Fish relative abundance of principal species (*Stolothrissa tanganicae*, *Limnothrissa miodon* and *Lates stappersii*) was monitored in two sites. At Uvira in DRC (DR Congo), the main fish catches were the small clupeids: *Stolothrissa tanganicae* (86%). Peaks and dampening of those fish are generally observed between September and December every year during a secondary upwelling period. Deep isotherms could be indicators of planktonic and fish population changes as well as cholera. At Mpulungu (Zambia), the timing of clupeid increased catches was associated with seasonal changes and planktonic blooms particularly in April/May.

Phytoplankton dynamics were assessed at the same sites and periods where the limnology tests were performed. The results include a seasonal succession in the abundance of major algal groups, genera and, where possible, up to the species level. The changes from a cyanobacteria-chrysophyte-chlorophyte community of 1975 (Hecky & Kling 1987) to a cyanobacteria-chlorophyte-diatom community, as observed near Kigoma (Cocquyt & Vyverman 2005), was confirmed for the north. However, dominance of dinophytes was also observed at Uvira during the CHOLTIC survey. At the coast, dinophytes were dominant at the end of the wet season. The cyanobacteria-chlorophyte-diatom community was then replaced by a dinophyte-chlorophyte-diatom community. In the south of the lake, a cyanobacteria-chlorophyte-diatom pelagic community was characteristic, while dinophyte levels increased during the wet season. Interannual changes in phytoplankton biomass were mainly due to cyanobacteria in the north and diatoms in the south. Although no significant correlations were found between phytoplankton composition and cholera cases in the North, significant statistical relationships were observed between total phytoplankton abundance (chl-a) and cholera in the integrated analysis.

Zooplankton and particularly copepods are important food components for pelagic fish species in Lake Tanganyika. In this region, copepod populations are mainly represented by one species of *Calanoida*, four species of *Cyclopoida* and *Harpacticoida*. Copepods have been identified in oceans as possible reservoirs of *V. cholerae* bacteria (Colwell et al., 1996). Therefore, the CHOLTIC project involved the monitoring of copepod levels. Copepod abundance was higher in the pelagic environment. Seasonal variations showed peaks in April/May and September/October for the most abundant species of copepods. *Tropocyclops tenellus* was numerically dominant throughout the study in all post-nauplii stages.

Remote sensing-allowed producing time series of Lake Tanganyika surface water temperatures and ocean color products using MODIS data. This assessment covered the period 2002-2014, extending the period covered by previous BELSPO projects by eight years and improving the calibration and validation methods. Raster

georeferenced files and lake snapshots of daily- and weekly-aggregated data as well as spatio-temporally aggregated time series in table format were delivered for the entire lake and 12 ecoregions. We briefly present the calibration and validation procedures implemented to derive chl-a concentrations and lake-surface temperatures using in situ data and ARC Lake datasets. Landcover maps of the three sites were also produced from Landsat TM, ETM+ and OLI images. Diachronic classification was applied for the Mpulungu area, and monodate was applied for Uvira and Kalemie.

Results of the RS component identified various blooms in the lakes during the 2002-2014 period. Using these results, we were able to analyze the longest available cholera-case time series.

The **ecological model** developed during the earlier projects on Lake Tanganyika (CLIMLAKE and CLIMFISH) was successfully modified to include the dominant phytoplankton groups present in the lake considered important for the present CHOLTIC project. The model is closed by the predation pressure of planktivorous fish. Simulations were performed using the wind and solar radiation data from the National Centers for Environmental Protection (NCEP) reanalysis 2. The phytoplankton production and their succession in the lake is governed by the availability of nutrients in the surface layer, its depth and light. Chlorophytes, which require low nutrients and high light, dominate in the northern proximity of the lake while the diatoms and cyanobacteria, which require high nutrients and can survive in low light conditions, dominate in the southern basin. The years with stronger wind are accompanied by an increased biomass of diatoms and cyanobacteria, while low wind years have higher biomass of chlorophytes. Various possible climatic scenarios were studied by changing the surface layer depth, its temperature and the wind stress. Different phytoplankton groups responded differently to the changes in the model forcing.

Epidemiological monitoring of cholera by teams involved in the CHOLTIC project (DLM/DRC and UFC/France) was implemented as foreseen. In the context of epidemiological cholera data collected in the DRC, particular attention was focused on the South-Kivu and Katanga Provinces near Lake Tanganyika. Among the databases of the 515 DRC health zones, the project took advantage of the development of fine-scale databases for two target health zones: Kalemie and Uvira (2008 to 2014). Using these data, we constructed health zone attack rate maps. Time series were elaborated for these target health zones and main zones along the shores of Lake Tanganyika (Nundu, Fizi, Nyemba, Kansimba and Moba). The decompositions of time series were carried out for the two target health zones and two health areas with an endemic profile. Predominance of the most significant attack rates during rainy seasons was confirmed.

Microbiological screening was successfully implemented in Zambia, and for the DRC, clinical samples were obtained via the national surveillance system. A total of 47 environmental samples from water, plankton and fish were collected in Mpulungu, Zambia. Between August 2012 and October 2014, the results indicated seasonal fecal contamination of Lake Tanganyika surface water. Furthermore, four environmental *Vibrio cholerae* non-O1 strains were isolated from these environmental samples. Additionally, clinical samples were collected in Mpulungu (27 *V. cholerae* O1, Inaba, during an outbreak in 2012), Katanga, DRC (73 *V. cholerae* O1, Inaba), and South-Kivu, DRC (28 *V. cholerae* O1 isolates: Inaba (3) and Ogawa (25)). The implementation of microbiological analysis of water and clinical samples in field laboratories was challenging and complex due to insecurity and instability in the Lake Tanganyika region and high staff turn-over at the field sites. Nevertheless, we were able to reinforce the capacities and collect a set of environmental and clinical *V. cholerae* isolates for further molecular analysis.

Genetic analyses of 531 *Vibrio cholerae* isolates from patients and environmental samples from the DRC and Zambia were analyzed via MLVA (Multiple-Locus Variable number tandem repeat Analysis). The results were then compared to those of a panel of *V. cholerae* isolates from Guinea, Ghana and Togo. Overall, we identified 118 unique MLVA haplotypes. MLVA typing revealed the short-term divergence and microevolution of these *V. cholerae* populations, thereby providing insight into the dynamics of cholera outbreaks. The results revealed strong geographical clustering. Isolates from the African Great Lakes Region formed a closely related group, while West African isolates constituted a separate cluster. Interestingly, certain strains in the DRC have circulated in the region over a period of several years, occasionally giving rise to expansive epidemics. The six environmental isolates in our panel were unrelated to the epidemic isolates.

Data analysis

Time series (TS) of data were analyzed for the period 2002 to 2014 and the project-monitoring period (2011-2014). The cholera epidemiological data were analyzed as a dependent variable in relation to environmental variables. All statistical analysis applied aimed to take into account delayed responses, autocorrelation, non-stationarity and the rare event nature of the epidemiological dependent variable. Binomial, ARIMA-remainders and ARIMA-modeled time series analysis enabled us to observe the following:
  - Spatiotemporal interactions were identified between cholera outbreaks and chl-a concentrations in the lake from remote sensing time series (aggregated by week and by ecoregion).
  - Two main cholera periods were identified in relation with the main meteorological seasons: wet season peaks and end of wet/dry season peaks. The different wind

direction at seasonal change induces hydrodynamic fluctuations comparable to modes of development of cholera outbreaks. Damped oscillation of cholera cases was similar to those observed in various lacustrine related variables (e.g., T°, chl-a and clupeid catches), which suggests a probable lacustrine relationship with outbreaks.

- Unusual high pelagic surface temperatures appeared to have a strong and positive two-week delayed effect on cholera cases with regard to the expected number of cases.
- Rainfall levels were not found to have a direct statistical relationship. Seasonality was linked to a variety of other meteorological and environmental changes, which points to focalized attention on detailed study particularly in relation with initial cases at the beginning of outbreak conditions.

The H0 hypothesis concerning environmental relationships with cholera was thus supported by the statistical and ecological relationships at Lake Tanganyika. However, an alternative hypothesis associated with human impacts at the origin of the outbreaks may not be rejected as lacustrine and meteorological conditions could favor the development of *Vibrio cholerae*, particularly as the lake presents high pH values (>9.5) during periods of planktonic blooms. In the last 40 years, various ecological changes have taken place in the lake, which may be favorable for the establishment of *Vibrio cholerae*.

D. Support to a sustainable development policy.

CHOLTIC has enabled the establishment of a new collaboration between health and environmental stakeholder sectors. Multidisciplinary collaboration is a necessary step to understand cholera outbreaks, warn populations, provide advice, and develop appropriate measures and warning systems to decrease the transmission of epidemics. CHOLTIC has helped to reinforce capacities in three stations around the lake in the field of meteorology, limnology, plankton studies and microbiology. This involved both training and logistical aspects. Various tools continue to allow local teams to prolong the monitoring for cholera studies and other fields of investigation.

Various environmental and bacteriological field databases generated during the CHOLTIC project represented an important baseline, not only for further analysis with these data, but also for further investigations and comparisons. An ecological model has been considerably improved by detailing new ecological components, thereby enabling precise investigations of various planktonic groups associated with climate changes and its impact on the hydrodynamic and ecological characteristics of the lake. Analyses of CHOLTIC data have enabled us to identify interesting correlations

between cholera outbreaks and environmental data, which researchers may focalize on for targeted sampling design and analysis.

E. Keywords

Cholera, *Vibrio cholerae,* Lake Tanganyika, East Africa, multidisciplinary, health, environment, limnology, phytoplankton, copepods, remote sensing, hydrodynamics, ecological modeling, epidemiology, bacteriology, genetic, sustainable development.

1. INTRODUCTION

Cholera is one of the deadliest diseases in Africa. Cholera outbreaks reappeared in the area of the African Rift in the late 70's, when marked signals of climate changes were noted (Lipp et al., 2002; O'Reilly et al., 2003). Great lakes such as Lake Tanganyika are highly suspected to play a role as a reservoir of the *V. cholerae* bacteria, while human infection and movement propagate the disease inland. The African Rift has been highlighted as an area of cholera propagation (Bompangue et al., 2008).

In the Rift Lake area, environmental and epidemiological data have been collected independently between 2002 and 2006 by CLIMFISH, a BELSPO project, and by the DRC Health Department (Bompangue et al., 2008). Cross analyses have indicated a possible correlation between phytoplankton blooms and cholera outbreaks. Such a link has been shown in the Indian Ocean by Colwell et al. (1996). The comparable periodicities of blooms and cholera cases support the "lake reservoir hypothesis" that CHOLTIC has proposed to test. This interdisciplinary study, performed in collaboration with Congolese and Zambian researchers based along Lake Tanganyika, took place between 2011 and 2014 at three sites of Lake Tanganyika: Uvira and Kalemie in DRC and Mpulungu in Zambia (Figure 1).
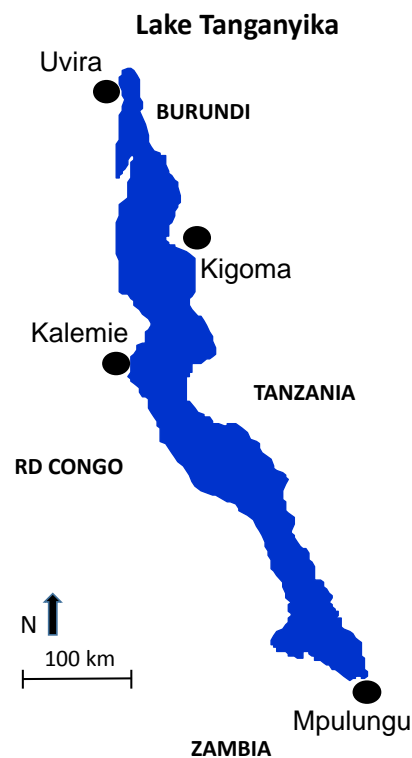

Fig. 1. Map of Lake Tanganyika indicating the CHOLTIC monitoring sites of Uvira and Kalemie in the DRC and Mpulungu in Zambia.

2. METHODOLOGY AND RESULTS

WP 1 Environmental in situ monitoring

WP 1.1 Meteorology

Meteorological data were recorded every 15 min by automatic DAVIS Vantage Pro2 weather stations at Uvira (northern end of the lake) and Mpulungu (southern end of the lake) (Figure 1). The station functioned well in Uvira during the entire monitoring period, while in Mpulungu, the station did not always provide continuous recordings. Additional recordings from a local survey conducted by the Department of Fisheries at Mpulungu however completed the meteorological database.

Average weekly air temperatures for complete years of observations (2012 and 2013) were 24.6°C in Uvira and 25.4 °C in Mpulungu. Air temperatures were slightly higher in 2013 at both stations: +0.2°C. The range of temperature was smaller at Uvira (23.3°C to 26.4°C) compared to Mpulungu (22.5°C to 28.1°C).

During a yearly cycle, the main meteorological change is linked to the rainfall seasons, which last from September to May. In the north of the lake, at Uvira, more frequent but less abundant rainfall levels were observed compared with Mpulungu (Figure 2). The total annual rainfall levels in 2012 and 2013 were 675 and 822 mm in Uvira compared with 790 and 1371 mm, respectively, at Mpulungu. A high interannual rainfall levels variability, particularly in the south, is not unusual. Relative humidity was higher on average at Uvira (77.6%) compared with Mpulungu (61.3 %).

The average atmospheric air pressure, at sea level, was very comparable at both stations (1008.4 mb in Uvira and 1008.3 mb at Mpulungu). Solar radiation was higher at Mpulungu (210 w/m^2) compared to Uvira (198 w/m^2), which was caused by increased cloudiness near the equator.

Wind speed in Uvira was 0.6 m/s on average. The data at Mpulungu were not complete, although stronger local wind speeds are probable. The position of Mpulungu at the south of the lake explains the additional effects of main southeast trade wind and offshore winds during the night ("Kapata winds"). Meanwhile, those effects are not observed at Uvira due to the position of Uvira (northwest side of the lake) and the slightly more inland position of the meteorological station compared with Mpulungu.
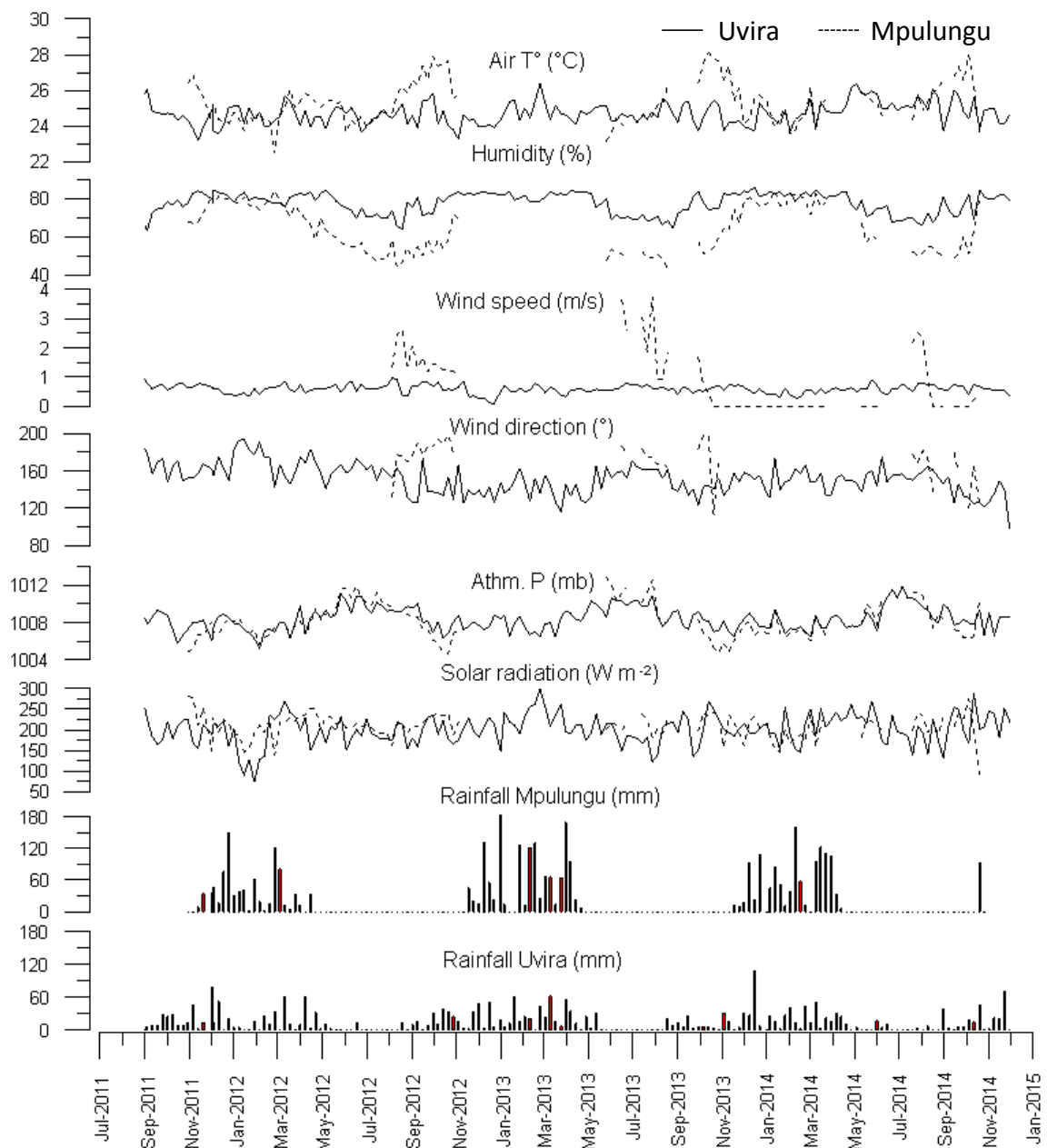
Fig. 2. Weekly meteorological conditions at Uvira (solid line) and Mpulungu (dashed line). Average air T (°C), relative humidity (%), wind speed (m/s), wind direction (°), atmospheric air pressure (mb), solar radiation (w m^{-2}), and total rainfall (mm) are indicated.

In summary, for the entire lake, two seasons were identified: a wet season from September/October to April/May with mainly northeast regional winds and warmer air temperatures compared to the dry season during the rest of the year. This season is characterized by strong southeast winds. The seasonality was much more marked in the south of the lake. Interannual changes can be relatively well marked particularly for the rain seasons influenced by the Indian Ocean thermic states (Marchant et al. 2007).

WP 1.2 Limnology

Limnological monitoring was implemented to detail the main limnological changes, as this may be linked with environmental conditions triggering planktonic blooms possibly associated with cholera outbreaks at three sites (Uvira, Mpulungu and Kalemie). Additionally, this provided additional field observations to calibrate the remote sensing data (chl-a & lake surface temperatures).

Every week, during a three-year period, key limnological parameters were measured at three sites (Uvira in the north, Kalemie in the center and Mpulungu in the south). In Uvira and Mpulungu, the parameters were T°, pH, dissolved oxygen, conductivity, turbidity, transparency and chl-a concentrations at pelagic and coastal sites. The pelagic area was sampled every week at Uvira and every two weeks at Mpulungu (every 20-m depth from the surface to 100 m) because of logistic reasons. At Kalemie, T°, pH, conductivity and dissolved oxygen were measured every week at a coastal site every week. The coordinates of the sampling sites are indicated in Table I.

Table I: Geographical coordinates of limnological sampling sites.

Site	Latitude	Longitude
Uvira - coast	3° 25' 01.2" S	29° 02' 56.9" E
Uvira - pelagic	3° 24' 56.7" S	29°10' 38.1" E
Mpulungu- coast	8° 46' 05.5" S	31°06' 18.4" E
Mpulungu - pelagic	8° 45' 51.6" S	31°02' 38.1" E
Kalemie - coast	5° 57' 19.6" S	29° 11' 52.1" E

Water T° (°C), pH, conductivity (µS/cm at 25°C) and dissolved oxygen (mg/l) were measured in situ using an HANNA multiprobe HI 9828 at Uvira and Mpulungu, while a WTW 340I multiprobe was used at Kalemie. Turbidity (NTU) was measured using a Hach Turbidimeter 2100P at Uvira and Mpulungu. For the measurement of water transparency, a 30-cm diameter standard Secchi disk was used in each station. Chl-a (µg/l) levels were measured via spectrophotometry on filtered samples after a 24-h extraction step in acetone.

The time series of collection spanned from September 2011 to August 2014 at Uvira, from November 2011 to October 2014 at Mpulungu and from July 2012 to December 2014 in Kalemie. The database of limnological results was processed and shared amongst the project partners. At each station, a strong short-term variability of limnological parameters was observed, which may be linked to internal waves and pulses of deep nutrient-rich water towards the euphotic zone. A peak in chl-a levels, such as that observed in November/December 2011, are characteristic of that period

following trade wind inversion and re-equilibrium of the water layers (Plisnier et al. 1999). However, near the lake shoreline, rainwaters and erosion may significantly affect lake waters during the rain season, as reduced correlations between parameters compared to the pelagic area have been observed. Chlorophyll a levels obtained from remote sensing and cholera case numbers in the Uvira (Figure 3a) area present a dampening that may be linked to similar hydrodynamic patterns after seasonal wind direction changes (Plisnier and Coenen, 2001).

At Uvira, the weekly average water T° (from 0 to 100-m depth) was 26.0°C (min: 25°C to 27.2°C). The thermic stratification was always observed at Uvira although the mixing zone (average: 52 m) varied between 20 and 100 m in depth. Chlorophyll a levels in the surface water significantly correlated with chl-a levels in the mixolimnion zone (r = 0.68). This observation confirms that chl-a data monitoring from remote sensing of surface waters is a good indicator of chl-a concentration changes in deeper waters.

Water transparency at Uvira was inversely correlated with chl-a in pelagic (r = -0.35) and coastal (r=-0.32) waters. Water transparency was also significantly correlated to water T° in pelagic waters (r= 0.48). Higher water T° associated with increased thermic stability induces decreased probabilities of mixing of hypolimnion waters towards the surface, which is less favorable for phytoplankton development. Water transparency in coastal water correlated with transparency in pelagic water (r = 0.39). All indicated correlations are significant for alpha = 0.01.

At Mpulungu, the average water temperature for the 100-m upper water column was 25.7°C. For the complete year of sampling (2012 and 2013), this temperature fluctuated between 24.2°C and 27.1°C. In contrast to Uvira, the thermic stratification broke down every year between May and September during the southern upwelling (Coulter, 1991; Plisnier et al. 1999). Chlorophyll a concentrations in coastal water (Figure 3a) were also significantly correlated with chl-a concentrations in pelagic water (r=0.37). This correlation was even better observed with chl-a concentrations at the depth of 80 m (r = 0.54) and 100 m (r = 0.48) in pelagic waters. Coastal turbidity correlated well between both coastal area and pelagic waters (r=0.59 at 0 m; r= 0.70 at 100 m). Those observations indicate that frequent coastal monitoring at Mpulungu may provide information partially extrapolable to the pelagic area.
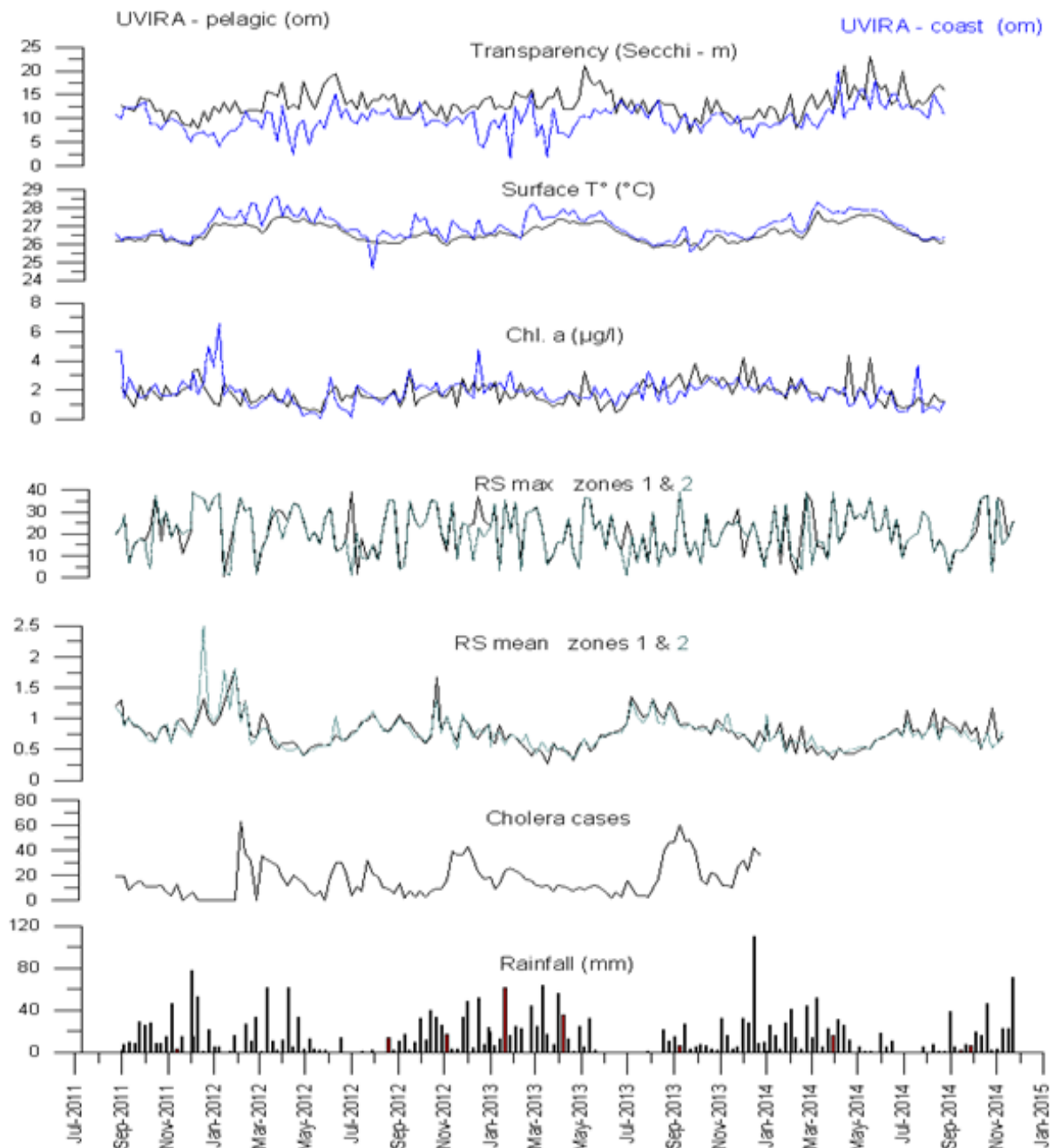
Fig. 3. Weekly water transparency (m), water T° (°C) and chl-a levels (µg/l) in coastal waters of Uvira from September 2011 to August 2014. Chlorophyll a max and mean values in zones 1 and 2, cholera cases in Uvira district and rainfall levels at Uvira are also indicated.
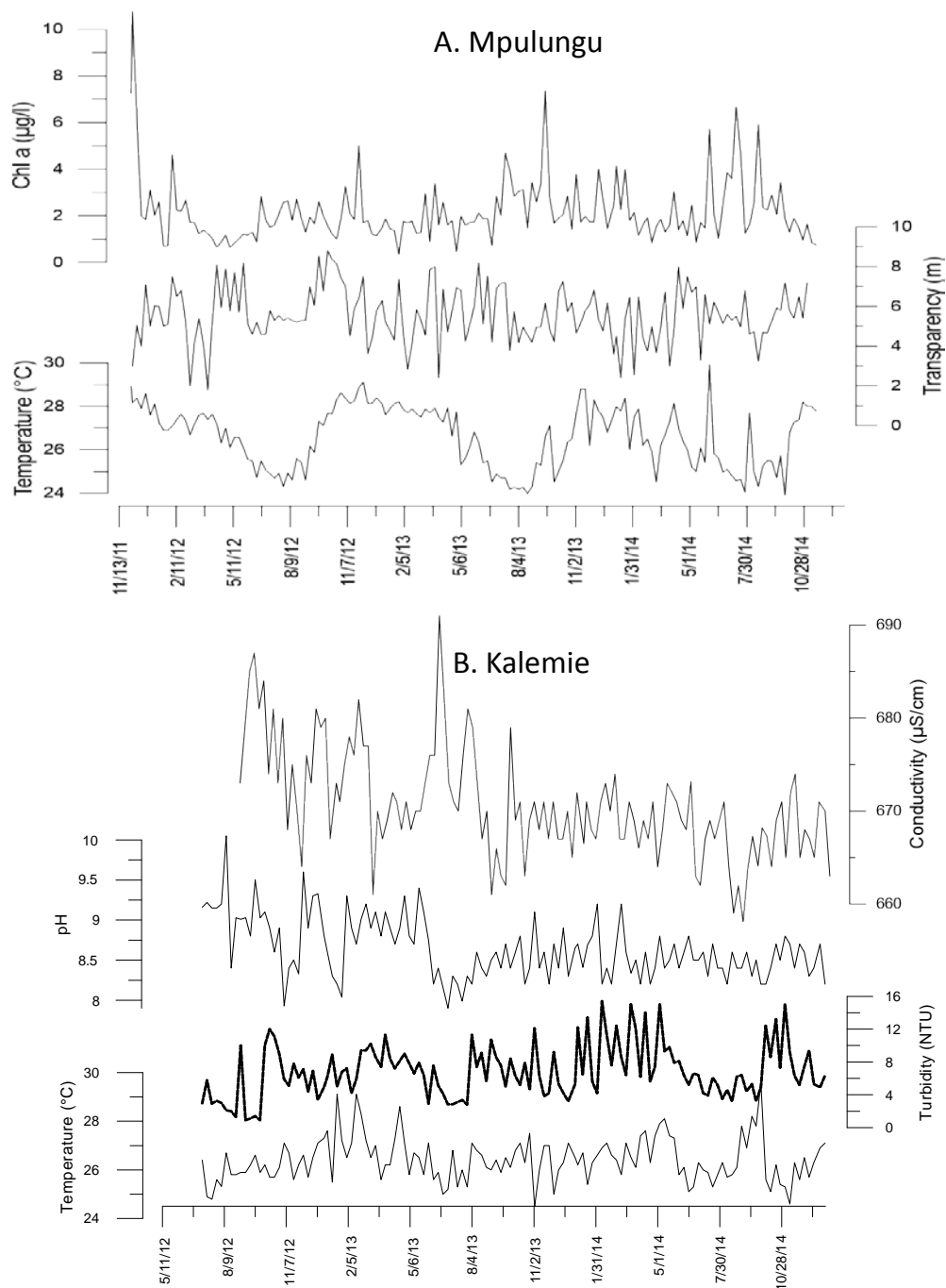
Fig. 4. (a) Weekly water T° (°C), transparency (m) and chl-a (µg/l) levels in coastal waters of Mpulungu from November 2011 to October 2014, (b) weekly water T° (°C), turbidity (NTU), pH and conductivity (µS/cm) in coastal waters at Kalemie from July 2012 to December 2014.

Transparency in pelagic waters at Mpulungu significantly correlated with water temperatures (r= 0.51 for surface waters) at Uvira. Average water transparency at Mpulungu (13.0 m) was similar to that observed at Uvira (13.1 m.) However, at Uvira, water transparency showed less fluctuations (9 to 18.7m) compared with Mpulungu (3.9 to 22.4 m). The stronger seasonality in Mpulungu induced changes in thermic stability or break-up of stratification during the upwelling, which explains marked differences in nutrient mixing from the hypolimnion towards the surface and more

variable planktonic biomass and transparency changes. This is confirmed by chl-a values (average: 1.7 µg/l in Uvira and 1.3 µg/l in Mpulungu), in which the maximum in the mixing zone reached 4.1 µg/l at Mpulungu compared to 3.0 µg/l at Uvira.

At Kalemie, temperature, turbidity, pH and conductivity measured in coastal water are presented in Figure 3b. The average pH was 8.6, although marked fluctuations may take place (pH from 7.9 to 10.1). High pH values may be recorded at the time of planktonic blooms, although field coastal measurements are not available for this station.

Deep-water isotherms at Uvira are good indicators of deep water fluctuations associated with internal waves and water mixing events. The damping of temperature isotherms from September (seasonal changes) until May/June was observed, while deep "colder" water raising up toward the surface was associated with nutrient rich waters (Plisnier et al. 1999). The pulse chl-a concentrations may be associated with those movements (Figure 5).
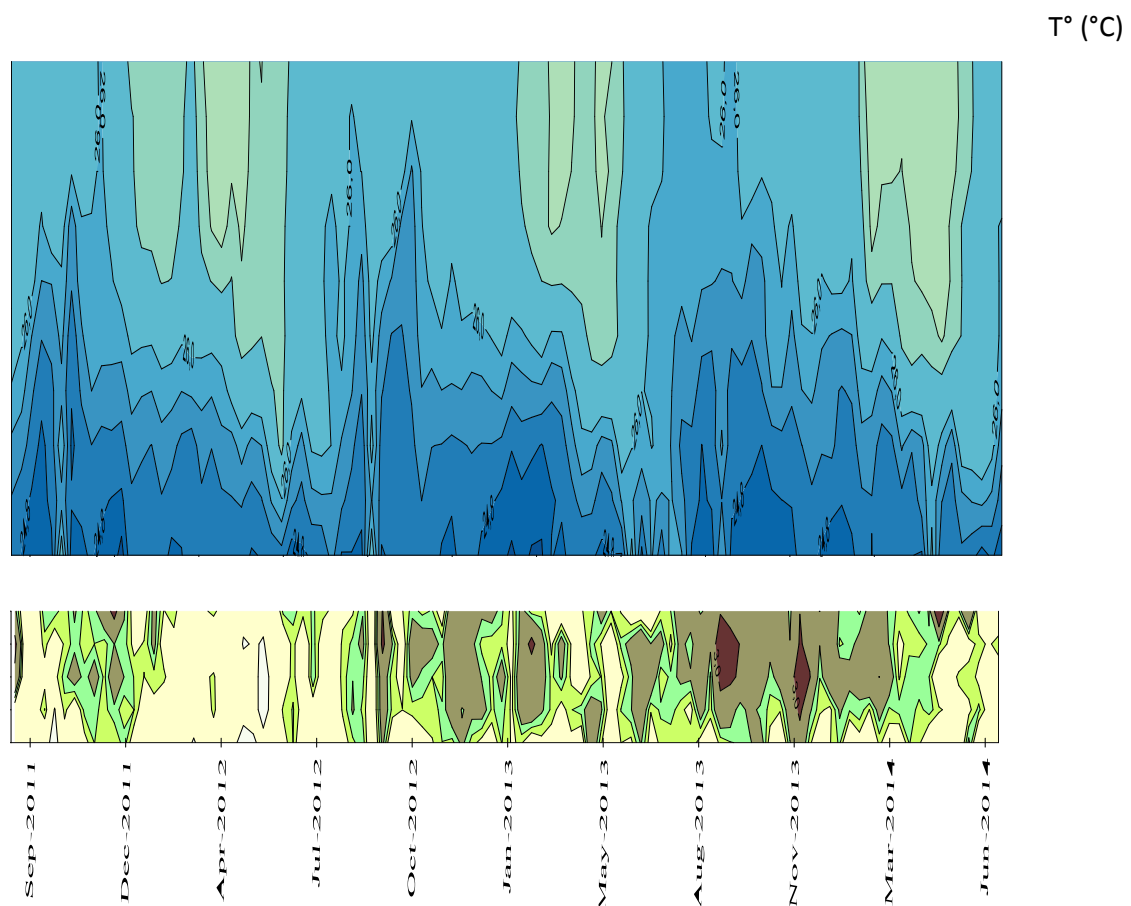
T° (°C)

Fig. 5. Weekly isotherms (°C) from 0 to 100 m and chl-a concentrations (µg/l) from 0 to 40 m in the pelagic area of Uvira (September 2011 to August 2014).

WP 1.3 Fish catches

At Lake Tanganyika, fishing takes place in the pelagic area during the night by boat that use light to attract the fish. Catch weights and composition of 10 fishing units (standardized to a number of 10 lights) per night were sampled by CHOLTIC each week at the return of fishermen to the beach at Uvira and Mpulungu. A subsample was collected to sort out the species caught. The main fish species targets were *Stolothrissa tanganicae*, *Limnothrissa miodon* and *Lates stappersii*.

Sampling catches on the beach was thus dependent of fishing techniques at each site. Because of local differences in fishing, "lift nets" (using catamaran boats) were sampled in the north, while "ring net" and "mutobi nets" were sampled in the south. In the south, fishermen alternate fishing techniques based on the species present at a particular time; therefore, it was not possible to sample from the same fishing techniques each week. The collected information is thus considered only an indicative index of species abundance. The monitoring was implemented every week of fishing besides new moon periods at Uvira and Mpulungu, as fishing does not take place then.

The relative abundance of main pelagic fish at two sites may indicate planktonic blooms (clupeids) or absence of blooms (catches of adult *L. stappersii* correlated with transparent water). It has been previously observed that abundance of *Stolothrissa tanganicae* may indicate planktonic blooms, such as that observed in the south of the lake each year during April/May when trade winds shift direction from northeast to southeast (Plisnier et al. 2009).

Fish relative abundance could also play a role in cholera outbreaks, either directly as vectors or indirectly via fishermen gathering in certain locations when fish catches are significant. It is possible (although not confirmed) that the processing of fish could also play a role in *V. cholerae* transmission.

In Uvira, the main fish catches were the small clupeids *Stolothrissa tanganicae* (86%), while *Limnothrissa miodon* and *Lates stappersii* represented 10% and 4% respectively of the weight of the sampled catches, respectively, during the three-year monitoring period of the CHOLTIC project in 2011-2014. Only small *Lates stapersii* were captured at this station (average total length: 11 cm).

Peaks and dampening of *Stolothrissa tanganicae* catches at Uvira were generally observed between September and December every year (Figure 6). This may be associated with mixing and peaks in limnological parameters as inferred by deep isotherms observed during the same periods (Figure 5).
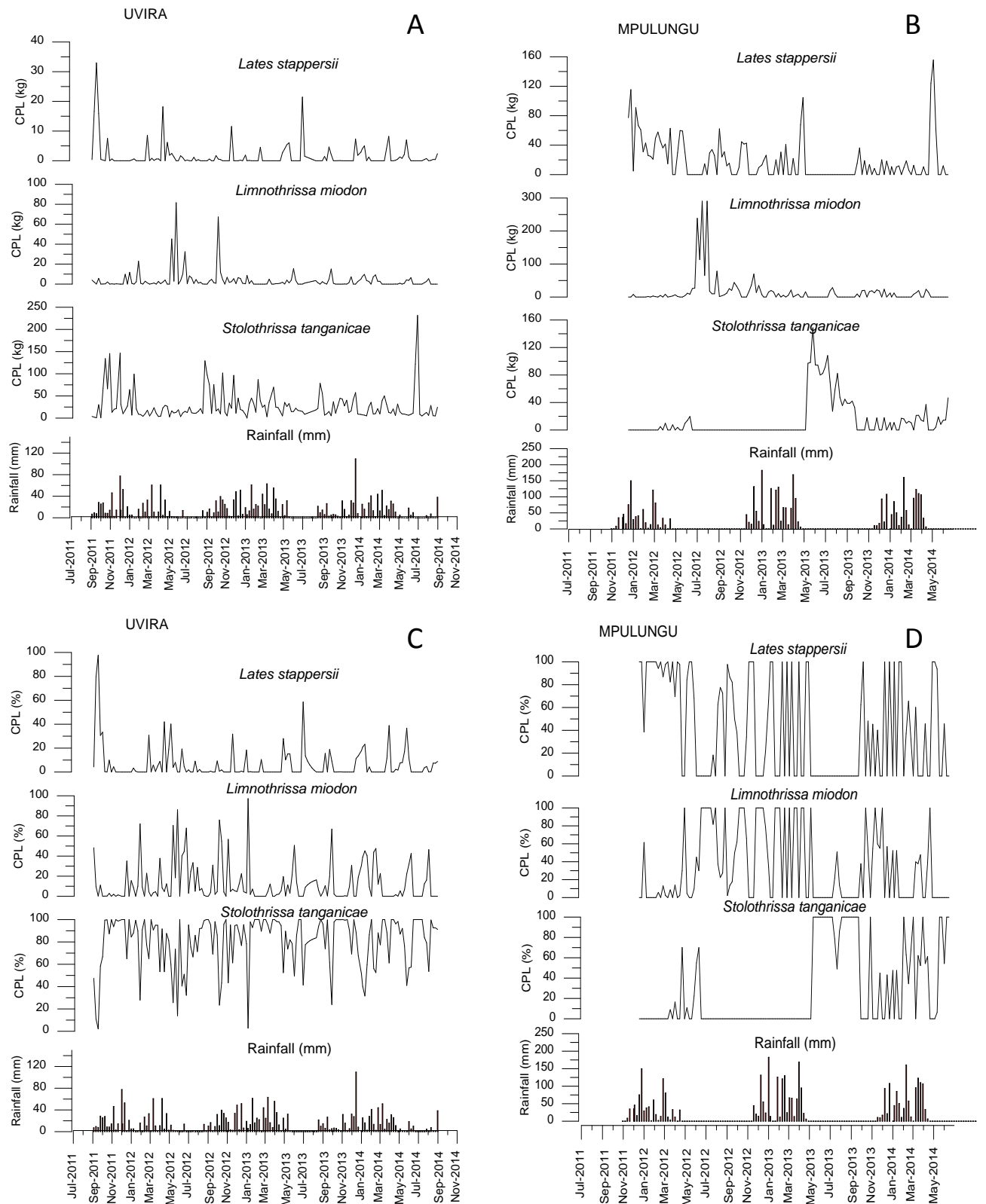
Fig. 6. Standardized catches (kg) per boat/night of *Lates stappersii*, *Limnothrissa miodon* and *Stolothrissa tanganicae* at Uvira (lift net) and Mpulungu (mutobi/ring net) (A and B). Percentages of those catches per species (C and D).

In Mpulungu, *Lates stappersii* represented 52% of the sampled catches, while *Limnothrissa miodon* was 25% and *Stolothrissa tanganicae* represented 23% of catches. Clupeid catches took place mainly at the onset of the dry season, in May/June when a planktonic bloom was observed (Figure 23). In 2013, catches of *Stolothrissa* persisted until the end of the upwelling period in September. A strong inverse relationship was observed during this period when no catches of *L. stappersii* were recorded from the sampling. Limnological changes are the main reason for this observation rather than predator-prey relationships (Plisnier et al., 2009).

WP 2 Phytoplankton dynamics

Phytoplankton was sampled weekly over the course of three years at a coastal site and a pelagic site in the northernmost part of the lake near Uvira (DRC) between the end of August 2011 and the end of August 2014 as well as in the southernmost part near Mpulungu (Zambia) between the end of October 2011 and the end of October 2014. The sampling at a third station along the west coast near Kalemie (DRC), from August 2012 until June 2013, was only conducted at a coastal site[2]   Phytoplankton quantification was performed according to Uthermöhl et al. (1931) using sediment chambers of 10 ml and an Olympus CXK41 inverted microscope with an immersion oil objective of 100x. A minimum of 500 solitary cells and/or colonies were enumerated; for each colony, individual cells were counted. Counts were restricted to organisms ≥ 3 µm. As smaller organisms belonging to the picoplankton fraction require another quantification method using epifluorescence techniques, they were not included in the present study. Biovolumes were calculated from the mean cell dimensions and converted to biomass. For the quantitative phytoplankton analyses, priority was given to the surface samples as they can be linked to the satellite images, although at the pelagic sites, sampling was also performed at different depths (10 m, 20 m and 30 m).

Phytoplankton dynamics near Uvira (Figures 7-8)

Phytoplankton in the northernmost part of Lake Tanganyika is characterized by a cyanobacteria-chlorophyte-diatom-dinophyte community. This is in contrast to the composition in 1975 (Hecky & Kling 1981), in which chrysophytes and, in smaller amounts, cryptophytes represented a major component of the phytoplankton population, while dinophytes were almost absent. The high algal biomass (always > 200 mg m^{-3}) noted in 1975 was never observed during our survey. However, it has not yet been demonstrated that both localities Uvira and Bujumbura, although located in

---

2 Several samples from Kalemie got lost during transport, and preliminary analyses of some of the remaining samples demonstrated that these did not represent the phytoplankton community as too many benthic organisms (especially diatoms) and sand particles were present.

the northernmost part of the lake, can be compared, as they are situated on the west and the east coast, respectively, with the mouth of the Rusizi River in between.

In the pelagic site, two peaks in phytoplankton abundance were observed at the beginning of the dry season in June and July 2012, and a smaller peak was observed later in the dry season, at the end of August. A bloom of *Aphanocapsa incerta* (Lemmermann) Cronberg & Komárek (Cyanobacteria) was responsible for these peaks. Cyanobacteria, which peak during the wet season, are *Chroococcus turgidus*, *C. limneticus*, *Anabaena* spp. and *Anabaenopsis* spp. The damping of the secondary upwelling of nutrient rich hypolimnion water at the end of the southeast winds season is clearly visible in the peaking pattern of *Chroococcus turgidus,* starting each year in October. Dominant taxa within the chlorophytes were *Siderocelis irregularis* Hindák*, Didymocystis* spp., *Dictyosphaerium* spp., cf. *Plectodictyon* sp. and *Coelastrum* spp. Biomass peaks of these taxa occurred during both the wet and dry seasons; however, an ebbing away of the peaking was observed. Chrysophyte levels were never elevated during our monitoring, which is in contrast with the observations in the pelagic near Bubumbura in 1975. Chrysophytes were represented by species belonging to, among others, the genera *Chromulina*, *Dinobryon* and *Ochromonas*. Diatom levels increased during the rain season, with *Nitzschia asterionelloides* O. Müller being the most important planktonic taxon. However, the peak during the dry season of 2013 was due to the benthic *Rhopalodia hirudiniformis* O. Müller. This may suggest heavy winds/storm impacts on hydrodynamics breaking up this attached species in coastal areas and transported to the pelagic waters. A peak of *Woloszynskia* sp. (dinophytes) was observed during the rainy season in the beginning of January 2013. Other dinophytes taxa such as *Peridinium africanum* Lemmermann and *Gymnodinium* spp. were sporadically present. Euglenophytes were represented mostly by *Trachelomonas* (among others *T. volvocina* (Ehrenberg) and *T. volvocinopsis* Svirenko), while *Lepocinclis* was only sporadically observed.

Overall phytoplankton biomass was higher at the coastal site than in the pelagic water, with a major peak (> 20000 mg.m^{-3}) at the beginning of January 2012 due to *Aphanothece clathrata* W. West & G.S. West (Cyanobacteria). This was followed by smaller peaks at the end of the wet/beginning of the dry season in 2012 caused by the same species, *Aphanocapsa incerta*, as observed in the pelagic waters and with a similar biomass (approximately 5000 mg.m^{-3}). *Chroococcus limneticus* was more abundant at the coastal site compared with *C. turgidus*. The peaking of *Anabaena* spp. and *Anabaenopsis* spp. followed the same trend as that observed in the pelagic waters although the peaks were much more important (often > 20 mg.m^{-3} versus never higher than 0.6 mg.m^{-3}, respectively).

The chlorophytes biomass was much more important at the coastal site than in the pelagic waters, and an oscillating peak pattern was observed throughout the seasons. *Didymocystis* spp. and cf. *Plectodictyon* sp. were the most abundant taxa peaking at the beginning of the rainy season in 2011. In contrast, the green algae belonging to the genus *Dictyosphaerium* were not the dominant taxa at the coastal site.
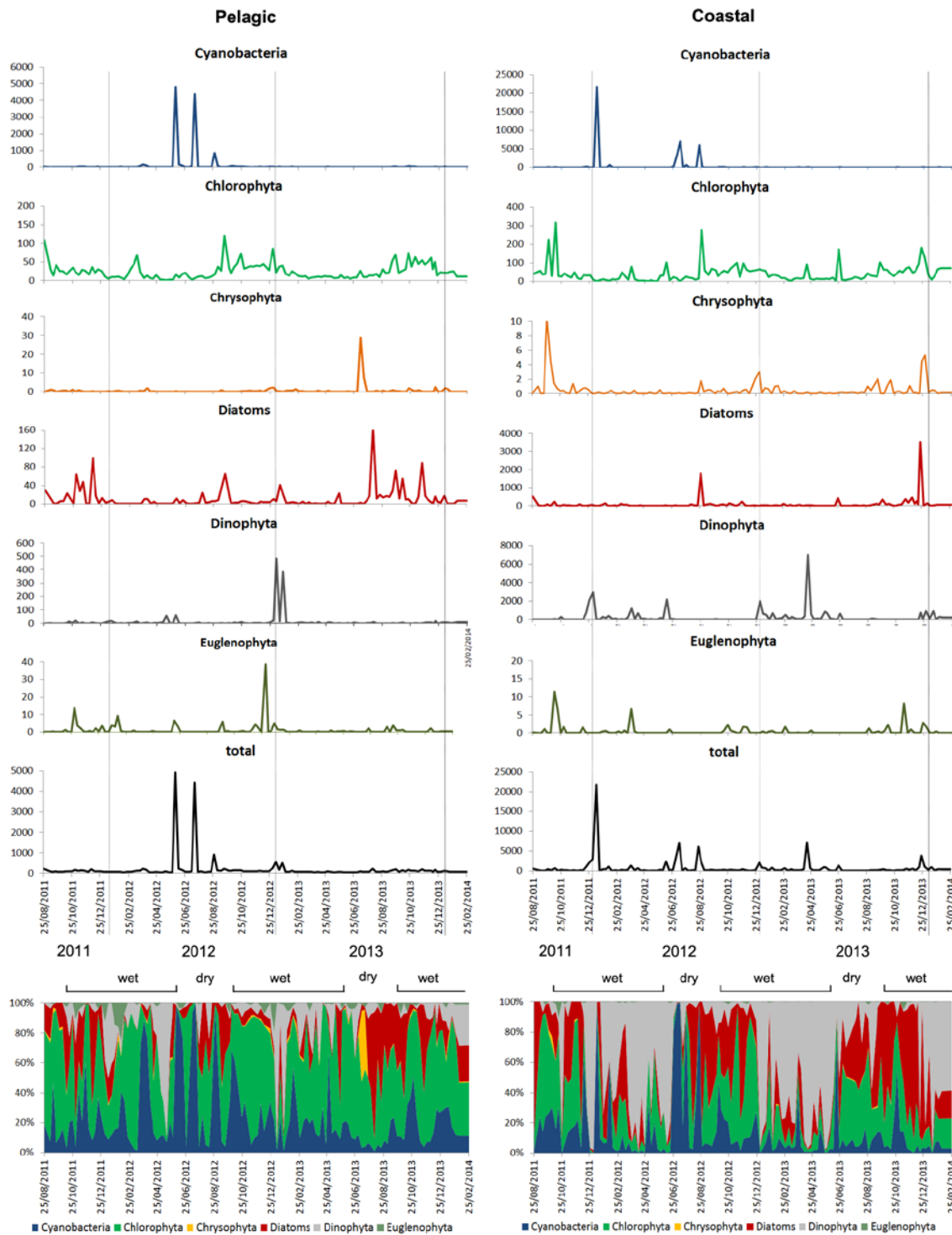

Fig. 7. Dynamics of cyanobacteria and algal groups in the pelagic and coastal sites at Uvira (DRC) expressed in biomass (mg.m^{-3}) and derived percentage for the period September 2011-February 2014.

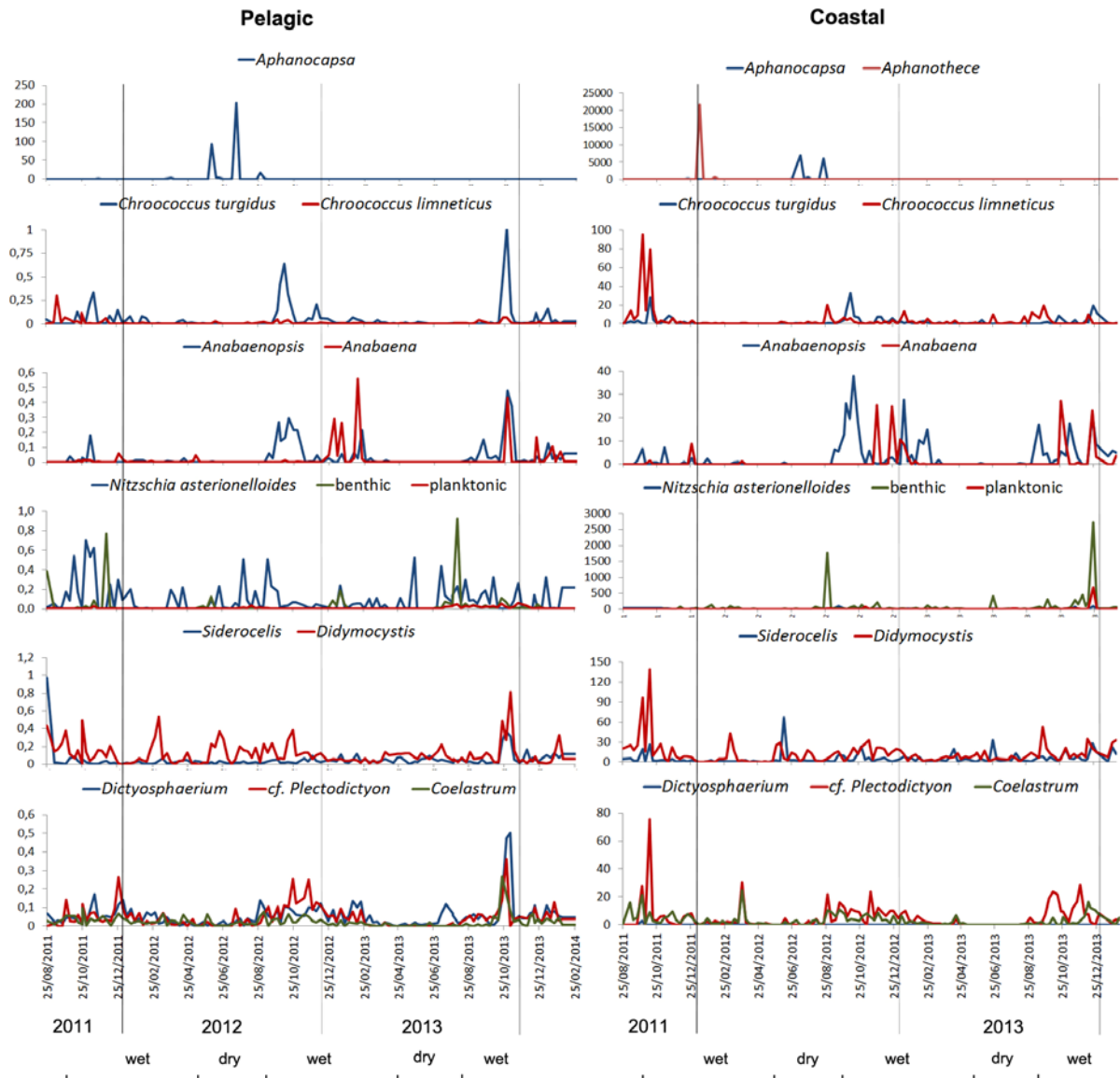Fig. 8. Dynamics of a selected number of taxa in the pelagic and coastal sites at Uvira (DRC) expressed in biomass (mg.m⁻³) for the period September 2011-February 2014. Cyanobacteria: *Aphanocapsa*, *Aphanothece*, *Chroococcus turgidus*, *Chroococcus limneticus*, *Anabaenopsis* and *Anabaena*. Diatoms: *Nitzschia asterionelloides*, other planktonic taxa and benthic taxa. Chlorophyta: *Siderocelis*, *Didymocystis*, *Dictyosphaerium*, cf. *Plectodictyon* and *Coelastrum*.

As in the pelagic water, chrysophytes were even less abundant, never reaching a significant biomass. Similar to the cyanobacteria and chlorophyte results, the diatom biomass was also higher in the coastal water compared with the pelagic site, with *Nitzschia asterionelloides* peaking at the end of 2013, although benthic diatoms were always dominant (e.g., *Encyonema minutum* (Hilse) D.G. Mann, *E. muelleri* (Hustedt) D.G. Mann, *Afrocymbella beccarii* (Grunow) Krammer, *Gomphonema clevei* Fricke and *Rhopalodia hirudiniformis*). Dinophytes peaked during the rainy season; the highest peak in 2013 was due to the same species, *Woloszynskia* sp., as that

responsible for the peak in the pelagic waters. However, a peak of *Peridinium africanum* was observed in December 2011.

Phytoplankton dynamics near Mpulungu (Figures 9-10)

The phytoplankton in the southernmost part of Lake Tanganyika was characterized by a cyanobacteria-chlorophyte-diatom-dinophyte community. During previous studies (CLIMLAKE project, financed by BELSPO) it has been shown that picoplankton, which was not taken into consideration in the present study, was often responsible for most of the algal biomass near Mpulungu. During periods in which picoplankton is primarily responsible for the algal biomass, it therefore creates the impression that the algal biomass near Mpulungu is sometimes relatively low.

At the pelagic site, the highest phytoplankton biomass (> 15000 mg.m^{-3}) was observed at the beginning of the dry season, at the end of May 2013, and a smaller peak was seen during the rainy season in February 2013 due to diatoms and dinophytes, respectively. The cyanobacteria peak during the rainy season of 2013 was among others due to *Chroococcus turgidus*. *Anabaenopsis* spp. and *Anabaena* spp. presented a higher biomass in the rainy seasons of 2011 and 2012, respectively. *Eremosphaera tanganyikae* Stoyneva, Gärtner, Cocquyt & Vyverman, *Dictyosphaerium* spp. and an unidentified green algal taxon were responsible for the chlorophyte peak during the rainy season of 2013. *Siderocelis irrgularis* showed an oscillating biomass pattern with the highest peak during the rainy season of 2013. *Ochromonas* was the dominant chrysophyte genus. Dinophytes were represented by *Woloszynskia* as the most abundant taxon, in addition to *Peridinium africanum* and *Gymnodinium* spp.

In the coastal site, algal biomass reached its highest peak (> 130000 mg.m^{-3}) during the rainy season at the beginning of the sampling period in November 2011, followed later in that season by peaks at the end of February and mid-April 2012. These were all due to the bloom of *Woloszynskia* (dinophyte), as were the smaller peaks during the rainy season of 2012. Only the smaller peak in the beginning of the dry season in 2013 was due to diatoms. Chlorophyte taxa peaking at the beginning of the rainy season in 2011 were *Eremosphaera*, cf. *Gloeocystis*, *Dictyosphaerium* and *Siderocelis irregularis*. The latter two taxa peaked again later during that rainy season and at the beginning of the rainy season of 2012. Within the diatoms, planktonic species were never significantly elevated with the exception of a peak of *Nitzschia asterionelloides* during the dry season of 2013, while benthic species peaked during the wet season at the end of 2012. Among the dinophytes, *Peridinium africanum* and *Gymnodinium* never became abundant. Euglenophytes peaked during the rainy season of 2011, with *Trachelomonas volvocina* being the most abundant taxa, although small peaks of

Lepocinclis and *Euglena* were observed during the same period. *Trachelomonas volvocina* was also responsible for a smaller peak at the end of the rainy season in 2013.
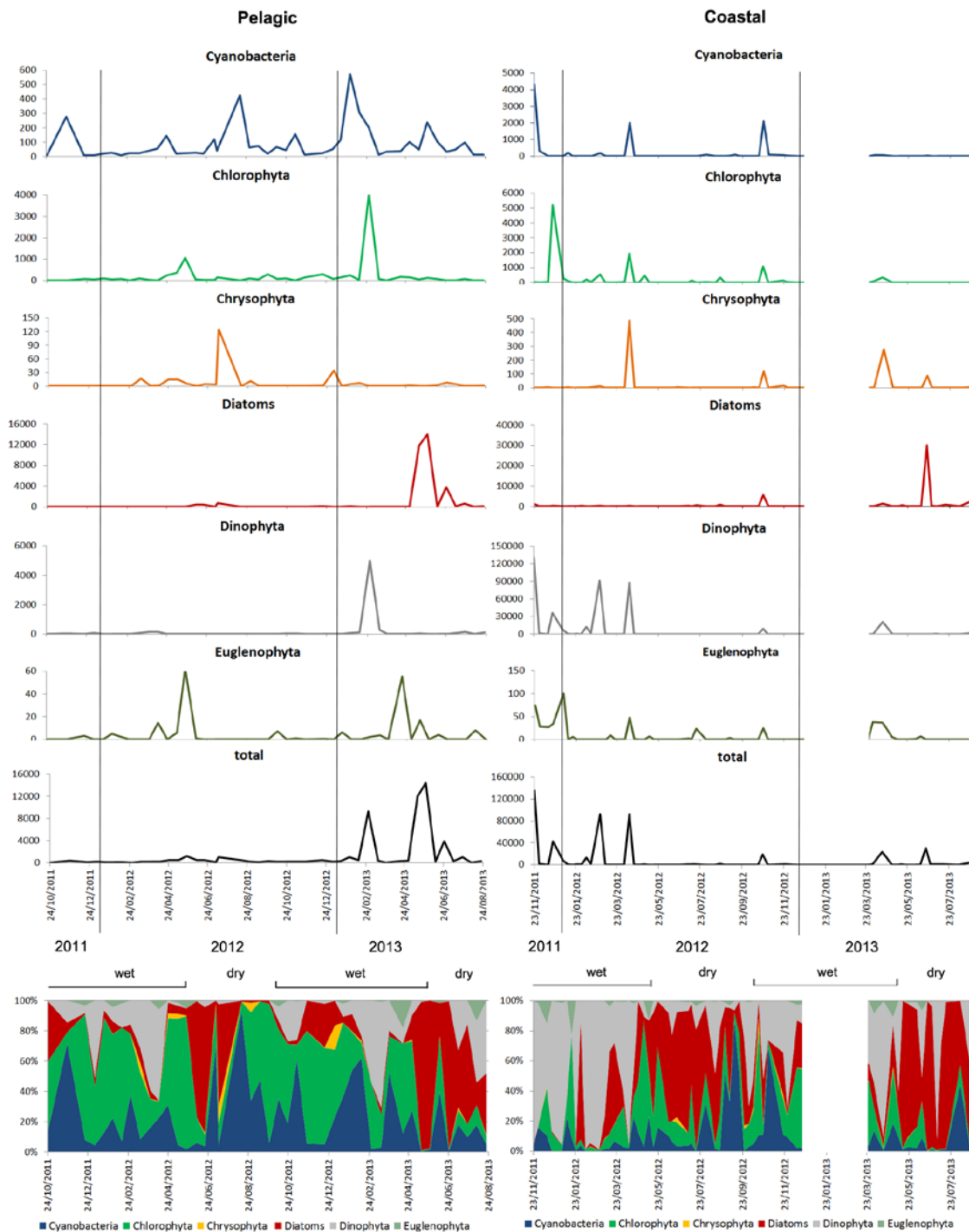

Fig. 9. Dynamics of cyanobacteria and algal groups at the pelagic and coastal sites at Mpulungu (Zambia) expressed in biomass (mg.m⁻³) and derived percentage for the period November 2011-August 2013.
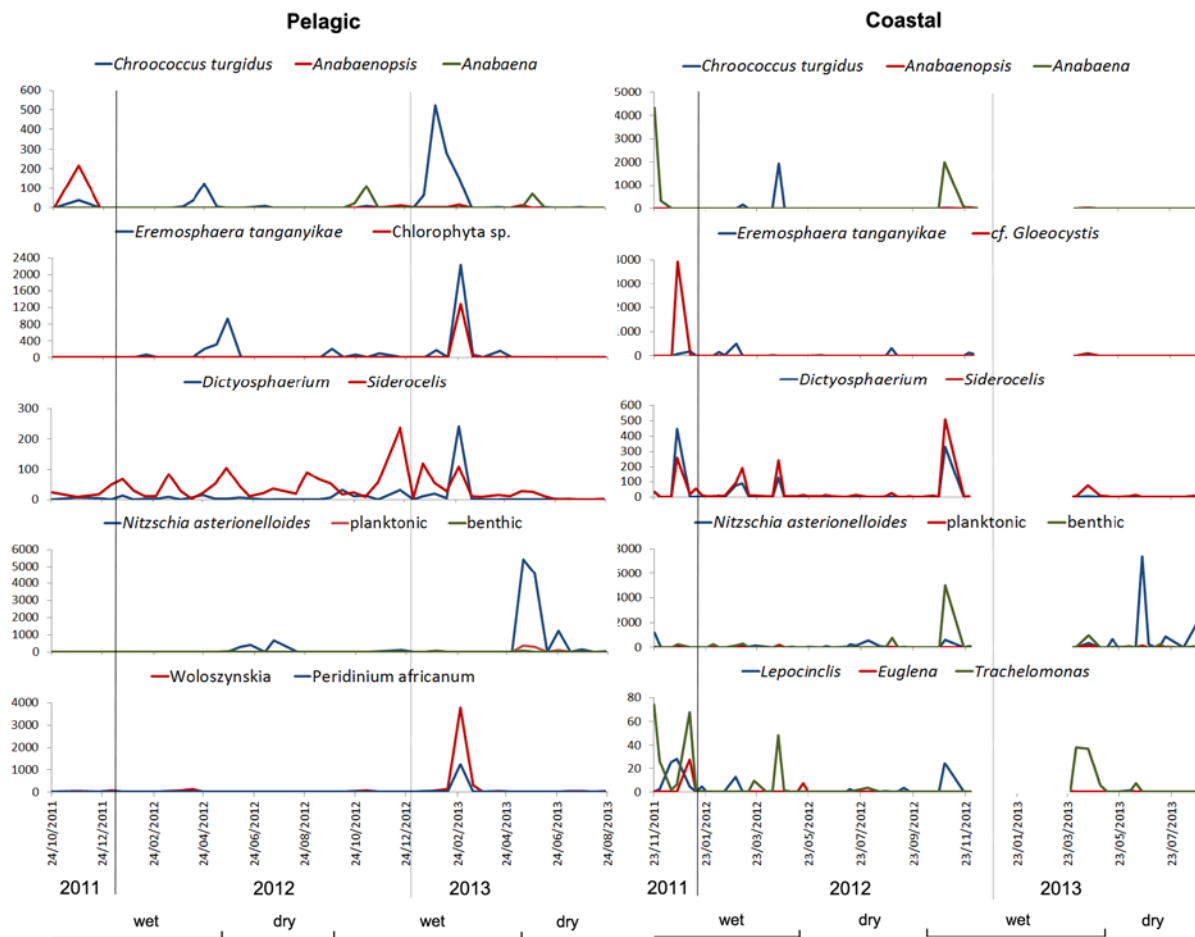
Fig.10. Dynamics of a selected number of taxa in the pelagic and coastal sites at Mpulungu (Zambia) expressed in biomass (mg.m⁻³) for the period November 2011-August 2013. Cyanobacteria: *Chroococcus turgidus*, *Anabaenopsis* and *Anabaena*. Chlorophyta: *Eremosphaera tanganyikae, an unknown chlorophyte species,* cf. *Gloeocystis, Dictyosphaerium* and *Siderocelis irregularis*. Diatoms: *Nitzschia asterionelloides*, other planktonic taxa and benthic taxa. Dinophyta: *Woloszynskia* and *Peridinium africanum*. Euglenophytes: *Euglena*, *Lepocinclis* and *Trachelomonas*.

The phytoplankton communities observed in Lake Tanganyika near Uvira and Mpulungu showed a clear seasonality with diatoms being more abundant during the dry season and dinophytes more abundant during the rainy season. This seasonality clearly demonstrated in Figures 7-10 is also visible in the sample graphs obtained from ordination techniques. An example of a PCA analysis, conducted on the pelagic phytoplankton dataset near Uvira, is provided in Figure 11. The plankton community characterized by *Woloszynskia* (dinophyte) is situated in the upper-right quadrant. Points 38, 41, 72 and 73 correspond to the samples collected on May 5, 2012; June 5, 2012; January 8, 2013 and January 22, 2013, respectively; the samples were dominated by *Woloszynskia* (WOLOSZYN).

These points (samples) are plotted on the graphs turning counter clockwise around the center of the graph (which indicates the seasonal characteristics of the samples) and returning to the same quadrant approximately one year later. Seasonal changes in major algal groups in the pelagic zone, observed during the survey period near Uvira and Mpulungu, have been reported for Lake Tanganyika (Verburg et al. 2003, Cocquyt & Vyverman 2005). The present study demonstrates that the seasonality is also clearly apparent in the coastal sampling region. The changes from the cyanobacteria-chrysophyte-chlorophyte community of 1975 (Hecky & Kling 1987) to a cyanobacteria-chlorophyte-diatom community as observed near Kigoma at the beginning of the 21st century (Cocquyt & Vyverman 2005) was confirmed for the northernmost part of the lake. However, the dominance of dinophyta at Uvira, not observed in previous studies during the CLIMLAKE project near Kigoma (Tanzania), is not a unique phenomenon but was observed during the three consecutive years of the CHOLTIC survey. Dinophytes were even dominant during most of the second half of the wet season near the coast. The cyanobacteria-chlorophyte-diatom community was then replaced by a dinophyte-chlorophyte-diatom community.

Figure 11 provides only preliminary exploration of the results. Species and samples will be submitted to an intensive investigation together with the available physical and chemical data. Even if the species names in the left graph in Figure 11 are not all legible, it is of no importance at the moment. We merely aimed to display the manner in which our data will be explored in future studies.
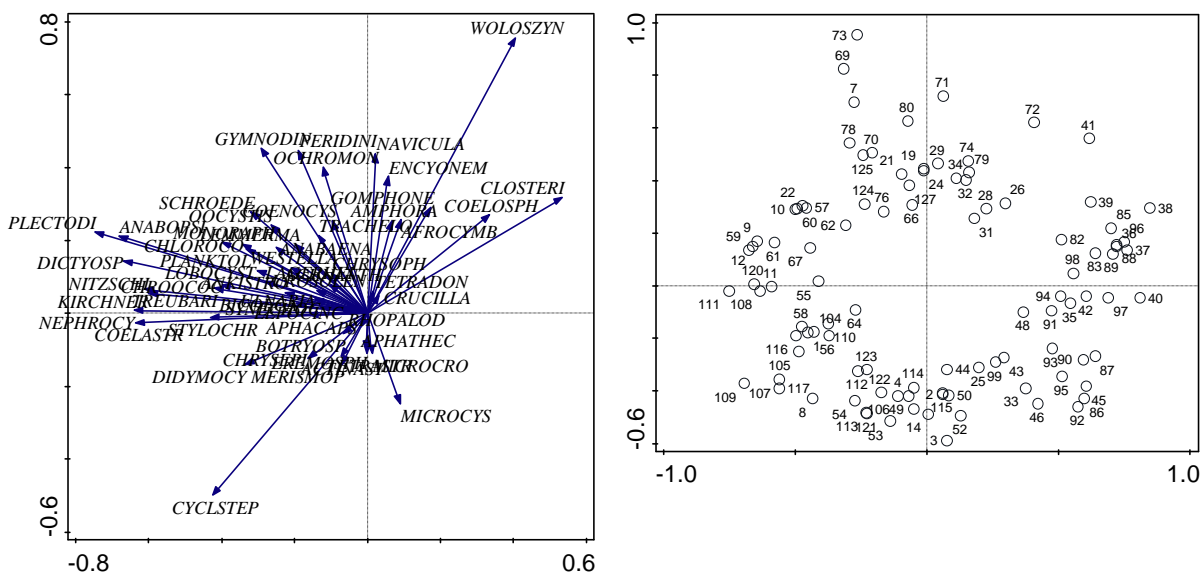

Fig. 11. Ordination (PCA axes 1 and 2) of the genera (acronyms) (left) and samples (dates replaced by numbering starting with 1) (right) derived from the pelagic site near Uvira for the period September 2011-February 2014.

Phytoplankton diversity (Table II):

Approximately 400 specific and infraspecific taxa were observed during the phytoplankton monitoring period in a pelagic site and a coastal site near Uvira in the north of Lake Tanganyika and near Mpulungu in the south of the lake. These figures concern only "living" cells. Of note, the number of genera and species for which only empty frustules ("dead" diatom cells) were observed are included in Table II but not incorporated into the other figures; they were all benthic diatoms.

Table II: Number of genera (bold) and species observed in quantitative surface samples collected at the coastal and pelagic sites in the northernmost part of the lake near Uvira (DRC) and the southernmost part of the lake near Mpulungu (Zambia) as an indication of phytoplanktonic biodiversity.

	Uvira				Mpulungu				Total	
	Pelagic		Coastal		Pelagic		Coastal			
Cyanobacteria	**14**	27	**30**	44	**25**	41	**21**	34	**39**	71
Chlorophytes	**38**	75	**37**	70	**43**	102	**39**	74	**62**	155
Chrysophytes	**5**	6	**6**	9	**4**	8	**8**	12	**11**	18
Cryptophytes	**-**	-	**2**	3	**3**	4	**2**	4	**6**	8
Diatoms	**12**	27	**32**	79	**17**	38	**30**	75	**34**	119
Dead frustules	*2*	*3*	*2*	*13*	*14*	*41*	*11*	*38*	*6*	*32*
Dinophytes	**3**	4	**3**	3	**3**	6	**4**	3	**3**	7
Euglenophytes	**2**	5	**4**	6	**4**	12	**2**	9	**5**	19

As expected, diatoms were more diverse at the coastal sites, where the majority were tychoplanktonic and benthic species. The latter were broken off from their substratum by turbulence in the shallower coastal sites or transported from small rivers into the lake. In the pelagic water, the majority were euplanctonic taxa, although some benthic species were observed, which was probably due to stormy weather conditions transporting, among others, broken off submerged parts of aquatic plants into the deeper parts of the lake far from the coast.

Based on the Canonical Correspondance Analysis (CCA), no link between cholera and environmental characteristics or algal groups composing the phytoplankton was demonstrated (Fig. 12). It must be noted that the plankton composition data concern only three years of monitoring (2011-2014) in contrast to the RS Chl-a data covering a longer period of 13 years (2002-2014).
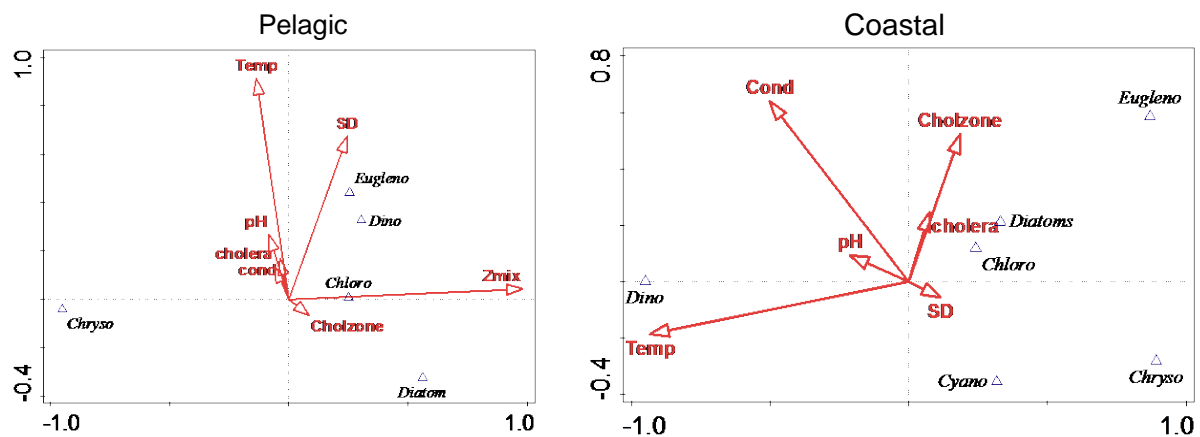
Fig. 12. Axes 1 and 2 of the CCA (Canonical Correspondance Analysis) of the taxa (algal groups) and some environmental variables for the pelagic and coastal sites near Uvira (surface samples). Cond: surface water conductivity, SD: Secchi disk depth, Temp: surface water temperature, Zmix: mixing depth of the water column, CholZone: cholera cases recorded for the entire Uvira zone; cholera: cholera cases reported for the coastal part of the Uvira zone.

WP 3 Zooplankton dynamics

Sampling of copepods took place from September 2011 to August 2014 in Uvira at the same sites where limnology and phytoplankton have taken place (pelagic and coastal areas). A sample was taken every week from 8 to 10 AM using a plankton net of 100 microns mesh and 25 cm aperture diameter. The net was weighted with a lead and raised slowly (0.5 m/s) on a vertical line of 60-0 m deep pelagic zone and pulled horizontally over a distance of 20 m at the surface in coastal areas. Each sample was adjusted to a volume of 50 ml and immediately preserved in 4% formalin (Haney 1988). The samples were then preserved in 70% ethanol and in some cases isolated specimens were dissected in a 10% glycerin solution onto slides using fine "brucelles" forceps (Dussart, 1980).

The quantification (OLYMPUS CH30RF200 and DIALUX 20 microscopes) of copepods was carried out by taking sub-samples of 1 mL using a pipette (opening 4 mm) after adjusting the entire sample (Isumbisho et al 2006) at the Biology laboratory of the Hydrobiology Research Centre (Uvira, DRC). For the identification and determination of copepod species, we used the optical microscope and sometimes the Scanning Electron Microscopy (Dussart and Defaye, 2001) and Proud (unpublished data). The samples collected in 2011 were not included in the results because of identification problems. To improve the identification skills of a researcher, a training session was organized at the Royal Museum for Central Africa and the Royal Institute of Natural Sciences in Belgium.

The principal copepod species identified during the study period showed a greater abundance in the pelagic zone (66%, F = 232, p <0.001) compared with the coastal area (34%, F = 94, p <0.001). Seasonal variations of total copepod abundances were more pronounced in 2013 (Figure 13). The three main species of copepods (*Tropocyclops tenellus*, *Tropodiaptomus simplex* and *Mesocyclops aequatorialis*) showed peaks in October 2013 at the beginning of the rainy season and in April/May 2014 at the end of the rainy season in pelagic and coastal waters. A strong peak was also observed in the abundance of ovigerous females in pelagic waters from October 2013 for *T. simplex* and *M. aequatorialis* (Figure 14).

Copepod species showed a coastal pelagic variable distribution according to their developmental stages. In coastal areas, this distribution for *T. simplex*, was nauplii> copepodites> adult> ovigerous females. Almost the inverse distribution was observed in the pelagic area. *T. simplex* nauplii dominated in both study sites and represented 57% and 47% of copepod biomass in the coastal and pelagic areas, respectively.

Seasonal variations of *M. aequatorialis* showed that this species could be absent for two consecutive months particularly at the start of the dry season (June and July). They showed sporadic peaks in April and August/September 2012 as well as in October 2013. *M. aequatorialis* nauplii were very rare at the beginning of the rainy season (September/October) and reappeared around December/January during the study period. The peaks observed around September/October and April/May corresponded to seasonal changes when wind direction changes caused mixing of nutrient-rich deep water towards the lake surface.
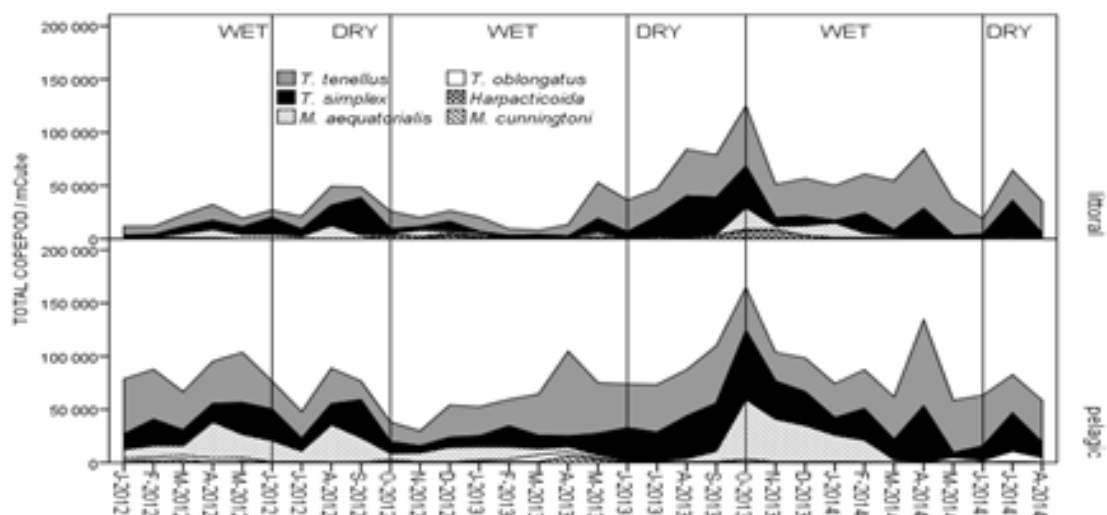

Fig. 13. Total number (individuals m-3) of copepods according to the sampling sites (coast: above and pelagic: below) and species in northern Lake Tanganyika.
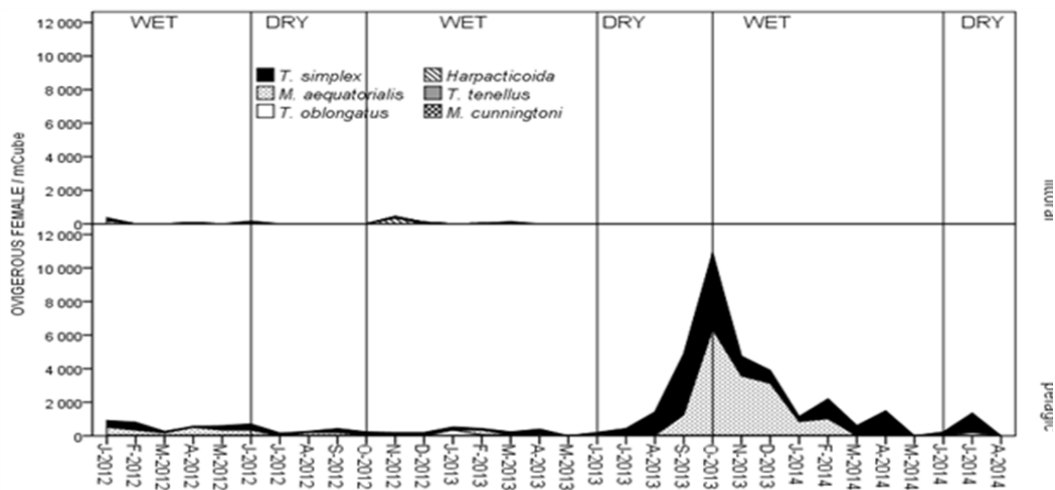
Fig. 14. Total number (individuals m^{-3}) of ovigerous female according to the sampling site (pelagic and littoral waters) and species in northern Lake Tanganyika.

WP 4 Remote Sensing

Methodology

WP4.1 Chlorophyll a and K490 extraction from MODIS Terra and Aqua

The objective of this WP was to produce Time Series (TS) of Ocean Color (OC) products for Lake Tanganyika from MODIS images for the period 2002-2014, extending the period covered by previous CLIMFISH project by eight years.

Satellite data preparation. All available **MODIS Level 1A data** (Reprocessing 5) covering the area of Lake Tanganyika were downloaded from NASA Ocean Color server http://oceancolor.gsfc.nasa.gov/. **SeaDAS v7.1** was used for extraction of geolocation information, computation of calibrated radiances and processing geophysical variables. For Visible and Near-Infrared (NIR) channels we defined appropriate atmospheric corrections for high altitude lakes. **Terrain height correction** was enabled when computing geolocation files, and altitude was taken into account to compute Rayleigh radiance using an external digital elevation model. Default stray light and cloud masking were disabled to preserve as many pixels as possible. We started processing from **Level 1A** to reduce the parallax effect and used already subsetted dataset to mitigate the data management aspect (data storage and processing time). Only Aqua data were used, as the OBPG does not encourage the use of MODIS Terra data for OC processing due to calibration problems for the sensor on board Terra (Franz et al. 2007). **Level 2** products corresponding to each 5-minute granule were computed. They include the following parameters: remote sensing reflectance in all visible wavelengths (Rrs), chl-a concentration computed using

several algorithms, K490, sensor and solar zenith angles and other useful ancillary information (e.g. flags). **External MODIS cloudmask product** (35_L2) from MODIS-Atmosphere group was also downloaded. This cloudmask defines four quality levels: cloudy, probably cloudy, probably clear and clear.

In situ dataset structuration, quality control and analysis. Collected **in situ data** from different cruises and stations (from WP1 and previous projects) were integrated in databases structures and consistent quality control was performed (i.e., wrong data identification, homogenization between different Excel file formats, correction of station coordinates, completion of missing time and instrument information and homogenization of NoData coding). This was required to perform inter-comparisons and validation of satellite data. In situ vertical profiles of water columns were also compared (vertical gradient and interpolation) to detect inconsistencies and identify the most reliable in situ data to calibrate/validate the OC and LSWT (WP4.2) parameter extraction process.

OC calibration phase. Important efforts were dedicated to select the most appropriate bio-optical algorithm and atmospheric correction during a first calibration phase. Of note, only a small portion of the signal recorded by the sensor on board the satellite actually comes from the lake water (small water-leaving radiance with respect to top-of-atmosphere radiance). This highlights the importance of **atmospheric correction**. The available in situ data together with the operational expectance of the RS contribution in this project did not enable the derivation of new sets of parameters for an atmospheric model specifically adapted to Lake Tanganyika. Therefore, it is necessary to reuse existing models. SeaDAS application enables the application of various atmospheric corrections by defining different **aerosol options** (aer_opt). In the current version, 80 'fixed' aerosol models numbered from 0 to 79 are implemented (Ahmad et al., 2010). Of note, these 'fixed' models are built according to eight different values for relative humidity [30, 50, 70, 75, 80, 85, 90 and 95%] and 10 values for particle size (defined as fine mode fraction) [95, 80, 50, 30, 20, 10, 5, 2, 1 and 0%]. A variation in fine mode fraction will lead to greater differences in the computed water-leaving radiance than a variation in relative humidity (McCarthy et al 2012). SeaDAS aer_opt parameter allows to either directly apply one of these 80 fixed models or choose some other aer_opt numbered -1 to -10. Most are variants of the procedure described by Gordon & Wang (1994), which is an iterative approach using two NIR channels for selection of the most appropriate fixed models (or a combination of two models). For example, option -2 is default one (multi-scattering with 2-band, RH-based model selection and iterative NIR correction), and -10 corresponds to Multi-scattering with MUMM correction and MUMM NIR calculation. Several of these options and bio-optical algorithms for chl-a concentration computation (OC2, OC3, OC3C, OC3 with CDOM Morel correction, Morel, Carder, GSM, GIOP and MGIOP) were tested on a

Lake Tanganyika MODIS dataset. In situ surface chl-a concentration measurements and chl-a concentrations from MODIS images was **matched** applying the following filtering criteria: maximum time difference of 6 hours, 'clear' or 'probably clear' sky at selected pixel according to MODIS Atmosphere cloudmask, minimum coast distance of 1 km and maximum spatial distance of 4 km. In situ values have been associated either directly with the matched pixel value or the median extracted in a 5-by-5-pixel window around this pixel to retain a greater number of points in match-up. Certain aerosol model/bio-optical algorithm combinations result in a higher number of missing values of chl-a concentrations. Only matching observations for which a value is given for each tested combination have been retained.

WP4.2: Lake Surface Water Temperature (LSWT)

The **objective** of this work package was to produce time series of LSWT for Lake Tanganyika from MODIS images for the period 2002-2014.

Satellite data preparation. Same MODIS Level 1A dataset for OC was used to derive LSWT. Level 2 products corresponding to each 5-minute granule were computed containing brightness temperatures (BTs) in different thermal IR channels, sensor and solar zenith angles and other useful ancillary information (e.g. flags). To calibrate the LSWT algorithm, we used daytime and nighttime lake surface temperatures of the ARC-Lake dataset (MacCallum & Merchant, 2013), which was produced from (A)ATSRs acquisitions for the period 1991-2011 with a spatial resolution of 0.05 degrees and a temporal resolution of three days. More specifically, PLOBS3D and PLOBS3N datasets (i.e., unaveraged per lake observations from ENVISAT/AATSR only) were exploited. Coarser spatial and temporal resolution was the main disadvantage of ARC-Lake for our purpose. The MODIS and ARC-Lake data fusion approach was thus justified. Furthermore, we validated it against existing in situ measurements to demonstrate its usefulness as a reference to produce Lake Tanganyika specific model coefficients.

LSWT computation. Two methods were foreseen to derive LSWT coefficients for MODIS/Terra and Aqua sensors using an ARC-Lake dataset to adjust the coefficients of Coast Watch-like models or Radiative Transfer Modeling (RTM) to simulate brightness temperatures (Francois et al., 2002). RTM was not applied since the first method provided very accurate results. **Triple** (T37) **window algorithm** [McClain et al., 1985] with channels 3.96, 11.04 and 12.04 µm was used for nighttime:

$$MCSST\_TRIPLE = \boldsymbol{a}\, T11 + \boldsymbol{b}\, (T3.96 - T12) + \boldsymbol{c}\, (T3.96 - T12)(\sec\theta - 1) + \boldsymbol{d}$$

A non-linear (NL) **split window algorithm** [Walton et al., 1998] applied to thermal IR channels 11.04 and 12.04 µm was used for daytime.

$$NLSST\_SPLIT = \boldsymbol{a}\,T11 + \boldsymbol{b}\,MCSST\,(T11 - T12) + \boldsymbol{c}\,(T11 - T12)(\sec\theta - 1) + \boldsymbol{d}$$

in which MCSST is estimated by:

$$MCSST\_SPLIT = \boldsymbol{a}\,T11 + \boldsymbol{b}\,(T11 - T12) + \boldsymbol{c}\,(T11 - T12)(\sec\theta - 1) + \boldsymbol{d}$$

MODIS and ARC-Lake observations were matched using the following **filtering criteria**: time difference <= 4 h, spatial distance <= 10 km, distance to coast >= 2 km, external MODIS cloudmask clear or probably clear, cloud_albedo flag from SeaDAS processing < 0.054, and standard deviation in 3x3 window of BT 11 µm and 12 µm channels < 0.05 K (to avoid using too many undetected cloudy pixels). Equations coefficients were adjusted from the matching observations separately for day/night, aqua/terra, yearly/monthly via regressing MODIS BT against ARC-Lake LSWT.

WP 4.3: Landcover maps of the three epidemiological test zones

Objective. As suggested in Ali et al. (2002), landcover maps have been produced for three epidemiological sites (Mpulungu, Kalemie and Uvira) to characterize the environmental context (permanent or temporary wet zones) of the cholera cases.

Image classification strategy. Several medium resolution multispectral images (Landsat 5, 7 and 8) were selected to take advantage of the multispectral and diachronic information, avoiding as much as possible the presence of clouds. Cloud masking and classification were performed and final land cover maps were finalized for the three sites. For more information, please consult the master thesis of F. De Smet D'Olbecke (2014).

Results

For both WP4.1 and 4.2, several scripts (MATLAB, Python) were written to automate the processing.

Cloud-masking issue. For both LSWT and OC, cloud masking was critical step. Cloud frequencyon Figure 15 shows a clear distinction between north and south as well as between day and night that would have a direct consequence on the quality of derived data. Several test results demonstrated that even a dedicated MODIS cloudmask product is not accurate enough for the project purposes, and cloud cover is sometimes underestimated leading, for example, to LSWT values lower than expected. We thus adapted cloud masking for the different concerns (derivation of LSWT coefficients or production of final LSWT/OC datasets) by applying simple thresholding methods with various conservative levels.
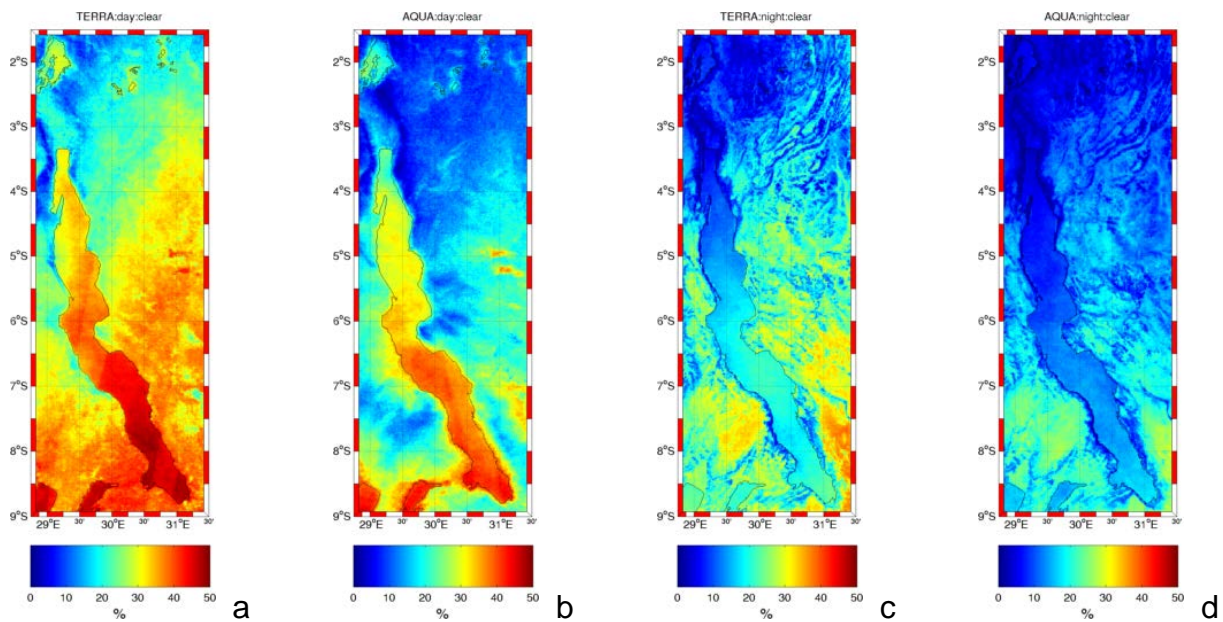
Fig. 15. Frequency of clear pixels from MODIS cloudmask based on a subsample of 999 Aqua images (2003-2005) and 1025 Terra images (2004-2006). a) Terra day (~8:30AM UTC); b) Aqua day (11:30AM UTC); c) Terra night (8:30PM UTC) and d) Aqua night (11:30PM UTC).

WP4.1 Chl-a and K490 extraction from MODIS Terra and Aqua

Calibration/validation against the *in situ* data. The calibration for OC products based on the in situ chl-a data was substantial. Only a few results are summarized in the present report. As three different methods were used for the chl-a measurements on the field between 2002 and 2015, depending on the period and the station considered, we decided to handle HPLC and 'field' (i.e., fluorometry or spectrophotometry) measurements separately. After selecting surface, non-coastal chl-a in situ measurements concomitant with clear sky (see match-up criteria in methodological section), the number of available observations for calibration was considerably reduced but still significantly increased with respect to the former CLIMFISH experiment. In this case, R^2 is not a reliable measure of performance of the tested algorithm. Figures 17 to 19 synthesize the comparison between all tested atmospheric algorithm combinations and in situ data following the recommendations of Campbell & O'Reilly (2006). Unfortunately this analysis did not permit to clearly identify a unique best combination. Results were sometimes contradictory using non-HPLC measurements (e.g., they show negative instead of positive bias) or only one-pixel values instead of the median in the neighborhood. In certain cases, statistics can also be influenced by the low number of observations (especially when using non-HPLC measurements).
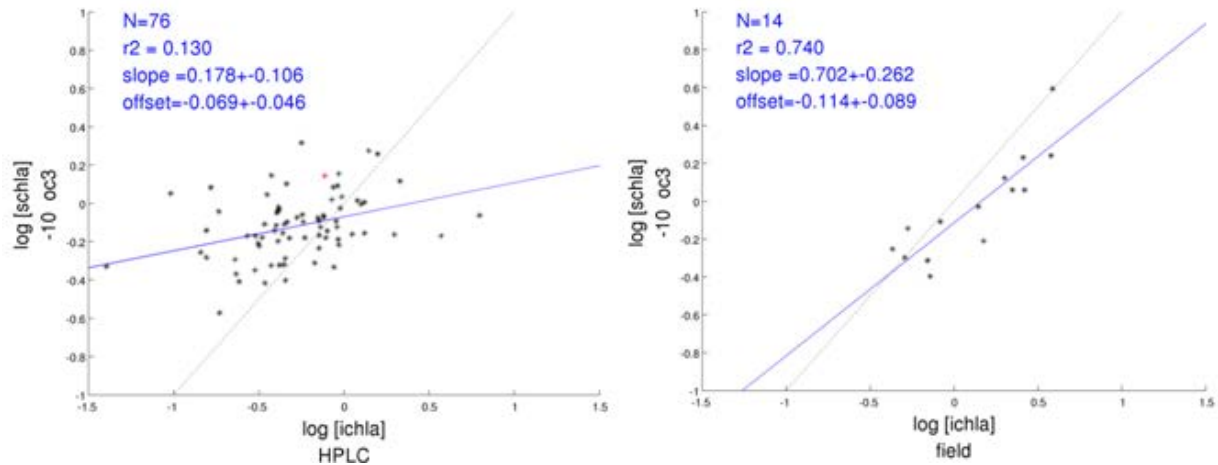
Fig. 16: Relationship between measurements of chl-a (µg/m) obtained using MODIS Aqua OC3 algorithm and MUMM aer_opt -10, median in a 5-by-5-pixel window (Y axis) and in situ HPLC values on the left and non-HPLC values on the right (X axis).

The range of chl-a concentration values from in situ (the values in the match-up) was very small (very few high values and no very high values). We noticed that satellite products tended to show even smaller dispersion, i.e., satellite data probably overestimated very low values and underestimated higher values. This is clearly visible from Figure 16 and Figure 20. Aerosol options -2 or -9 first appeared to yield larger ranges of value, although this was artificially created by artifacts created during the atmospheric correction, thereby quite often giving rise to incoherent spatial discontinuities (Figure 21). Moreover, another test performed on 645 nm Rrs (Figure 22) shows that these options -2/-9 also tended to overestimate the contribution of atmosphere, thereby giving rise to physically impossible negative reflectance values. This figure shows that fixed model with high fine mode (21, 62, 73) also tended to overcorrect the atmospheric effect, at least during the dry season.
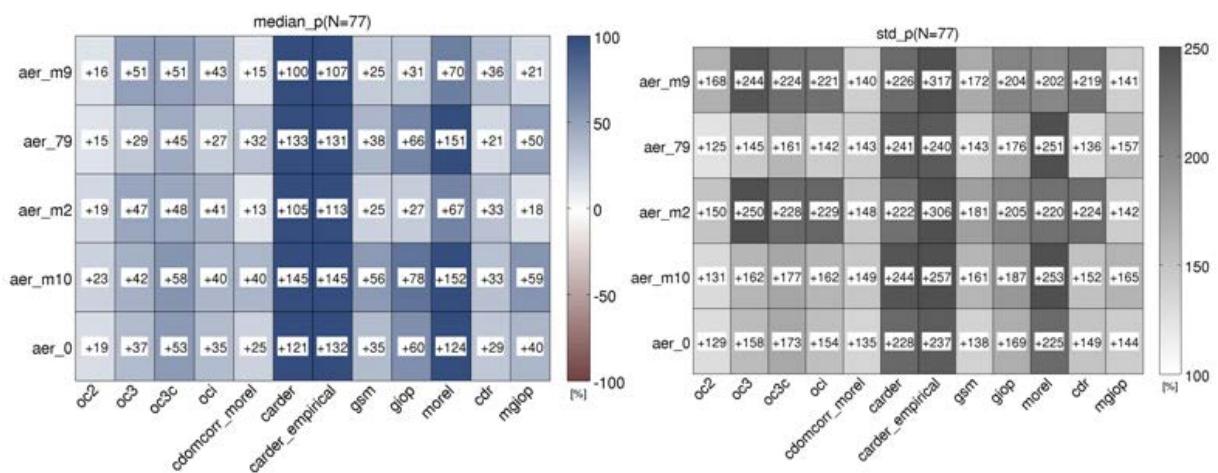

Fig. 17. Median (bias) (left) and standard deviation (imprecision) (right) in percentage of the difference between in situ and MODIS Aqua values for 12 OC and 5 aerosol models.
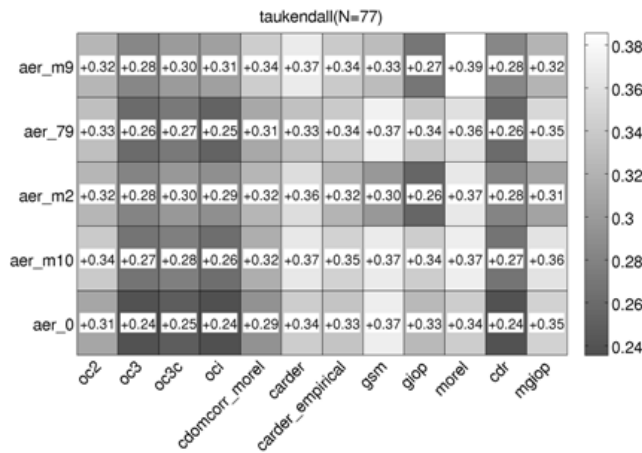
Fig. 18. Rank correlation (Kendall Tau) between in situ and satellite values of chl-a A tau near 0 means a low association (HPLC and median values in 5-by-5 pixels).
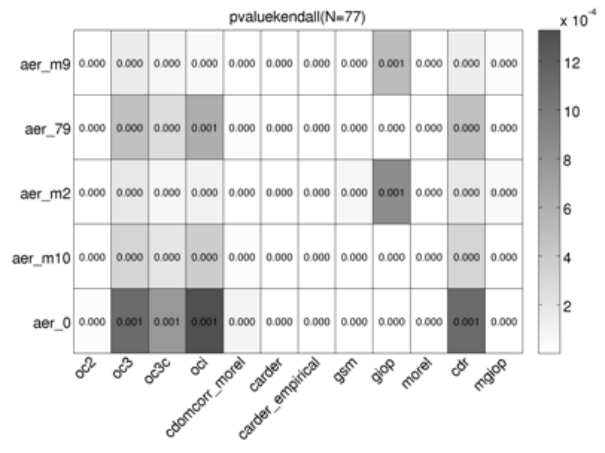
Fig. 19. P-value tables associated with the Kendall Tau (stat.signif. of the tau values). The association between in situ and RS chl-a (HPLC and median values in 5-by-5 pixels).
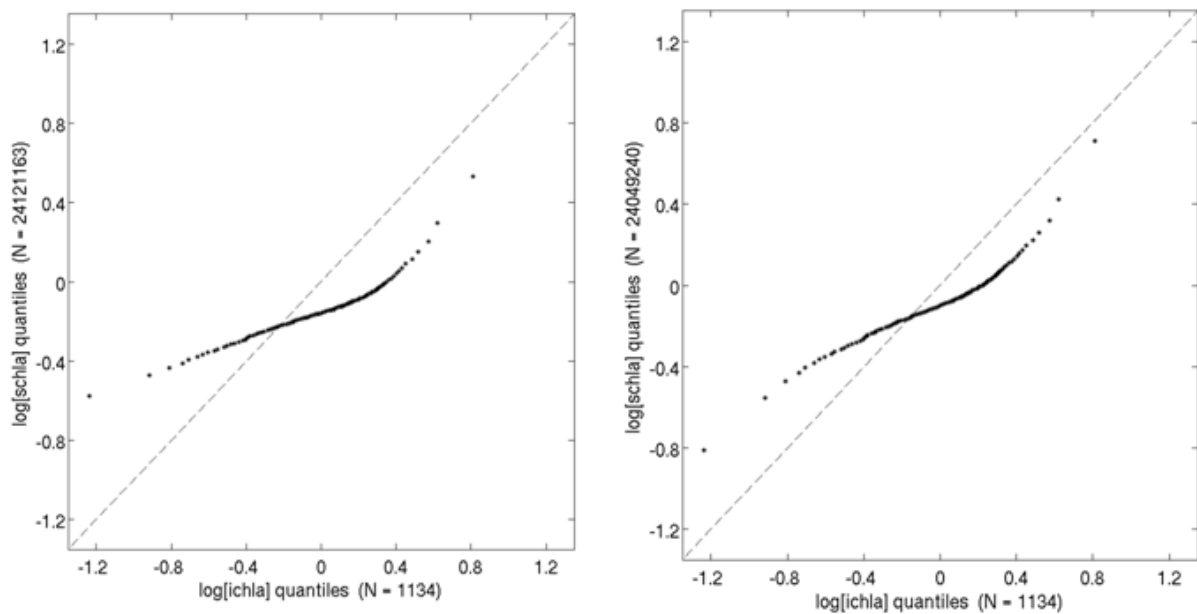

Fig. 20. Quantile-quantile plots comparing general distributions of satellite values (Y-axis - OC3 algorithm with aer_opt 79 on the left and -10 on the right) and in situ chl-a values (X-axis). Matching observations, all pixels from the 2002-2014 period, and all surface *in situ* measurements, including coastal, were applied.
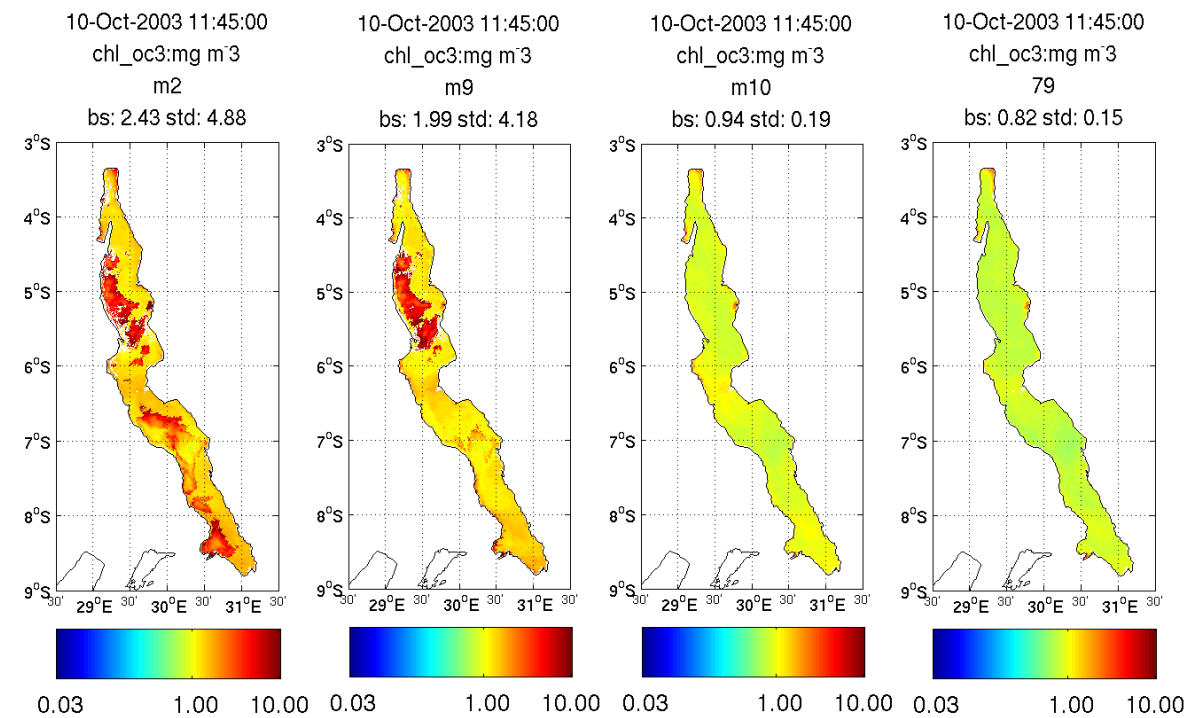
Fig. 21. aer_opt -2 and -9 (left) show unreliable spatial discontinuities in the computed chl-a concentration values. These discontinuities do not appear for options -10 and 79 (right).
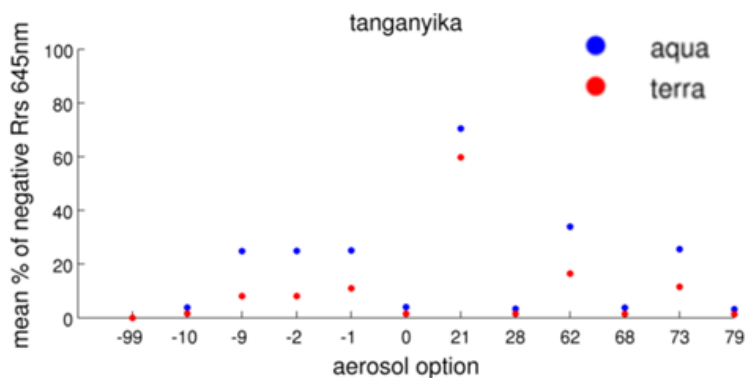

Fig. 22. Proportion of negative reflectance in band 645 nm based on a small number of clear sky images (mainly from the May to September period, when clear sky is more frequent).

On the basis of these arguments, we produced the chl-a series using the OC3 algorithm with the two best-performing aerosol models for atmospheric correction: 79 (fixed model based on relative humidity of 90% and fine mode fraction of 0%) and -10 (multi-scattering with MUMM correction and MUMM NIR calculation). Concerning the bio-optical algorithm, we observed a lower effect on the obtained chl-a concentration values compared with the variation of aerosol model. The OC3 algorithm, default algorithm for MODIS, was selected, as we could not verify the global spatiotemporal superiority of some other algorithm based on *in situ* data. Kendall tau p-values presented in Figure 19 yielded significant association for all combinations. More complex algorithms also produce more frequent no-data values. They also often require the setting of additional parameters; however, we do not have enough

elements to automate that parameterization properly for each individual image. Our approach is supported by Knox et al (2014), who also considers appropriate for Lake Kivu the OC3 algorithm combined with coastal atmospheric correction model with 90% relative humidity (very similar to aer_opt 79 in the present report). Furthermore, the same approach was already applied by Horion et al (2010) on Lake Tanganyika.

Final dataset. The final products in **raster** format and **snapshots** display chl-a concentrations at spatial resolution of 0.02 degrees for the entire lake with a daily and weekly temporal resolution for the period July 2002–November 2014 with the OC3 algorithm (Fig.) and the two aforementioned aerosol options.
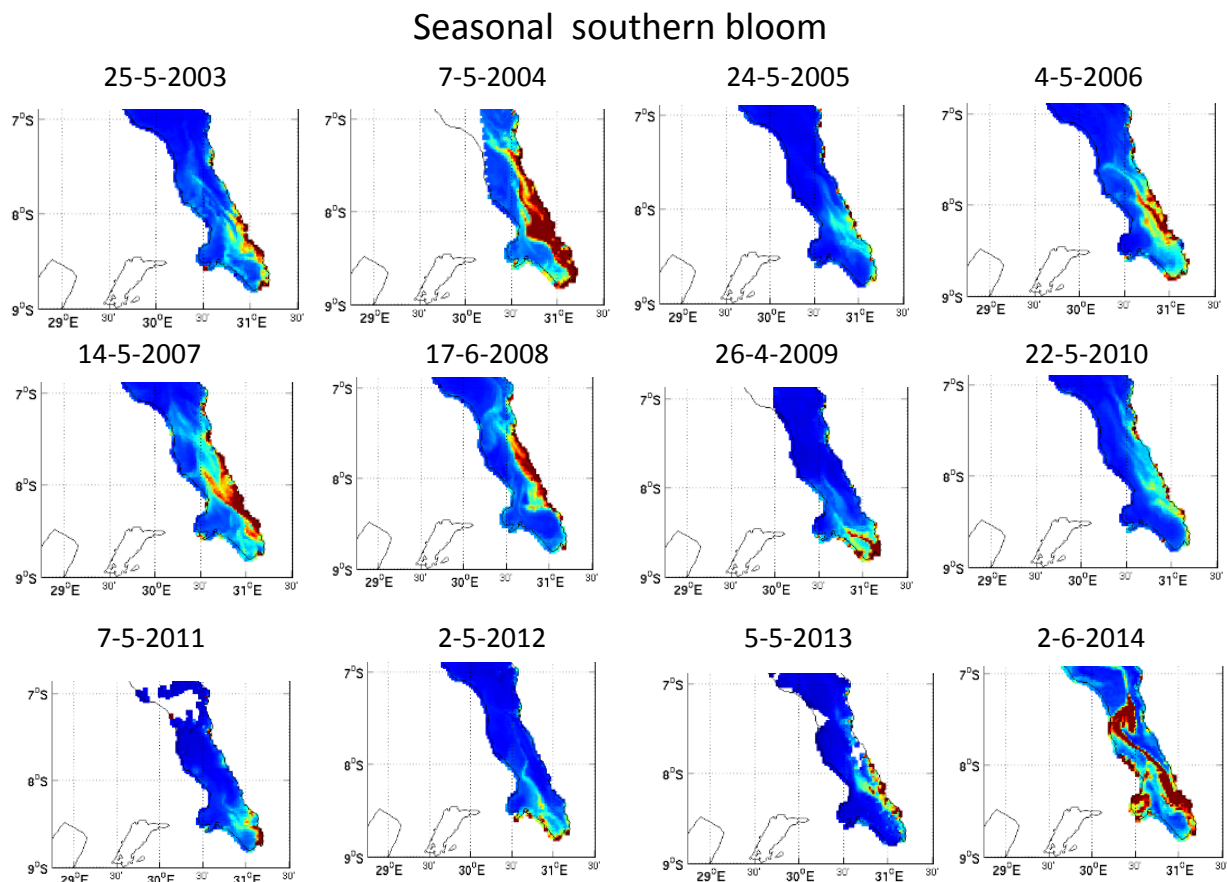
Seasonal southern bloom

Fig. 23: TS subset of daily chl-a concentration images (OC3-aer_opt -10) showing seasonal planktonic blooms (April/May to June) in southern Lake Tanganyika at the annual period of trade winds changes (from northeast to southeast) for the period 2003 to 2014.

Tables containing weekly spatiotemporally aggregated TS for the entire lake (Figure 24) and ecoregions from Bergamino et al. (2010) were produced. Our analysis suggests that the uncertainty of absolute chl-a concentration values computed in one location and time from MODIS data cannot be considered as negligible, particularly for the highest concentration values. However, the obtained TS demonstrates a certain coherence in time and space and can definitely be used in a relative and qualitative manner. Furthermore the spatio-temporal aggregation that will be used in the

interdisciplinary analysis greatly reduces the effect of local highly inaccurate values. This can efficiently be exploited for association analysis between cholera and environmental factors, even if the amplitude of the variation from RS products is reduced with respect to in situ measurements. Nevertheless, the signal remains quite noisy due to the use of a cloud mask that was not sufficiently constraining. At this stage, K490 TS were also computed but not used and distributed, as we estimated that the added value of this parameter with respect to chl-a concentrations was limited.
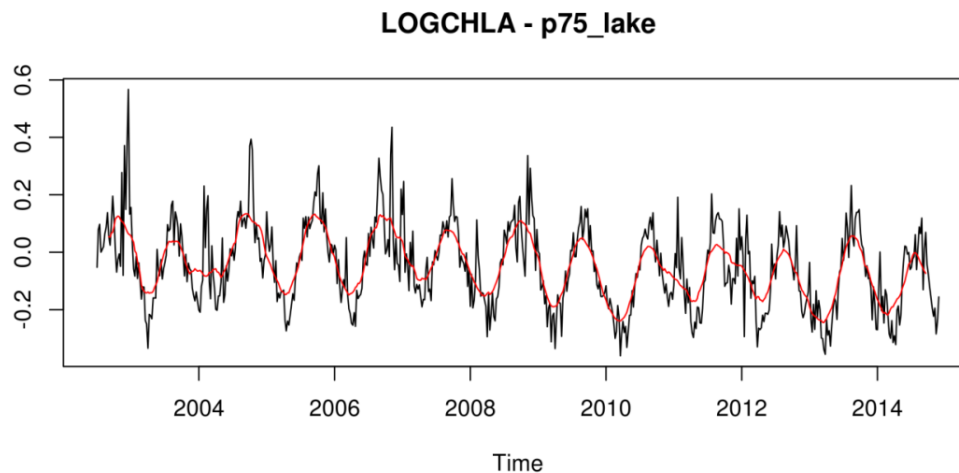

Fig. 24. Weekly aggregated TS of log (chl-a concentrations) with a moving average of 13 weeks is shown in red. Aggregated statistics shown is the third quartile of chl-a concentration values (OC3-aer_opt -10) for all clear sky pixels during the week for the entire lake.

The decreasing trend observed on Figure 24 in the satellite chl-a time series (2002-2012) for the entire lake was confirmed via Mann-Kendall test, which yielded a negative tau value of -0.17 (significant p-value << 2.5%). However, the change point detection algorithm AMOC (At-Most-One-Change) cut the TS at the beginning of 2009, with a reduction of the mean chl-a concentration detected between the two successive periods based on raw chl-a values (not log-transformed). Kendall tau values were non-significant when applied separately on these two sub-periods. This decrease in chl-a concentrations can be real (and related to natural phenomena); however, we cannot exclude the possibility that this observation may be due to a change associated with the MODIS sensor.

Observations of planktonic blooms from RS chl-a images

The daily times series were analyzed when cloud conditions permitted so. Three types of blooms were identified: coastal, southern seasonal and pelagic planktonic blooms.

Phytoplanktonic **blooms in coastal areas** were often observed along the shoreline. Five coastal areas (named « A ») were identified in the north of the lake (Figure 25). Littoral water presented higher planktonic concentrations on average compared with pelagic waters. However, chl-a concentrations in coastal areas were variable and

sometimes displayed peaks ("blooms"). During these blooms, concentrations of chl-a may be estimated as > 5 µg/l. However, precise concentrations of chl-a may not be determined from RS images in those areas as water conditions are not always clear and may be influenced by turbidity during rainfalls periods (Figure 26).

Coastal planktonic "blooms" also occur in other locations other than the "A" sites along the shore of the lake. Their occurrence between each degree of latitude is presented in Figure 26.

How can these coastal blooms be explained? Those sites may be influenced by rainfall, river input and anthropogenic impacts. However, those are not the only sources of nutrients in coastal areas. Hydrodynamics (e.g., internal waves, Kelvin wave) driven by meteorological conditions play a very important role in the mixing of lake waters and frequently allow access of nutrient-rich deep waters towards the surface of the lake, including the coastal areas (Chitamwebwa, 1999 ; Plisnier et al. 1999 ; Naithani & Deleersnijder 2004).

A1: Rusizi river inflow. Nutrients from the river may be mixed at the east as a result of a generally predominant clockwise current circulation along the coast (Coulter, 1968). The impact of the nearby city of Bujumbura probably also contributes to the nutrient input in this area, particularly during rainfall periods.

A2: Ubwari bay. The southern end of this bay is very shallow. More frequent water mixing probably induces higher phytoplankton biomass. Rainfall water may also have a greater impact on the nutrient concentration in the waters of the bay.

A3: Kigoma area. Nutrients from the populated area could affect the productivity of the coastal area, particularly during rainfall periods.

A4: Malagarazi River inflow into the lake.

A5: Kalemie area and impact that nutrients from populated areas, particularly during rainfall periods, have on the lake.

Fig. 25. Map with location of five coastal areas « A » displaying frequent planktonic blooms.

Seasonal southern phytoplankonic blooms.

Southern planktonic blooms are observed each year, at trade wind time change period around April/May (Figures 23 & 26). Important surface waves are observed along the south/eastern side of the lake. Those waves are well known by local fishermen and probably correspond to the local «Yala wave» events (Plisnier & Coenen, 2001). They induce mixing of deep nutrient rich water and phytoplanktonic blooms.

There is sometimes a second phytoplanktonic bloom period after approximately three weeks, probably associated with rhythmic movement of deep internal waves. Blooms identified in June of 2008 and 2014 (Figure 23) are probably the second occurrences of blooms linked to internal waves periods. First occurrence of blooms in those years could not be detected probably because of cloudiness and lack of images.

Those southern seasonal blooms are linked to increased catch of clupeids in the south and sometimes migrations of *Stolothrissa tanganicae* (Plisnier et al., 2009). Both cases were observed during the CHOLTIC monitoring (Figure 6). At the same moments, many fishermen may migrate towards the south. This might be linked to cholera epidemics, such as that on Mutondwe Island near Mpulungu, as reported by local fishermen (pers. comm).

Other seasonal blooms in the same areas were also observed in the period September-December (Figure 26) and might correspond to the local «Chimbanfulla» waves linked to seasonal changes (Plisnier & Coenen, 2001). However, it is probable that the number of those events observed from RS images is underestimated because of cloudiness and lack of images in this period in the south. Fishermen have indicated the regular occurrence of those blooms around September as during the months of April to June. Although seasonal southern blooms generally start along the eastern coast, they often expand into the pelagic area. This was well observed in 2004 and 2014 (Figure 23).

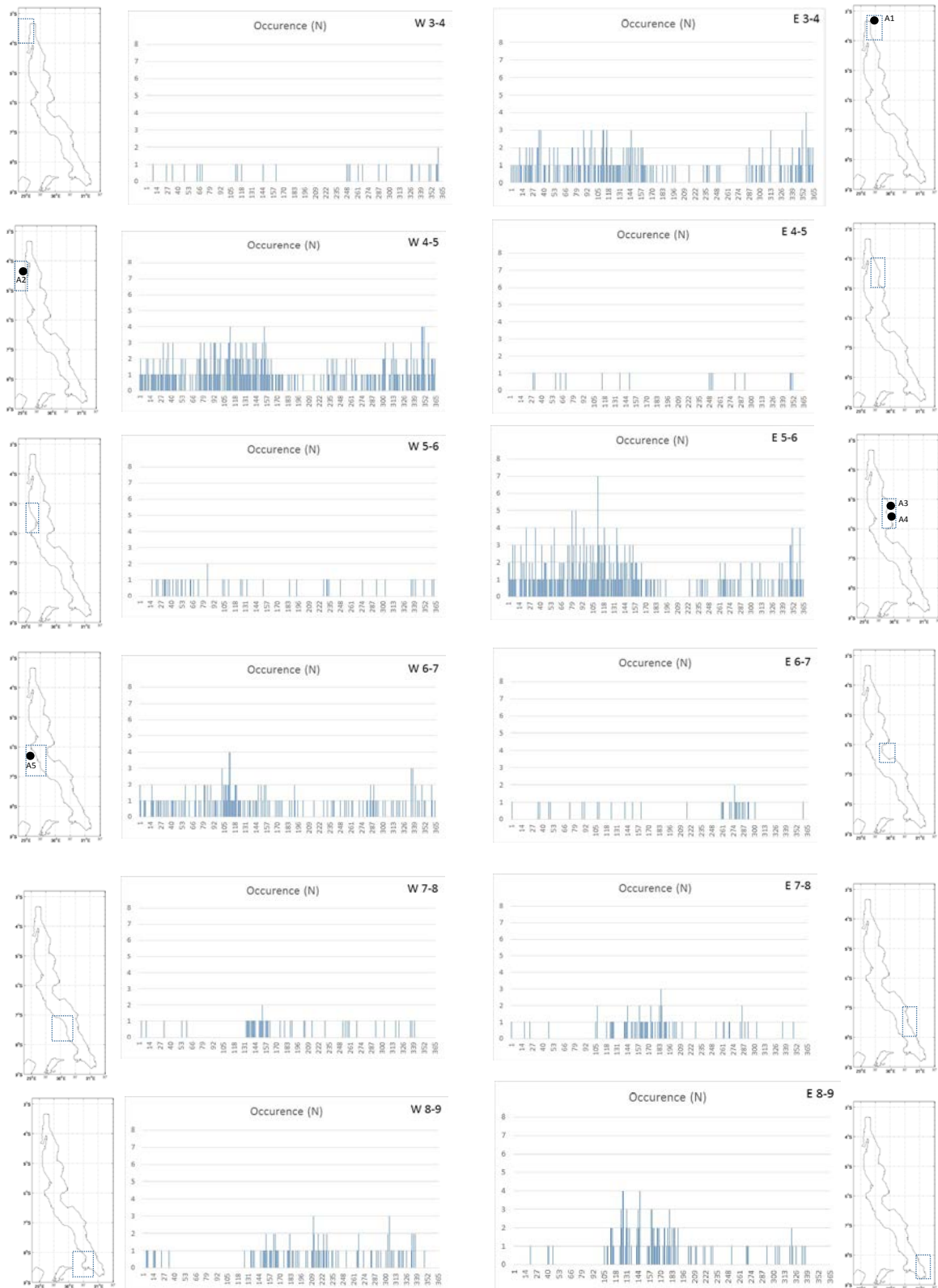## Coastal planktonic blooms occurence

Fig. 26. Phytoplanktonic bloom occurrences over an average yearly cycle (2002-2014) for each degree of latitude. Coastal blooms at «A» sites are indicated.

Pelagic phytoplanktonic blooms may be developed in any area of the lake. When this happens, coastal blooms are often also observed. Some examples are presented in Figure 27. In the north of the lake, the occurrence and intensity of pelagic phytoplanktonic blooms is more important during the period from September to December. No such blooms were observed during the dry period (May to August) in the north, although this is difficult to verify due to frequent cloudiness.In the south of the lake, the occurrence and intensity of pelagic blooms clearly take place at different periods (Figure 27). This is often related to the seasonal southern blooms starting along the southeast coast in April/May and probably also from September to December (no available images).Pelagic phytoplanktonic blooms seem to persist during a few days only, although in June 2014, it persisted for more than two weeks.
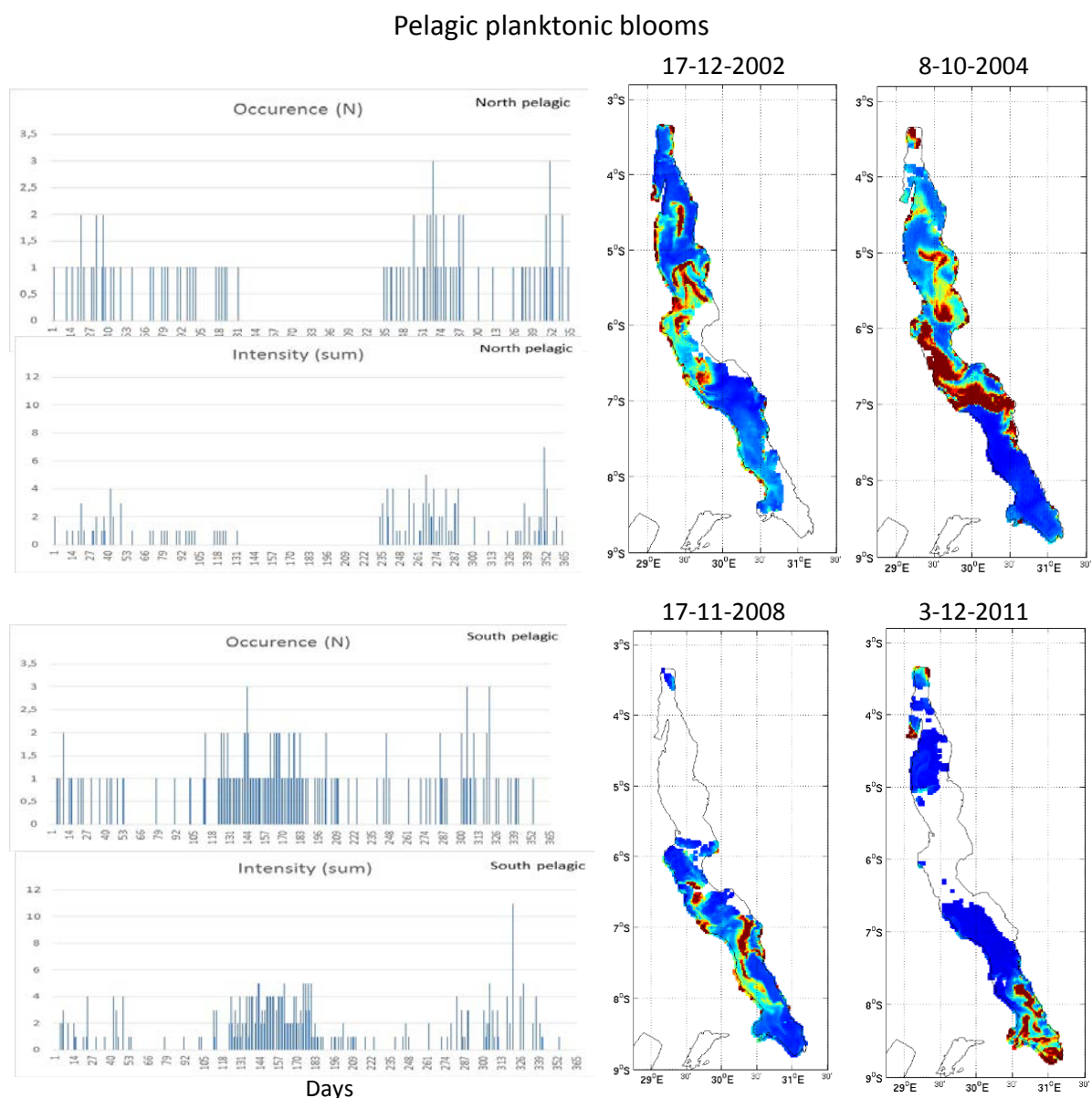

Fig. 27. Daily occurrence and intensity during the period 2002-2014 for the northern and southern parts of the lake. Four examples of extended pelagic blooms are presented.

WP 4.2 Lake Surface Water Temperature (LSWT)

Derivation of new LSWT algorithm. We confirmed that ARC-Lake could be used as a reference dataset to produce regional coefficient by validating against in situ measurements. Using the MODIS/ARC-Lake match-up, default MODIS SST was compared to the ARC-Lake dataset. From a detailed analysis of the residuals, we decided to filter MODIS data using a threshold of 0.05 K on the local standard deviation in a 3-by-3 window on brightness temperature in channel 11 µm (BT11) and channel 12 µm (BT12), before the derivation of new coefficients. This threshold is based on nighttime residual temperatures and enabled us to retain enough points for analysis. In the same manner, we defined a threshold of 40° on satellite zenith angle. Figures 28 and 29 compare ARC-Lake to MODIS temperatures. It can be noted that the association is much better for nighttime. Yearly coefficients already give good results and monthly coefficients do not substantially improve the quality of the association. Only one set of yearly coefficients has thus been used for the production of final dataset.

Fig. 28. Scatterplots correlating the daytime temperature of matching observations of ARC-Lake (X-axis) and MODIS Aqua (Y-axis) for. Left: default MODIS SST; center: MODIS LSWT computed using monthly-derived coefficients; right: MODIS LSWT computed new yearly-derived coefficients (i.e., same coefficients for the whole series). Same threshold local standard deviation on BT11 and BT12 have been used but no threshold on satellite zenith angle.
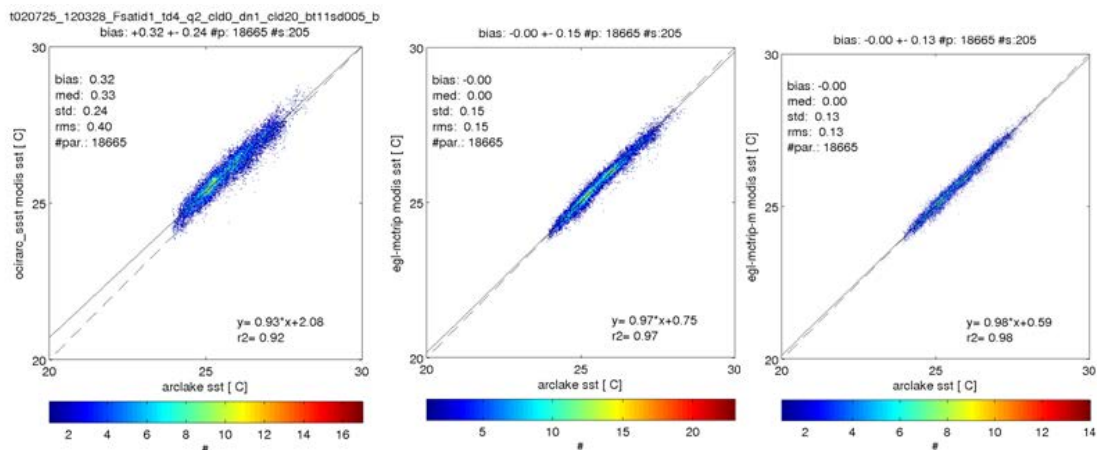
Fig. 29. Same as figure 28 but with nighttime temperature.

Final dataset. The final products in **raster** format give LSWT at spatial resolution of 0.02 degrees for entire lake with a daily and weekly temporal resolution for the period July 2002 – March 2014 computed with new derived algorithm for Lake Tanganyika. Daytime and nighttime temperatures were handled separately. Dataset provided to other teams are from MODIS Aqua so the corresponding acquisition time is ~11:30AM UTC for daytime and ~11:30PM UTC for nighttime, which is concomitant with the chl-a series. Tables containing weekly spatio-temporally aggregated TS for the entire lake and predefined ecoregions were also produced (Figures 30 and 31).
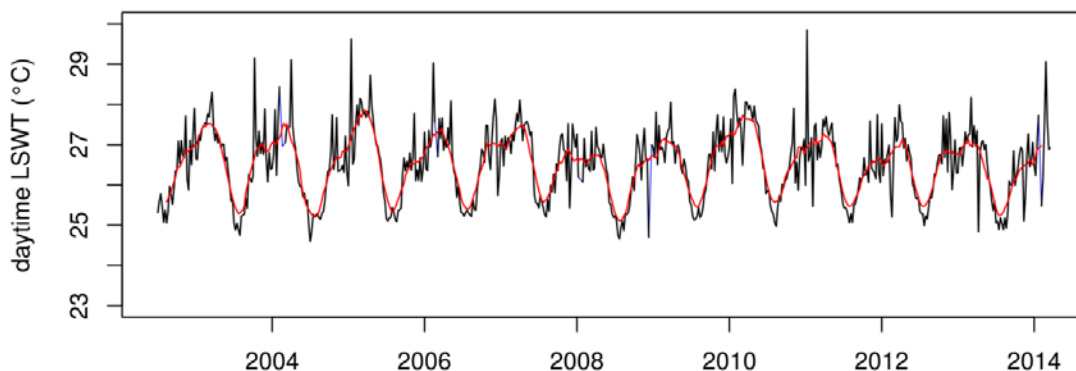
median of LSWT - Lake Tanganyika

Fig. 30: Weekly aggregated TS of daytime LSWT in °C with moving average of 13 weeks (in red). Aggregated statistics is the median LSWT of all clear sky pixels during the week for the entire lake.
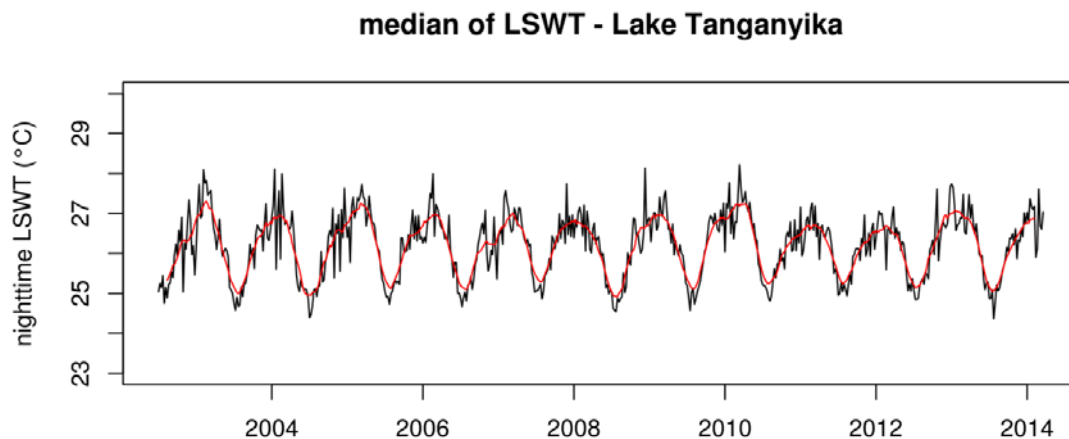
median of LSWT - Lake Tanganyika

Fig.31: Same as figure 30 but for nighttime LSWT.

WP 4.3 Landcover maps of the three epidemiological test zones

Landsat images were selected for the landcover mapping of the three epidemiological sites (Table III). Images were selected to minimize the area covered by **clouds**. Sometimes two images were used for the same area to this purpose. Remaining clouds were masked. **Geometric consistency** of all images covering the same site was insured using as a reference the panchromatic band at spatial resolution of 15 m (either from Landsat-7 or 8). The images of **Mpulungu** (south of Lake Tanganyika) allowed a diachronic (seasonal) classification strategy. More details about strategy, classification rules and validation are available in De Smet D'Olbecke (2014). Two classifications were carried out separately on each date (September and May) and then combined. Its final common legend contains the following 13 classes: bare/urbanized soil, dispersed miombo/cropland, shifting cropland, dense miombo, itigi, permanent dense forest, deforested land, dambo, temporary wetland, floodplain wetland, lake, cloud, shadow.

For **Uvira** and **Kalemie**, due to the lower seasonality effect on vegetation, only dry season images were used in the classification. The final legend consists in eight classes: bare/urbanized soil, bare soil/duricrust, dispersed savanna woodlands and cropland, open woodland, dense forest, water, cloud, shadow. The landcover maps produced were foreseen to analyze the association between **cholera** case locations and environmental context (more particularly the humid areas). Unfortunately, the localization of the cholera cases at the scale of the 'health area' (where they were recorded but not necessarily where people live) did not allow this kind of analysis, which has however been demonstrated to be very efficient by Ali et al. (2002). A better geolocalization of cholera cases could be introduced in future epidemiological surveys.

Table III: List of medium resolution images used for landcover mapping.

Site	Date	Satellite /Instrument	Path	Row	Use
Mpulungu	17-05-09	Landsat-5 TM	171	66	Classification
Mpulungu	19-09-08	Landsat-5 TM	171	66	Classification
Mpulungu	27-09-02	Landsat 7 ETM+	171	66	Geometry
Uvira	29-07-13	Landsat-8 OLI	173	62-63	Classification + Geometry
Uvira	01-08-14	Landsat-8 OLI	173	62	Classification
Uvira	13-07-13	Landsat-8 OLI	173	63	Classification
Kalemie	02-07-09	Landsat-8 OLI	173	64	Classification + Geometry
Kalemie	25-06-09	Landsat-8 OLI	172	64	Classification

Final remark on datasets produced by WP 4

All raster files, snapshots and spatio(-temporally) aggregated TS were shared amongst the project teams). The numerous figures and tables essential to support our methodological choices and ensure the reproducibility of our research and produced during calibration/validation phase of RS products computation were also available for projects team members. Regarding the landcover mapping, the Master thesis of De Smet D'Olbecke (2014) was also provided for the use of the project.

WP 5 Hydro-eco modelling

The Eco-Hydrodynamic model developed during the CLIMLAKE and CLIMFISH projects (Naithani et al., 2002; 2003; 2007a; 2007b; 2011; 2012) was further modified to include the dominant phytoplankton groups present in the lake considered important for the project. The model includes hydrodynamic and thermodynamic components and an ecological module. The hydrodynamic/circulation model considers the lake to consist of two homogeneous layers of different density, representing the warm epilimnion (surface mixed layer) and cold dense hypolimnion (lower layer) separated by a thermocline. The lower layer is considered to be much deeper than the surface active/mixed layer. The hydrodynamic model equations are the following:

$$\frac{\partial \xi}{\partial t} + \frac{\partial (Hu)}{\partial x} + \frac{\partial (Hv)}{\partial y} = w_e \tag{1}$$

$$w_e = \left(\frac{3}{20}\right)^{1/2} \frac{(\tau_x^2 + \tau_y^2)^{1/2}}{(\varepsilon g H)^{1/2}} - w_d - \frac{\xi}{r_{tt}} \tag{2}$$

$$\frac{\partial (Hu)}{\partial t} = -\frac{\partial (Huu)}{\partial x} - \frac{\partial (Hvu)}{\partial y} + fHv - gH\frac{\partial \varepsilon \xi}{\partial x} + \frac{\partial}{\partial x}\left(HA_x \frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(HA_y \frac{\partial u}{\partial y}\right) + \frac{\tau_x}{\rho_0} + w_e^- u \tag{3}$$

$$\frac{\partial (Hv)}{\partial t} = -\frac{\partial (Huv)}{\partial x} - \frac{\partial (Hvv)}{\partial y} - fHu - gH\frac{\partial \varepsilon \xi}{\partial y} + \frac{\partial}{\partial x}\left(HA_x \frac{\partial v}{\partial x}\right) + \frac{\partial}{\partial y}\left(HA_y \frac{\partial v}{\partial y}\right) + \frac{\tau_y}{\rho_0} + w_e^- v \tag{4}$$

where x and y are horizontal axes, u and v are the depth-integrated velocity components in the surface layer in the x and y directions, t is the time, ξ is the downward displacement of the thermocline, $H = h + \xi$ is the thickness of the epilimnion (the surface, well-mixed layer), h is the reference depth of the surface layer (m) and w_e is the entrainment velocity (m s^{-1}). The first term on the right hand side of Eq. 2 was inspired by Price (1979), τ_x and τ_y are horizontal components of specific wind stress in the x and y direction (m^2 s^{-2}), $\varepsilon = (\rho_b - \rho_s)/\rho_b$ is the relative density difference between the hypolimnion (ρ_b) and the epilimnion (ρ_s), calculated using the temperature of the surface layer (t$_s$) and bottom layer (t$_b$), respectively. w_d is the detrainment term (m s^{-1}), defined such that the annual mean of the epilimnion volume remains approximately constant. There are large uncertainties in the parameterization of entrainment and detrainment terms. As a consequence, to avoid occasional spurious values of ξ, a relaxation term (ξ/r_{tt}) is needed, which slowly nudges the surface layer depth toward its equilibrium position. The relaxation timescale, r_{tt}, is sufficiently long so that the relaxation term is generally smaller than the entrainment and detrainment terms. f is the Coriolis factor (<0 in the southern hemisphere), A_s is the horizontal eddy viscosity in the s $(=x,y)$ direction, $w_e^- = \frac{w_e - |w_e|}{2}$ is the negative part of the entrainment velocity, i.e. w_e^- is equal to w_e if $w_e < 0$ and is zero otherwise, $w_e^+ = \frac{w_e + |w_e|}{2}$ is the positive part of the entrainment velocity.

The surface layer temperature is predicted using the equation:

$$\frac{\partial (H\theta)}{\partial t} + \frac{\partial (Hu\theta)}{\partial x} + \frac{\partial (Hv\theta)}{\partial y} = \frac{\partial}{\partial x}\left(HK_x \frac{\partial \theta}{\partial x}\right) + \frac{\partial}{\partial y}\left(HK_y \frac{\partial \theta}{\partial y}\right) + w_e^+ \theta_h + w_e^- \theta + H\frac{(\theta_s - \theta)}{r_{ts}} \tag{5}$$

where θ is the surface layer temperature, θ_s is the reference temperature of the surface layer, θ_h is the temperature of the hypolimnion water and r_{ts} is the relaxation time scale for surface fluxes.

The original ecological model version is a typical NPZD model (Nutrient Phytoplankton Zooplankton Detritus) with phosphate as the nutrient and copepods as zooplankton.

The modified ecological model now contains four main phytoplankton groups (chlorophytes (CHL), dinophytes (DIN), diatoms (DIA) and cyanobacteria (CYA)), three nutrients (phosphate, nitrate and silica) and copepods as zooplankton. The ecosystem model is presented in Figure 32.
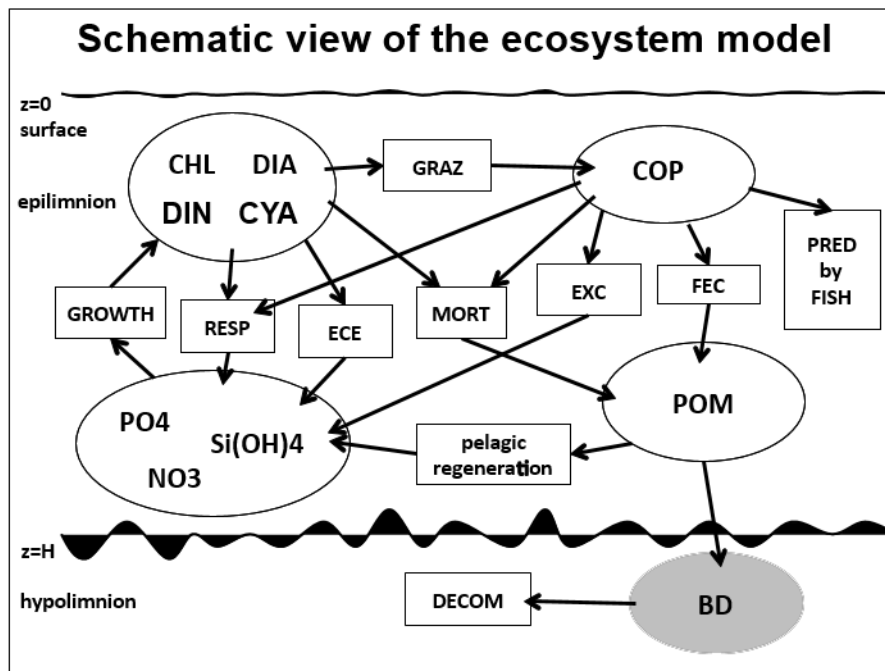

Fig. 32: Schematic view of the ecosystem model.

The phytoplankton processes includes primary production (*PROD*), respiration *(RESP)* and mortality (*MORT_{PHYTO}*). The processes concerning the zooplankton are grazing (*GRAZ*), fecal pellet (*FEC*) egestion, excretion (*EXC*) and mortality (*MORT_{ZOO}*). Phytoplankton respiratory release and the excretion from zooplankton are directly remineralized in the surface layer. A small percentage of feces, dead phytoplankton and zooplankton are also remineralized into phosphate in the surface layer, whereas the rest contributes to the detritus pool, which sediments fast. The regeneration/remineralization within the surface layer represents the effect of the microbial food web and also represents the pelagic regeneration. The model is closed by predation *(PRED)* from zooplanktivorous fish and sinking of detritus out of the surface layer. The zooplanktivorous fish biomass is assumed equal to that of zooplankton biomass (Sarvala et. al., 1999).

Biochemical concentrations are computed using an advection diffusion equation with biological terms (production and destruction) modeling the interaction between different species. The general equation describing a nonconservative variable is defined as:

$$\frac{\partial (H\,VAR)}{\partial t} + \frac{\partial}{\partial x}\left(Hu\,VAR - Hk_x \frac{\partial VAR}{\partial x} \right) + \frac{\partial}{\partial y}\left(Hv\,VAR - Hk_y \frac{\partial VAR}{\partial y} \right) + \phi_{he} = H\left(Biological\ terms \right)_{VAR} \quad (6)$$

where VAR can be any model dependent variable, such as PHYTO, ZOO, nutrients and POM. The terms on the left hand side of the equation represent the rate of change of the variable, horizontal advection and diffusion of the ecological variables. u and v are time-dependent horizontal velocities obtained from the circulation model, K_X and K_y are the horizontal diffusion coefficients. The sixth term represents entrainment from hypolimnion. $\phi_{he} = w_e^1 VAR_h + w_e^- VAR$. VARh is the entrainment from the hypolimnion. The right hand side of the equation includes all the biological processes contributing to the growth and loss of the VAR. Biological variables (except for nutrients) are expressed in units of concentration of carbon (μgCL^{-1}). Biological rates are expressed as $\mu gCL^{-1}d^{-1}$.

Biological terms in the phytoplankton equation are:

$$\left(Biological\,terms \right)_{PHYTO} = GROWTH_{PHYTO} - RESP_{PHYTO} - ECE_{PHYTO} - MORT_{PHYTO} - GRAZ_{PHYTO} \tag{7}$$

Phytoplankton growth rate, $GROWTH_{PHYTO}$, is considered to be influenced by nutrients, light intensity and temperature.

$$GROWTH_{PHYTO} = P_{mPHYTO} * \min(F(N), F(I)) * PHYTO \tag{8}$$

where P_{mPHYTO} is the maximum photosynthetic rate (d^{-1}). $F(N)$ is the limitation to growth by minimum nutrient (nitrate, NO_3^- or phosphate, PO_4^{3-} or silicate, $Si(OH)_4$). $F(I)$ is the light limitation to photosynthesis. The effect of nutrients, $F(N)$, on photosynthesis is modeled according to Michaelis-Menten formulation. The growth is considered to be limited by the single nutrient, which displays the lowest concentration to the half saturation constant for uptake of the respective nutrient. The nutrient uptake is defined as:

$$F(N) = \min \left(\frac{NO3}{NO3 + K_{NO3PHYTO}}, \frac{PO4}{PO4 + K_{PO4PHYTO}}, \frac{Si}{Si + K_{SiPHYTO}} \right) \tag{9}$$

$K_{NO3PHYTO}$, $K_{PO4PHYTO}$, $K_{SiPHYTO}$ are the half saturation constants for NO3, NH4, PO4 and Si uptake by phytoplankton, respectively. Silica limitation acts only on diatoms.

For primary production, the shortwave intensity at the surface is converted to the photosynthetically active component (400-700 nm) by assuming it composes 56% (alpha = .56) of the incident spectrum (Jorgensen and Bendoricchio, 2001). Light limitation, $F(I)$, is modeled as an exponential decrease of light intensity with depth. F(I) is defined as:

$$F(I) = \frac{1}{k_e H} \left(\arctan \frac{\alpha I_0}{2I_k} - \arctan \frac{\alpha I_0 \exp^{(-k_e H)}}{2I_k} \right) \tag{10}$$

ke is composed of a dominating background extinction and the contribution from SPM.
ke = ke1 + ke2 * PHYTO.

Respiration, RESP, is divided into maintenance/basic and growth/photo respiration. The former being a function of the maximum photosynthetic rate (0 - 10 %) and the latter depends upon the production (30 - 55 %). RESP of phytoplankton is defined as:

$$RESP_{PHYTO} = \left(RESP_{b0} + RESP_{p0} * P_{mPHYTO} * \min(F(N), F(I)) \right) * PHYTO \tag{11}$$

Mortality, MORT, is the loss of phytoplankton by natural death and is defined as a logistic equation and depends upon temperature. $MORT_{PHYTO} = m_{PHYTO} * PHYTO$, where m_{PHYTO} is the mortality rate.

Grazing, GRAZ, is the loss of phytoplankton because of grazing by zooplankton. It is the Ivlev (1945) equation as modified by Parsons et al. (1967). It is described with a temperature-dependent term (Q_{10}) and an Ivlev equation with a feeding threshold. Grazing ceases when the phytoplankton concentration is below this threshold value and saturates when it is sufficiently large.

$$GRAZ_{PHYTOZOO} = \max\left(0, g_{\max PHYTOZOO} * \left[1 - \exp^{-\lambda*(PHYTO_{\min} - PHYTO)} \right] \right) * ZOO \tag{12}$$

$g_{\max PHYTOZOO}$ is the maximum grazing rate. Equations similar to (6) are written for CHL, DIN, DIA and CYA.

The biological terms in the zooplankton equation are:

$$\left(Biological\ terms \right)_{ZOO} = GRAZ_{PHYTOZOO} - EXC_{ZOO} - FEC_{ZOO} - MORT_{ZOO} \mp GRAZ_{ZOOZOO} \tag{13}$$

The first term on the right hand side is the grazing of phytoplankton by zooplankton, second term represent the metabolic excretion, third term formulates egestion of fecal pellets by zooplankton, forth term is the natural mortality and the last term is the predation on zooplankton by zooplanktivorous fish. Excretion and fecal pellet egestion are expressed as a linear function of grazing (plus predation, where applicable). Grazing by zooplankton is divided into three parts: their growth, excretion and fecal pellets, all proportional to grazing.

Excretion is defined as:

$$EXC_{ZOO} = n_{eZOO} * \left(GRAZ_{PHYTOZOO} + PRED_{ZOOZOO} \right) * ZOO \tag{14}$$

Egestion of fecal pellets is defined as:

$$FEC_{ZOO} = n_{fZOO} * \left(GRAZ_{PHYTOZOO} + PRED_{ZOOZOO} \right) * ZOO \tag{15}$$

Mortality of zooplankton is defined with similar expression as that for phytoplankton.

The nutrient equations include the uptake by phytoplankton, the metabolic loss terms of all biological variables, a percentage of their natural mortality, a percentage of feces of zooplankton, and the remineralization of detritus.

$$\left(Biological\ terms\right)_{PO4} = \sum_{PHYTO=1}^{4} \left(\frac{\left(-GROWTH_{PHYTO} + RESP_{PHYTO} + ECE_{PHYTO} + p_{mort}*MORT_{PHYTO}\right)}{R_{C:P}}\right) +$$
$$\sum_{ZOO=1}^{4} \left(\frac{\left(EXC_{ZOO} + p_{fec}FEC_{ZOO} + p_{mort}*MORT_{ZOO}\right)}{R_{C:P}}\right) + r_D \left(\frac{POM}{R_{C:P}}\right) \tag{16}$$

$$\left(Biological\ terms\right)_{NO3} = \sum_{PHYTO=1}^{4} \left(\frac{\left(-GROWTH_{PHYTO} + RESP_{PHYTO}\right)}{R_{C:N}}\right) \tag{17}$$

$$\left(Biological\ terms\right)_{Si} = \sum_{PHYTO=1}^{4} \left(\frac{\left(-GROWTH_{DIA} + RESP_{DIA} + ECE_{DIA} + p_{mort}*MORT_{DIA}\right)}{R_{C:Si}}\right) +$$
$$\frac{\left(EXC_{ZOO} + p_{fec}FEC_{ZOO} + p_{mort}*MORT_{ZOO}\right)}{R_{C:Si}} + r_D \left(\frac{POM}{R_{C:Si}}\right) \tag{18}$$

$R_{C:N}$, $R_{C:P}$ and $R_{C:Si}$ are carbon to nutrients ratios. These ratios can also be seen as the weight ratios for cells to convert the biological loss to return to nutrients.

Particulate organic matter or detritus (μgCL^{-1}), is formed mainly by dead organic matter and zooplankton feces, rest of what is not directly remineralized in the water column. Organic matter remineralizes with a constant remineralization rate. Rematerialized inorganic nutrients are released back into the water column.

$$\left(Biological\ terms\right)_{POM} = \sum_{PHYTO=1}^{4} \left(\left(1-p_{mort}\right)*MORT_{PHYTO}\right) + \left(1-p_{mort}\right)*MORT_{ZOO} + w_s*DIA$$
$$\left(1-p_{FEC}\right)*FEC_{ZOO} + \left(1-p_{mort}\right)*MORT_{ZOO} - r_D POM - w_d*POM \tag{19}$$

Equations are discretised on Arakawa's C grid. The model uses the forward-backward time stepping. The lake is represented with a rectangular Cartesian grid with $\Delta x = 6$ km and $\Delta y = 20$ km. The time step is 30 minutes. The first-year model run is not used for analysis. The model predicts mixed layer depth and the depth-averaged values of ecological variables for the mixed layer. For the Hydrodynamic model we use the wind forcing and solar radiation data from NCEP re-analysis-2. The ecological model was fully calibrated using field measurements of the CLIMLAKE/CLIMFISH projects (2002-2006).

WP 5.1 Recent model predictions

Most of the nutrients in Lake Tanganyika lie in the lower layer below 30 meters. These are entrained into the upper layer during the wind-induced upwelling due to the strong and persistent southeast winds during the dry season (from April/May until August/September) and cause phytoplankton blooms. Phytoplankton blooms associated with coastal upwelling, also occur from time to time propagating clockwise around the western boundaries of the lake. In the previous studies using the NPZD ecological model, it has been predicted that the dominant components responsible for

the phytoplankton biomass in the lake are temperature stratification, availability of light and nutrients. The solar radiation around the lake varies little in the year due to its close proximity to the equator. Because of low average chl-a (1 µg/l) in the surface pelagic waters, the lake is highly transparent with a Secchi disk depth transparency of approximately 12 m on average. Primary production in the nutrient depleted surface layer depends mainly upon the recycling of nutrients by wind-induced vertical mixing. The transport and mixing events are critical for the re-supply of nutrients for the primary productivity and biogeochemical processes in the stratified lake. It has been predicted that at shallow mixed layer depths with sufficient light conditions, the phytoplankton growth is limited by nutrients. Inversely, at greater depths associated to deeper mixing phytoplankton production is limited by light in spite of higher nutrient levels at such depths.

Figures 33 and 34 present the model forcing for the years 2001 to 2014 and model simulated surface layer depth, depth averaged concentration in the surface layer of chlorophytes (CHL), diatoms (DIA), dinophytes (DIN), cyanobacteria (CYA) and zooplankton (ZOO), respectively. Simulations near Mpulungu, Kigoma, and average for the entire lake are presented. Model successfully predicts the succession of phytoplankton groups in the lake. Chlorophytes are the first to appear because of their lower nutrient requirements, followed by diatoms, while cyanobacteria appear during the dry season and persist until the end of the year. The model predicts lower values for chlorophytes in the later half of the year near Kigoma and simulates very high concentration of chlorophytes in the northern extremity of the lake at the beginning of the year. This could be explained by lower nutrient and higher light requirements of chlorophytes in addition to relatively stable surface layer in the north allowing chlorophytes early development and growth. On the contrary higher nutrient levels in the south, during the dry and upwelling season allows cyanobacteria to develop in higher concentrations. One noticeable feature (Figure 33) is the time evolution of average dry season winds and air temperature. Average dry season winds started decreasing from the year 2007 until the year 2011. The air temperature increased initially and started decreasing from the year 2009 onwards and stayed lower until the year 2012. During this time the biomass of chlorophytes and diatoms increased, while that of cyanobacteria decreased slightly (Figure 34).
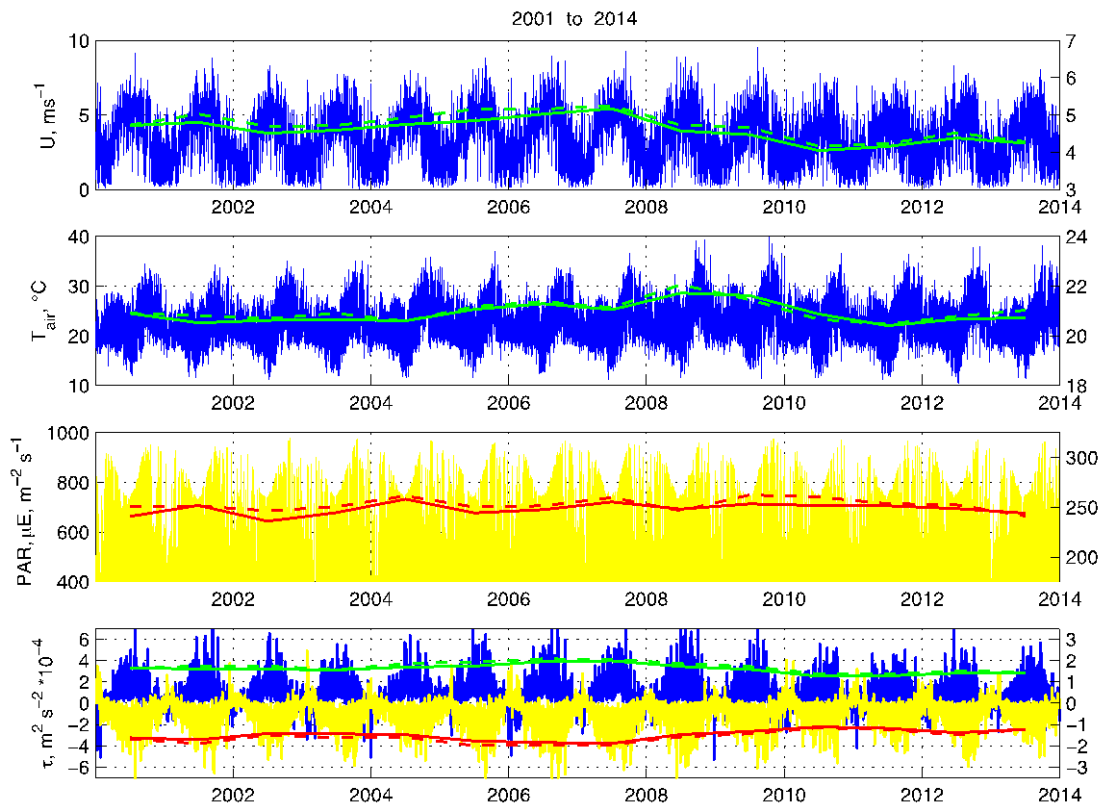
Fig 33: Time series of daily averaged values (2001-2014) of the NCEP re-analysis-2 horizontal wind speed (U), air temperature (T°), photosynthetic active radiation (PAR) and the x and y-components of wind stress. The plots include the wet season average of variables with the axis on the right (solid line: May-Sept average; dashed line: June-Sept average).
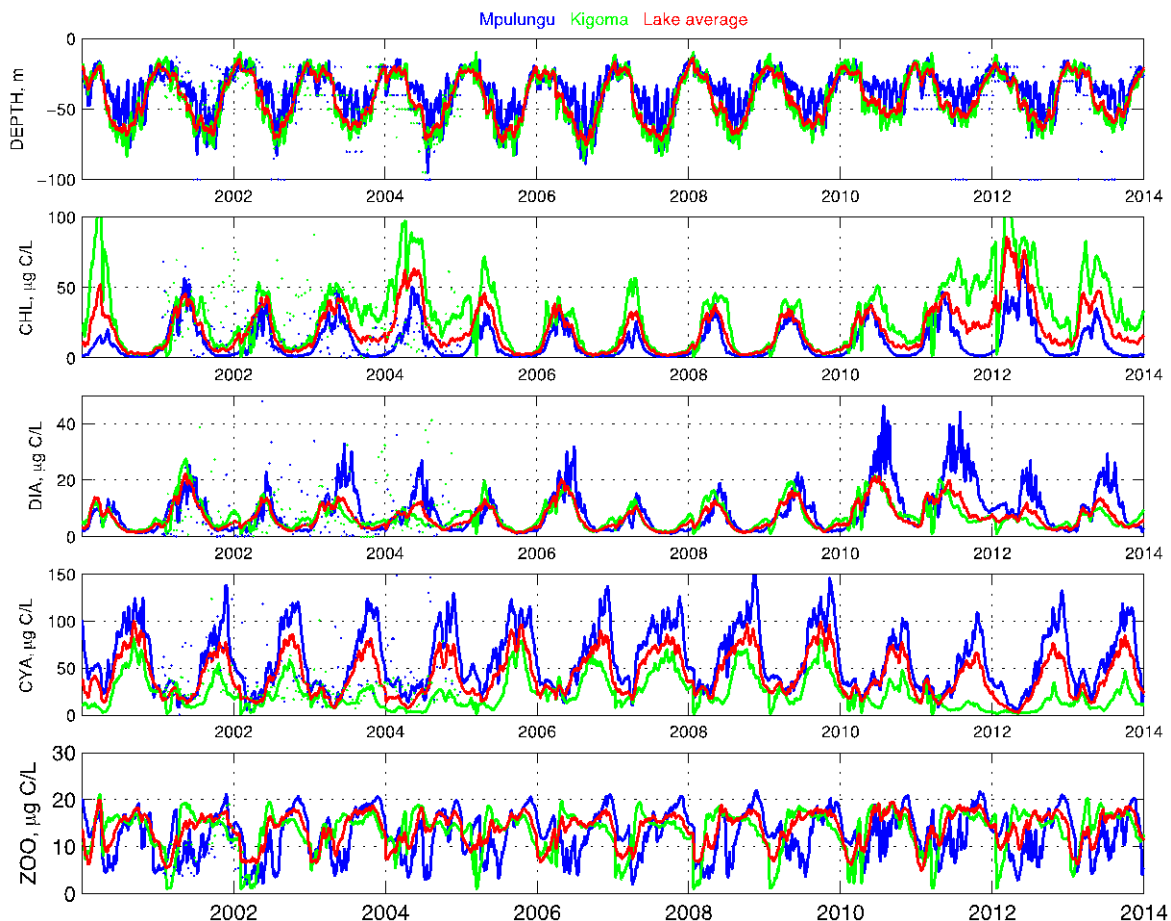
Fig 34. Time variation of model simulated surface layer depth, depth averaged concentration in the surface layer of chlorophytes (CHL), diatoms (DIA), cyanobacteria (CYA) and zooplanktons (ZOO) from 2001 until 2014 near Mpulungu (blue), Kigoma (green) and averaged for the entire lake (red). Solid lines are model simulations, and the dashed lines are measurements averaged over surface layer. Surface layer depth data acquired during the present CHOLTIC project near Mpulungu is also plotted.

W.P. 5.2 Climatological model run

Simulations were carried out for the years 1980 until 2013 to study the influence of climatic variability on plankton community in the lake. Figures 35 and 36 give the model forcing and simulations for the climatological run.
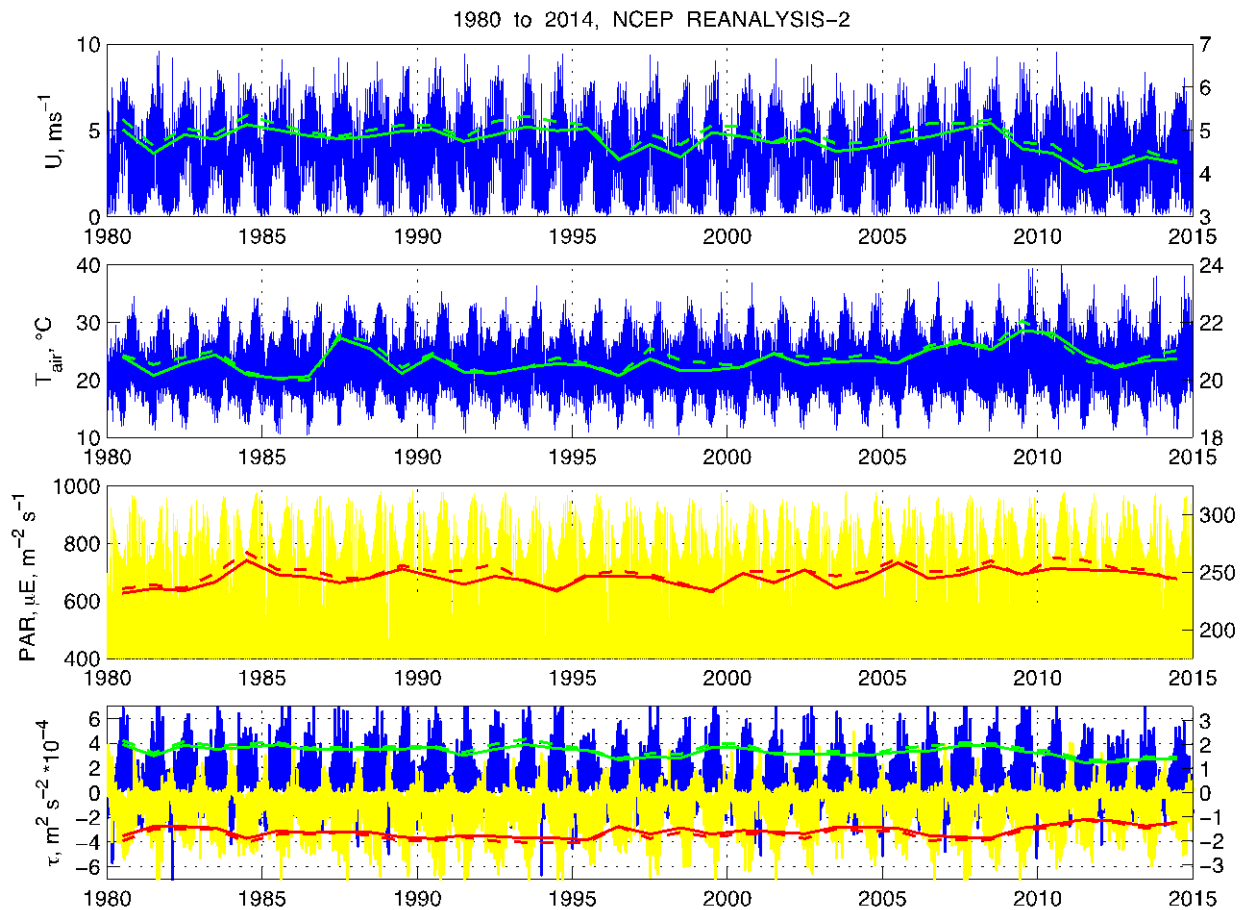
Fig. 35: Time series of daily averaged values of the NCEP re-analyzed horizontal wind speed, air temperature, photosynthetical active radiation, PAR and the x and y-components of wind stress for the years 1980-2014. The plots include the wet season average of variables with the axis on the right. The solid line is the average over May-Sept months, and the dashed line is the average over June-Sept months.
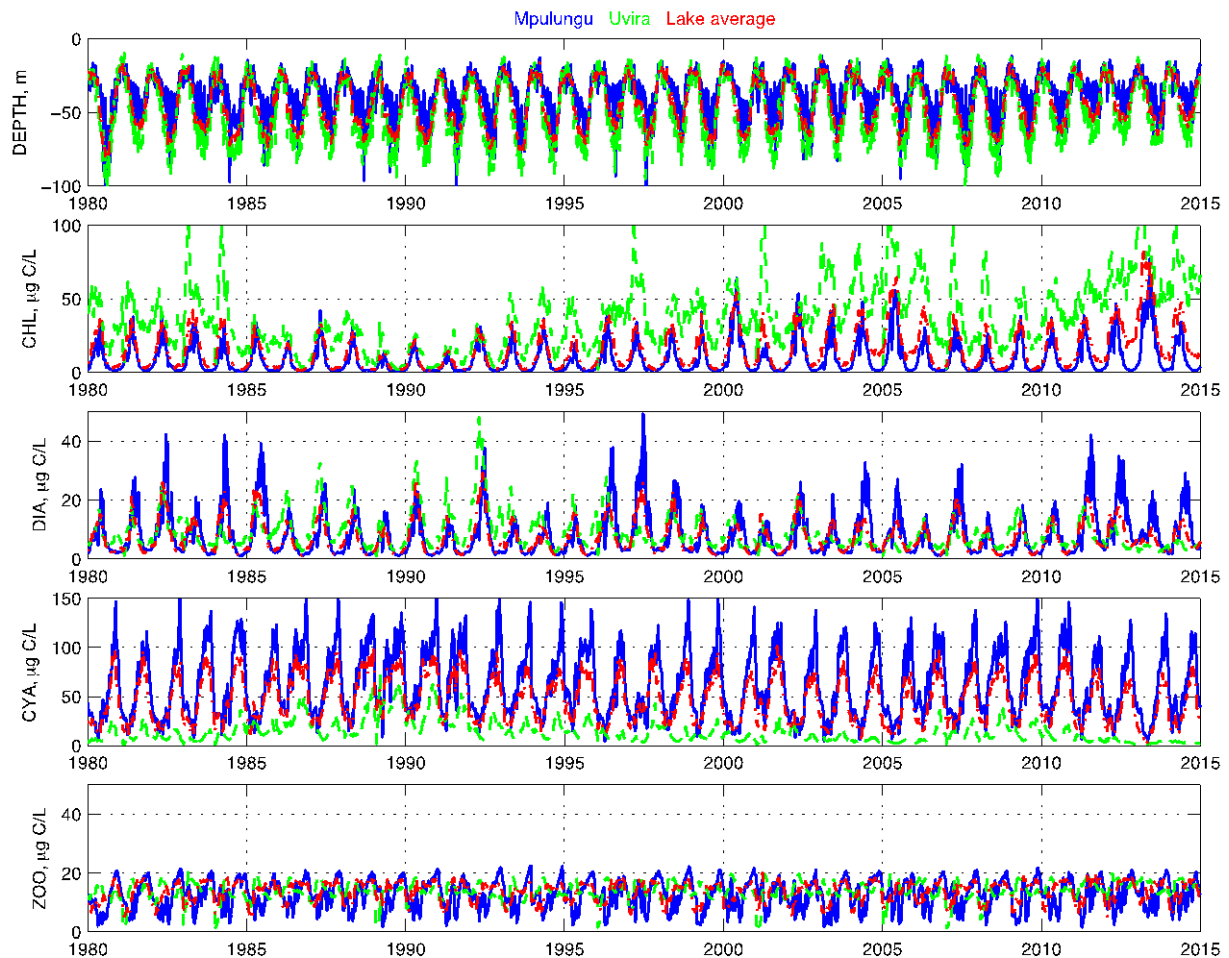
Fig. 36. Time variation of model simulated surface layer depth, depth averaged concentration in the surface layer of chlorophytes (CHL), diatoms (DIA), cyanobacteria (CYA) and zooplanktons (ZOO) from 1980 until 2014 near Mpulungu (blue), Uvira (green) and lake average (red).

In general, we found that the years with stronger wind are accompanied by increased biomass of diatoms and cyanobacteria, while low-wind years have higher biomass of chlorophytes. Stronger winds result in deeper mixing of water masses and help in the upwelling of nutrient rich hypolimnion water and increase the time spent by phytoplankton in the low light conditions of deep waters. Diatoms require high nutrients for their growth and can survive in low light conditions while chlorophytes require high light conditions and low nutrients. In the years when average wind during the dry season was rather high but with few strong wind events, the chlorophytes biomass remained high as well as the biomass of diatoms. However, there are some exceptions, for example, during the years 1993 and 1994, when the winds were high but the diatoms biomass was low. In 1996, when the winds suddenly dropped, the diatoms biomass rather increased along with the biomass of chlorophytes. Similarly, diatoms biomass did not increase with the increased winds of year 2008. During 1997 and 2011, the winds were not exceptional, although there were some high wind episodes and the diatoms biomass was high. Similar trend is seen in the evolution of

the biomass of phytoplankton in the northern basin near Uvira, except for some few exceptions, for example, during years 1985 and 2001, when the chlorophytes biomass was rather high near Uvira. The cyanobacteria remained more or less the same throughout, except 1997 and 2011 near Mpulungu when the diatoms increased tremendously and near Uvira during the years 2004-05 and 2011-2012. During these years, cyanobacteria biomass decreased.

Probable future scenarios: Climate change will presumably affect the ecosystem in the lakes by means of changing temperatures, surface layer depth and winds. These are ecologically important for the Lake Tanganyika ecosystem as they determine the depth of the surface mixed layer, the photosynthetically available radiation for primary production and the nutrient recycling in the surface mixed layer. Simulations are performed with varying reference depth of the surface layer (thermocline depth), its temperature and the wind stress to assess the sensitivity of Lake Tanganyika's plankton community to climate change.

Changes in the surface layer depth (Figure 37): Increasing the reference thermocline depth naturally increases the depth of the surface layer and entrainment of phosphate from below.

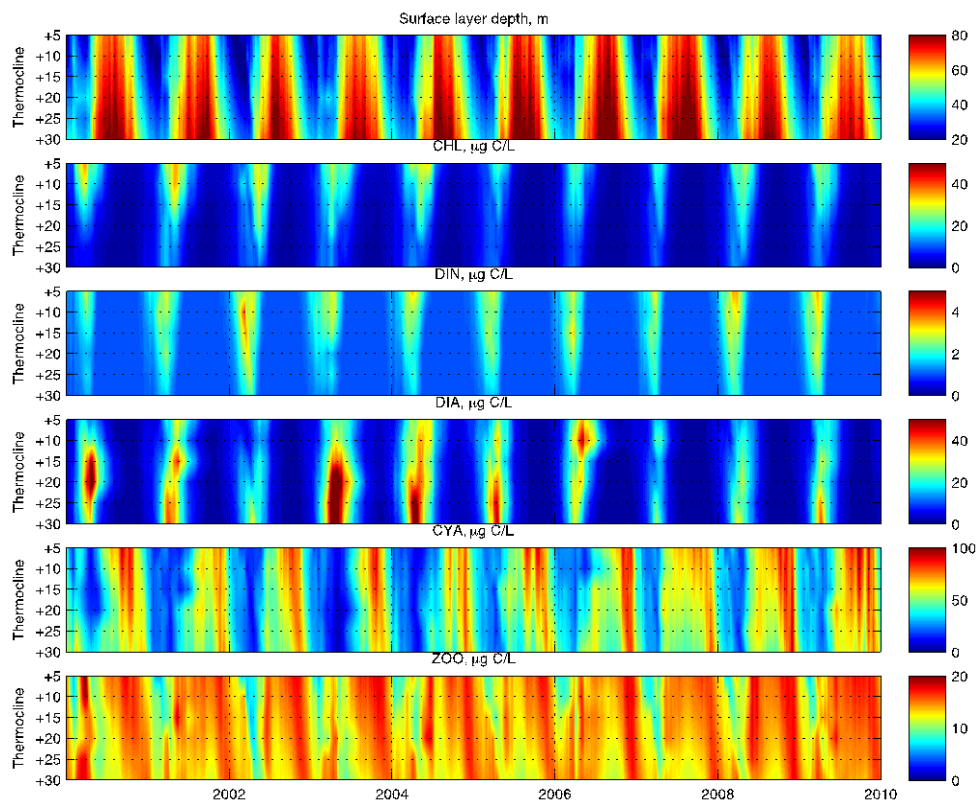

Fig. 37. Time variation of lake averaged surface layer depth (m) and the biomass of chlorophytes (CHL), dinoflagelates (DIN), diatoms (DIA), cyanobacteria (CYA) and zooplankton (ZOO) at various thermpocline depths for the period 2001-2010.

It decreases the chlorophytes and dinoflagelates. The biomass of diatoms increases with increasing reference depth of the thermocline and overcompetes the biomass of cyanobacteria.

Changes in the surface layer temperature (Figure 38): As the temperature of the surface layer increases, the biomass of chlorophytes increases. Dinoflagelates and diatoms remain more or less same, while cyanobacteria biomass decreases.
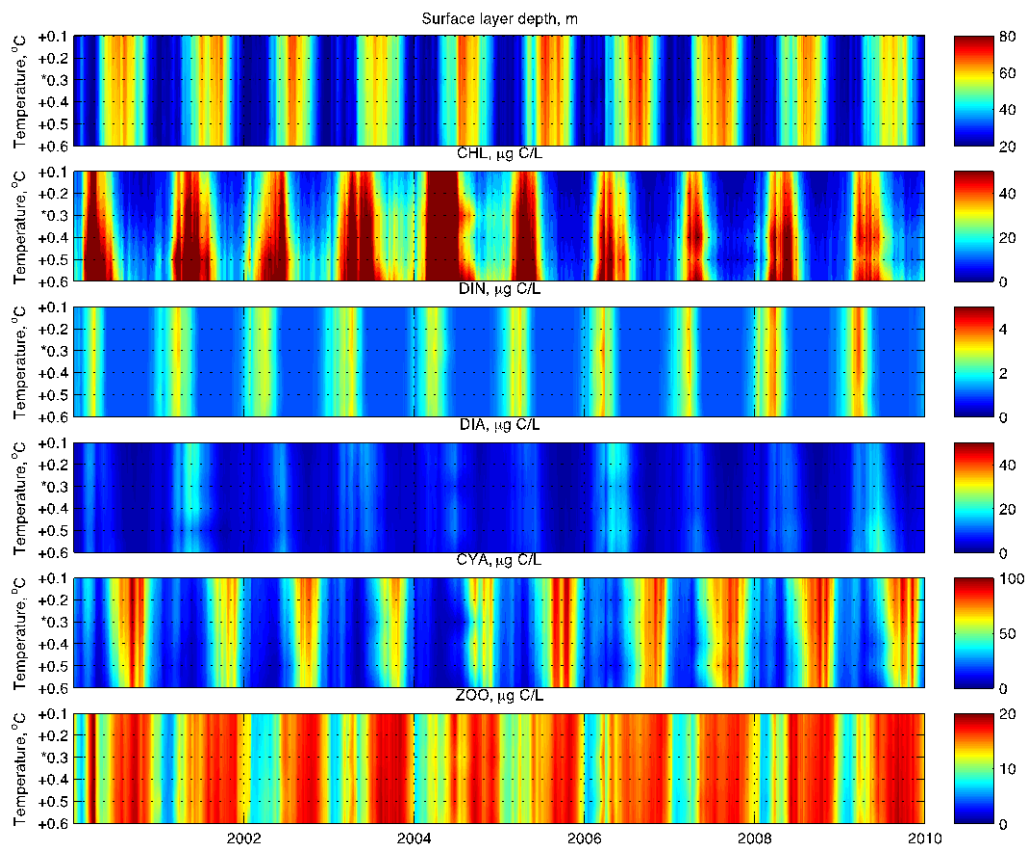

Fig. 38. Time variation of lake averaged surface layer depth and the biomass of chlorophytes (CHL), dinoflagelates (DIN), diatoms (DIA), cyanobacteria (CYA) and zooplankton (ZOO) at various temperatures of the surface layer.

Changes in the wind stress (Figure 39): The depth of the surface layer increases with increasing wind stress. The biomass of chlorophytes, dinoflagelates and diatoms decreases with increasing surface layer stress, while the biomass of cyanobacteria increases.

Changes in the surface layer depth and wind stress: Figure 40 presents the model simulations for varying thermocline depth and wind stress. Simulations were annually averaged over the entire lake. Different phytoplankton group react in different ways. Chlorophytes and dinoflagelates were abundant for low surface layer and low mixing conditions. Surprisingly, the biomass of diatoms stayed high for shallower surface

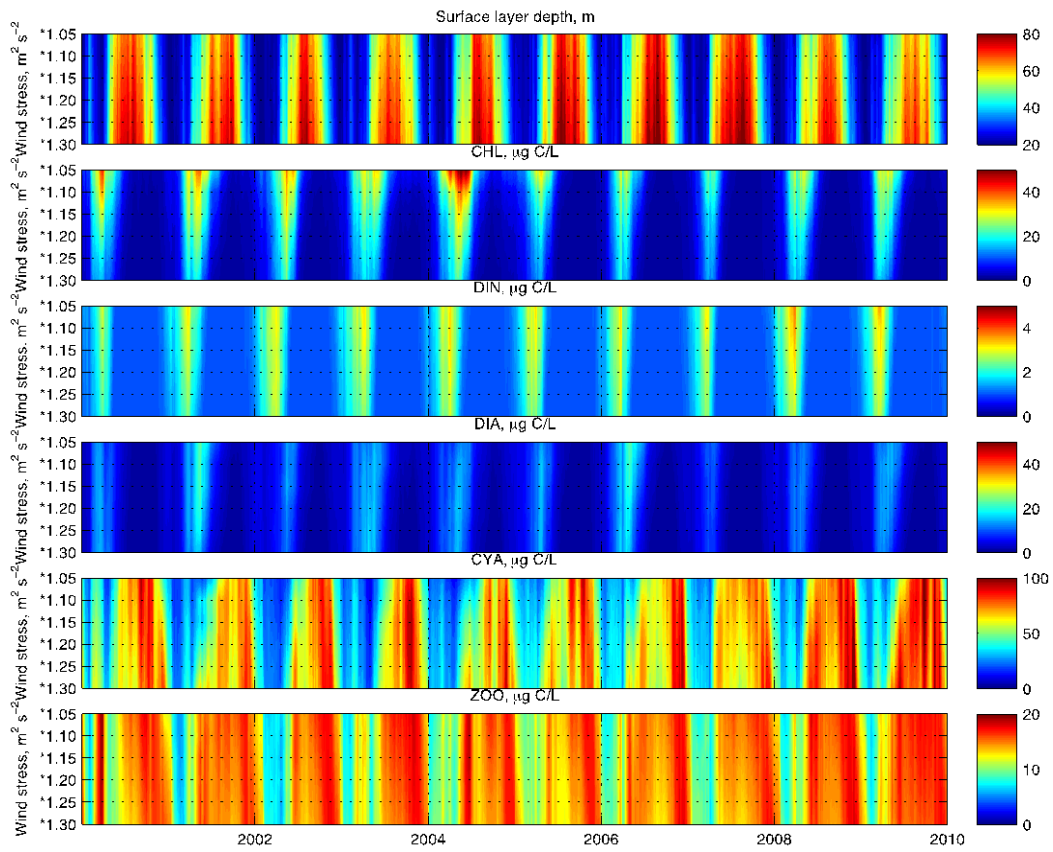layer and low mixing, while cyanobacteria remained high for all the depths and high mixing.

Fig. 39. Time variation of lake averaged surface layer depth and the biomass of chlorophytes (CHL), dinoflagelates (DIN), diatoms (DIA), cyanobacteria (CYA) and zooplankton (ZOO) at various wind stress.
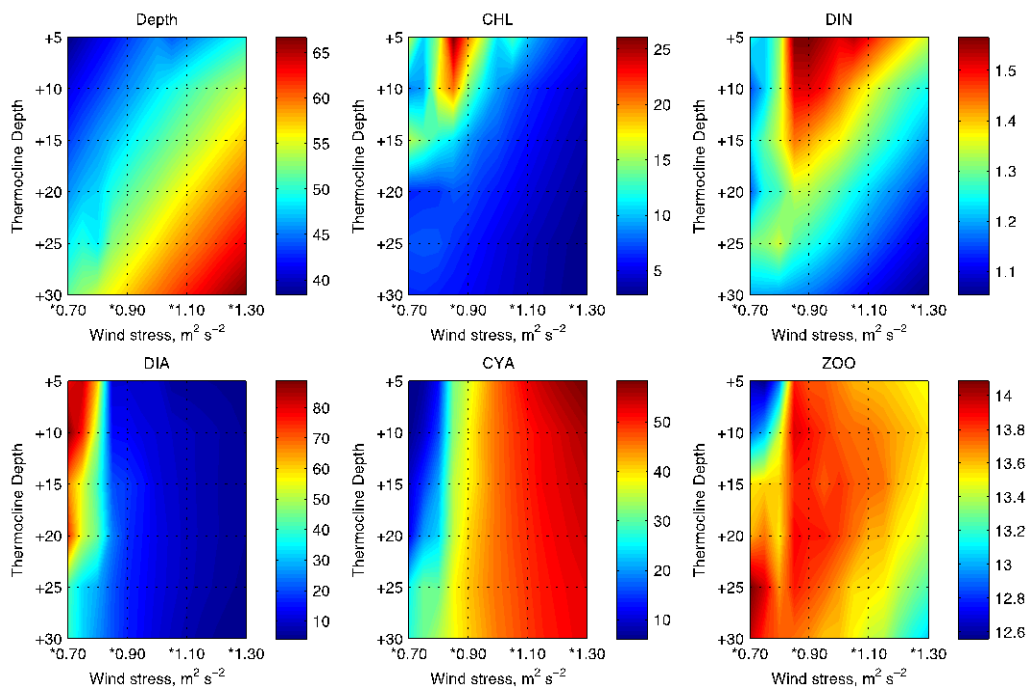

Fig. 40. Lake averaged surface layer depth and plankton biomass for various thermocline depths and wind stress.

Changes in the surface layer depth and its temperature: Figure 41 presents the model simulations for varying thermocline depth and surface layer temperature. Chlorophytes preferred low surface layer and higher temperatures, while dinoflagelates prefer low surface layer and remained high for all temperatures. Diatoms biomass increases with deeper surface layer and increasing temperatures. Cyanobacteria abundance remained similar for increasing surface layer depth and was high during low temperatures periods.

Changes in the surface layer temperature and wind stress: Figure 42 presents the model simulations for varying surface layer temperature and wind stress. The surface layer depth will decrease with increasing temperature. The biomass of chlorophytes and dinoflagelates increases with increasing surface layer temperature and decreased with increased mixing. The biomass of diatoms remained very low, while the biomass of cyanobacteria increased with greater mixing.

Conclusions

Different phytoplankton groups responded differently particularly when two model forcing parameters were changed simultaneously.

-Warm and little windy periods, were indicated by the model as favorable to **chlorophytes** and **dinoflagelates** biomass.

-Increased surface layer depth together with higher temperatures are conditions favorable for the development of **diatoms** biomass.

-Increased winds and lower temperature are favorable for **cyanobacteria** biomass.

The simulated planktonic relative abundance in Lake Tanganyika is the result of the competition between various phytoplankton groups according to their different requirements for light and nutrients. It is the combination of physical, chemical and biological parameters that decide which group will dominate under a particular set of conditions. Those are ultimately largely controlled by climatic conditions.
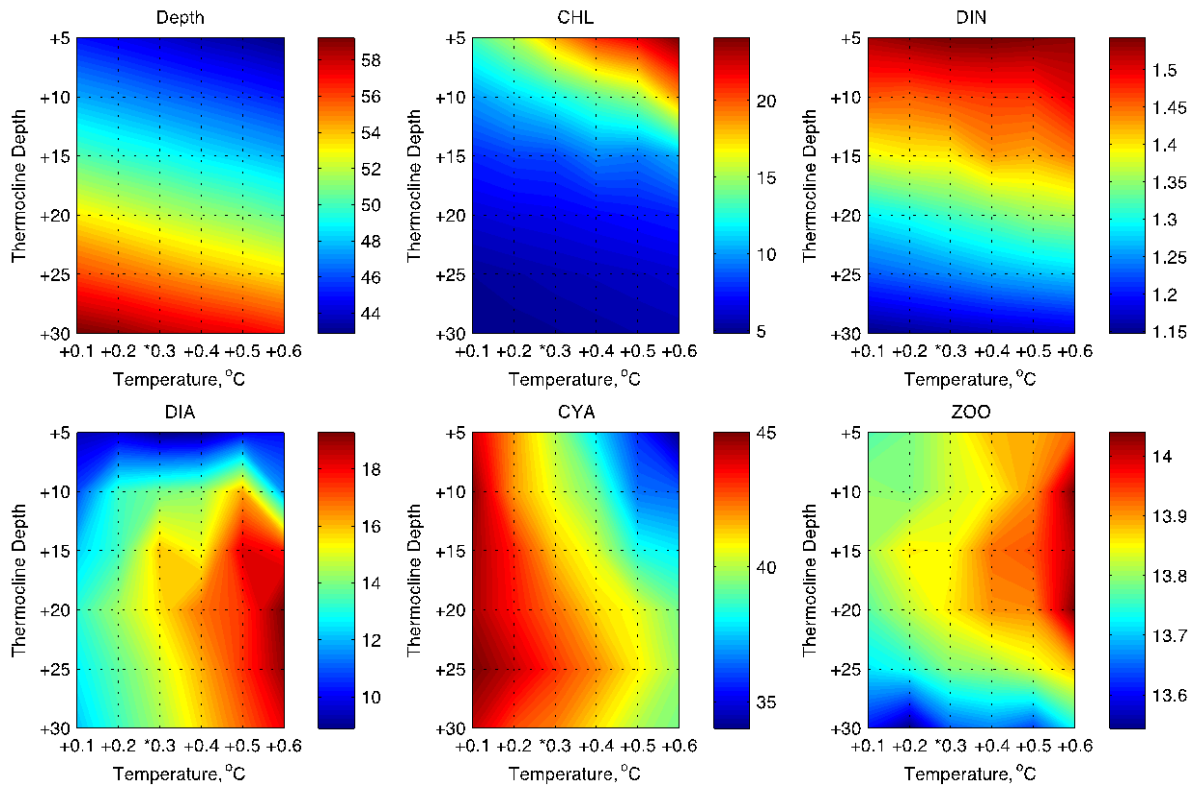
Fig. 41. Lake averaged surface layer depth and plankton biomass for various thermocline depths and surface layer temperature.
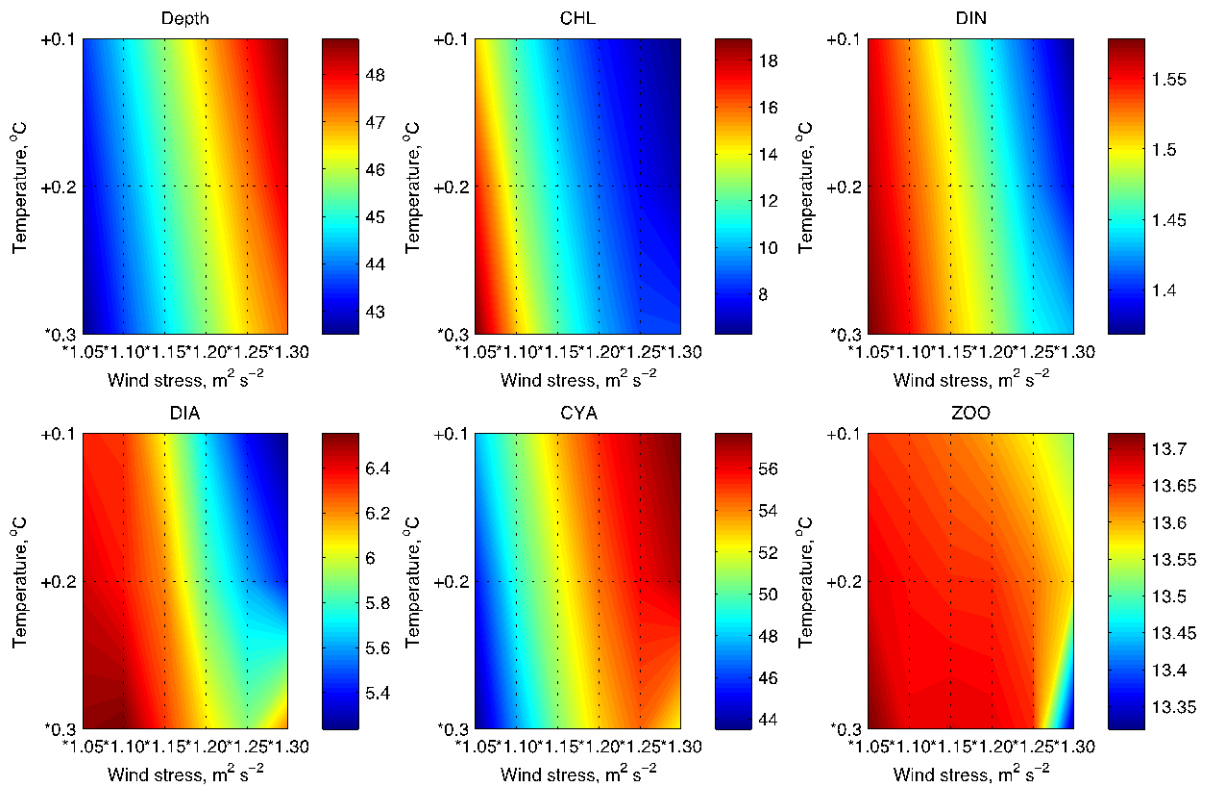

Fig. 42. Lake averaged surface layer depth and plankton biomass for various surface layer temperature and wind stress.

WP 6 Epidemiology

Methodology

Two DRC provinces were chosen for the epidemiologic monitoring and investigations: South-Kivu and Katanga (Figure 43a). The health zones of Uvira and Kalemie were the targets of the project among the 7 Health Zones (HZ) bordering Lake Tanganyika (Figure 43b).

Fig. 43: Areas of epidemiological investigations in DR Congo (DRC).

WP 6.1 Weekly statistics of cholera at three sites

Statistics of cholera cases and fatalities have been compiled using weekly epidemiologic report in two DRC sites: Uvira (2002-2014) and Kalemie (2002-2013). Cholera statistics were available from DRC thanks to the collaboration of the national health system that has been established several years ago.

Definition of suspected case of cholera: any patient, of at least five years in age, presenting serious dehydration or death following an acute watery diarrhea. The case is confirmed when *Vibrio cholerae* O1 or O139 is isolated from the stool sample.

WP 6.2 Spatial and temporal analysis of cholera outbreaks in eastern DRC

Data collected during CHOLTIC has been reinforced by data at five other HZ in eastern DRC. Spatial structure pattern of cholera cases has been achieved from thematic maps representing the yearly attack rates for the period 2011 to 2014 by HZ (for 100000 inhabitants) and those representing the yearly attack rates for the period 2011 to 2014 by Health Area (for 10 000 inhabitants). These maps have been elaborated using QGIS® (2.2.0).

The attack rate is the relationship between the notified cases for one period by the population, the whole multiplied by a factor k.

Time-series and spatial propagation analyses (Bompangue et al., 2009) were applied to the whole data set (2002-2014) (seasonal-trend decomposition based on loess regression (STL) (Cleveland et al. 1990).

Results

During the period from 2011 to 2014, three provinces of the DRC (Katanga, South Kivu and North Kivu) were the most affected by the cholera. Katanga and South Kivu, situated along the Tanganyika Lake, were the most affected by the recurrence of cholera outbreaks. During the period, Katanga reported 31 548 cases (32% of country cases) with 676 deaths, whereas South Kivu reported 24 256 cases (24% of country cases) and 100 deaths.

Throughout the period of the project, the weekly data collection of cholera cases has been continued, not only for both sites targeted by the project CHOLTIC (Uvira and Kalemie), but also for the rest of 515 health zones of 11 provinces of DRC (Figure 44).

During the 2011 to 2014 period, the majority of lacustrine Health Areas (HA) showed the highest attack rates; it was the case of Kilomoni, Kasenga Etat, Mulongwe, Tanganyika, Kimanga, Nyamianda, Kalundu Etat, Kigongo and Makobola. Cholera attack rates are presented for Uvira for the period 2011 to 2014 (Figure 45) and for Kalemie for the period from 2011 to 2013 (Figure 46).

Overall, from 2011 to 2013, the cholera attack rate in the health zone of Kalemie was more important in the health areas that are near Lake Tanganyika, particularly "Hôpital Général de Référence" (HGR), Lutherienne, Clinique, Kataki, Mutakuya, Katibili and Mulembwe. HGR and Lutherienne are two health areas where the attack rate was always more important than others.

Time series of cholera cases in three provinces are presented for the period 2002 to 2014 (Figure 47). The collection of epidemiological information by health areas, allowed to extract weekly sets of cholera cases for Uvira and Kalemie health zones (Figure 48). From these cholera case time series, we extracted three types of HA: the type I HA (endemic profile: Kasenga Etat and Hôpital General), the type II HA (metastable profile: Kala SOS and Kifungo) and the type III HA (sporadic notification of cases). These HZs (Kalemie and Uvira) function under a meta-populationnel rythm (Figure 49).

The statistical analysis of the time series at the health zone scale (Figure 48) and at the health area scale (Figure 50) shows a predominance of the epidemics of cholera during the rainy season.

Fig. 44. Cholera attack rate in DRC and heath zones near Lake Tanganyika, 2011-2014.

Fig. 45. Cholera attack rates in the health areas of Uvira (2011-2014).

Fig. 46. Cholera attack rates in the health areas of Kalemie (2011-2013).

Fig. 47. Times series of cholera cases in eastern provinces of RDC (2002-2014).

Fig. 48. Times series of cholera cases in health zones near Lake Tanganyika (2002-2014).

Fig. 49. Decomposition of the time series of cholera cases at Uvira and Kalemie (2008 – 2014).

Fig. 50. Times series of cholera cases in the health areas of Uvira and Kalemie, 2008-2014 (cholera in red, no cases in green).

Fig. 51. Decomposition of time series of cholera cases at Kasenga Etat (Week 1, 2008 – Week 52, 2014), Uvira and Hopital General (Week 1, 2008- Week 52, 2013) at Kalemie

Discussion

Overall, the CHOLTIC project aimed to understand the factors of persistence and the emergence of cholera around Lake Tanganyika. The analysis of the epidemiologic data monitoring enabled us to arrive at these principal results:

-At the country scale (DRC), an East-West gradient in the distribution of cholera cases, confirming the role that the DRC East provinces play in this disease, thereby corroborating the results of other studies (Piarroux et al, 2009; Bompangue et al, 2009). This space distribution can be explained by several factors: the favorable ecological conditions and high density of the population in the eastern part of the DRC compared to the western part mainly covered by very dense forests limiting migratory flows ((Ministère Congolais de l' Environnement, 2005; Bompangue et al, 2011).

-The health areas bordered by Lake Tanganyika presented high cholera incidences. This observation confirms the results of other studies (Birmingham et al., 1997; Shapiro et al., 1999) and led to a much 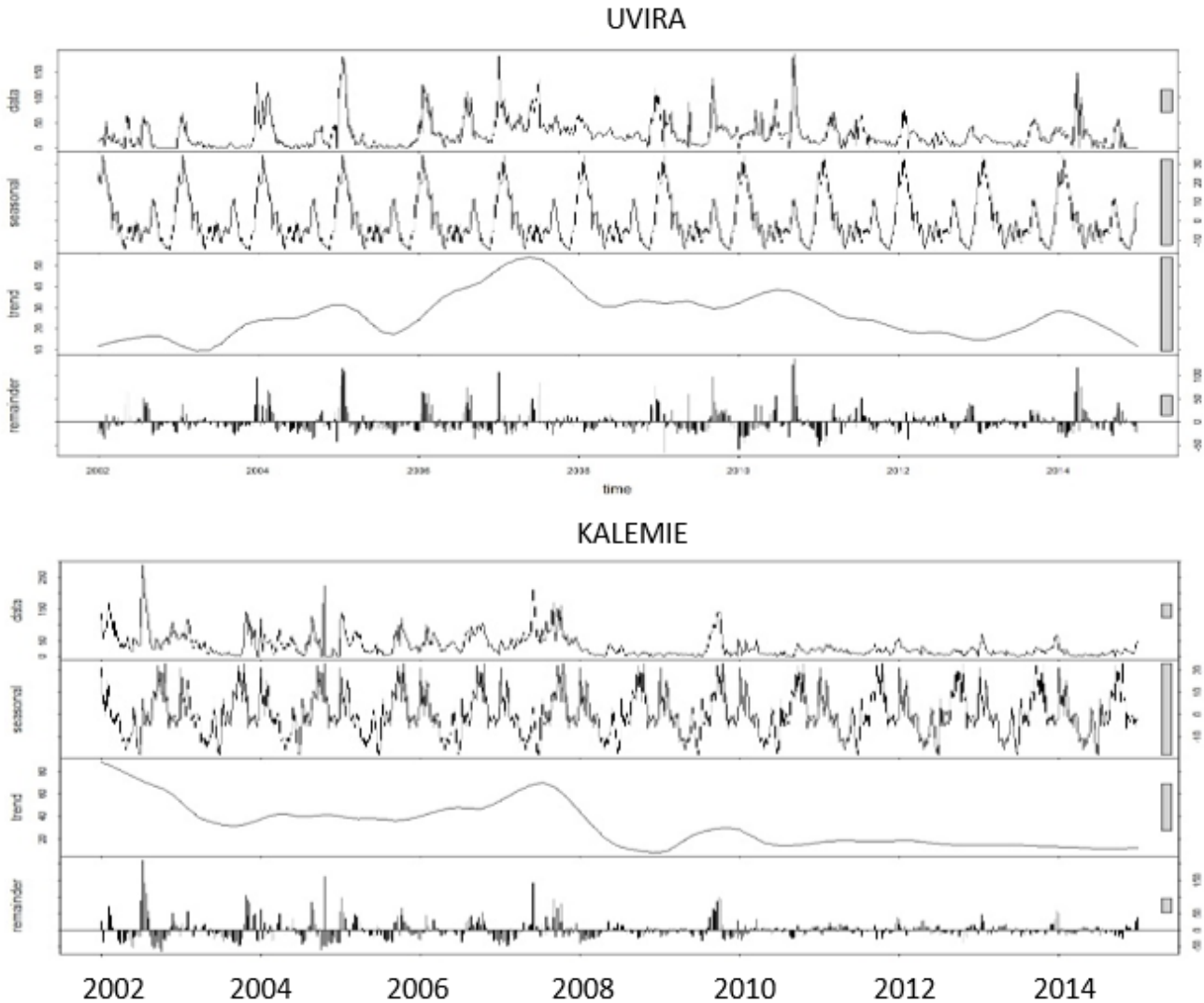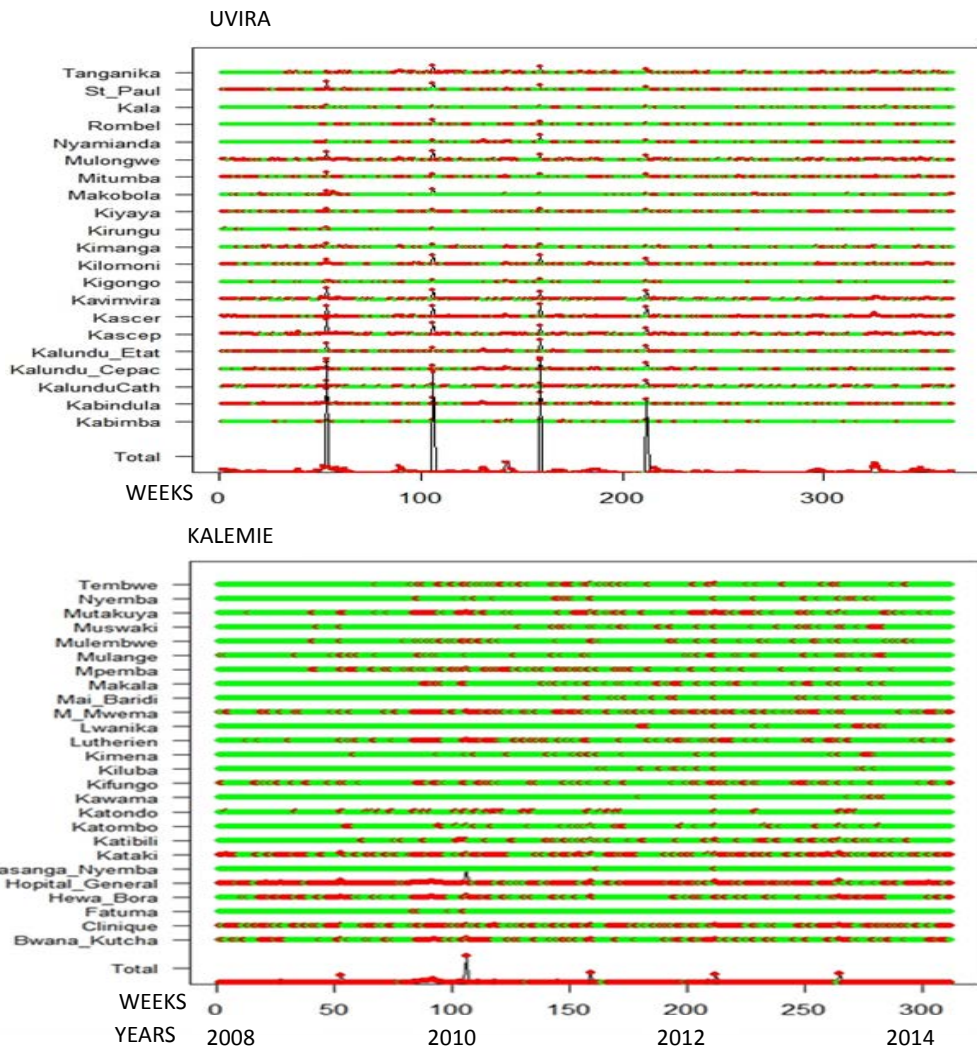larger scale (Health Zones), which showed a correlation between cholera epidemics and location along a lake (Bompangue et al, 2008).

-A seasonality marked by predominance of the epidemics at the rainy season. This seasonality of cholera was identified on two scales (HZ and HA). This seasonality could be explained by the population rhythm of life and to a small extent, by the role of rain (Hashizume et al., 2008; LagneFernendez et al., 2009; Griffith et al., 2006).

-Within two targets HZ, three health areas categories were identified: HA with a continuous reporting of the cholera cases (endemic), HA with discontinuities (metastable) and HA with very little case (bringing back, in very sporadic way, cholera cases).

WP 7 Bacteriological contamination

WP 7.1 Capacities reinforcement of Uvira and Kalemie laboratories

The objectives of this WP were to implement techniques of microbiological water analysis and microbiological analysis of clinical samples in a quality-assured setting at the different field sites in the DRC.

Figure 52 represents a timeline of activities performed to **strengthen the different laboratories**. Detailed reports of all field visits are available. A summary of activities and events influencing the activities and outcome of this WP is described below.

Fig. 52. Timeline with overview of capacity strenghtening activities.
[*]mission performed by INRB (DRC) collaborator

2011
Uvira - DRC
To strenghten capacities at Lake Tanganyika (WP 7.1), contacts were made with Centre de Recherches Hydrobiologiques (CRH-Uvira), the Ministry of Health (MOH) and the reference hospital in Uvira (Hôpital Général de Référence; HGR-Uvira). Both the laboratory of CRH and HGR-Uvira were assessed on site for infrastructure and equipment as well as availability of staff. The laboratory of the HGR-Uvira turned out to

be most appropriate for our activities. A partnership was launched and formalized by the signatures of memoranda of understanding (MOU) September 2011.

During this exploratory mission, a pilot sampling and processing of lake surface water was performed and procedures for the screening of surface water were updated and adapted to local conditions. Supplementary materials for sampling and work-up were listed in collaboration with the local partner and procured and shipped by ITM. A template for data collection was designed and adapted in terms of ease of use and feasibility in consultation with local staff.
Plans were made for a training and implementation mission early 2012.

Kalemie - DRC
The reinforcement of capacities in Kalemie, could not take place in 2011 due to security concerns in the region.

2012
Insecurity and instability in eastern Congo precluded missions for ITM members to the region. This hampered the planned training mission and implementation of the WP7 in Uvira and Kalemie laboratories, alternatively bacteriological monitoring was developed in Mpulungu, Zambia.

At the same time, together with the Congolese staff of INRB (Kinshasa, DRC), plans were made for bacteriological training for the Uvira staff.

Mpulungu, Zambia
As ITM has no institutional collaborations in Zambia, an **exploratory mission** was organized to meet local partners and present project objectives. Various stations in the Mpulungu area were visited: Department of Fisheries (DOF-Mpulungu), General Hospital (MOH) in Mbala, the provincial general hospital in Kassama as well as the Health Centre (HC) and District Medical Office (DMO) in Mpulungu. After meetings with district and provincial officers, a partnership was launched and formalized by the MOU with Mpulungu HC in September 2012. The provincial office in Kassama fully supported the collaboration with their district facility in view of the capacity reinforcement of their laboratory and proximity to clinical cases and the lakeside.

A three-week training and implementation mission in September 2012 enabled starting-up of bacteriology processing of patients samples, lake water, phyto- and zooplankton.

The capacity reinforcement of the Mpulungu HC laboratory included:
    -Reorganization of the laboratory with installation of necessary equipment

-Training on all laboratory procedures including work-up of samples, biosafety, equipment management, quality control procedures, data collection, and shipment of samples
-Development and implementation of a field manual

A protocol of the study was submitted to the University Of Zambia, Biomedical Research Ethics Committee (UNZAREC) for ethical clearance.

A follow-up mission with refresher training was performed end 2012. During this mission, a full refresher training was provided, as the Mpulungu HC suffered from high staff turn-over and key-persons initially trained and involved in the laboratory work and data collection had been replaced.

Additional procedures were developed and training was given both to the Department of Fisheries and laboratory staff for the sampling and screening of three fish species as suggested by the follow-up committee fall 2012.

Study outlines for a cholera RDT (Crystal Vc, SPAN Ltd, India) study were discussed with the district health officer and health staff involved and a planning was made to implement the study spring 2013. A protocol was designed and submitted for ethical clearance at the institutional review board of ITM and the ethics committee of Antwerp University, approvals were granted.

Kinshasa, DRC
During week 1, exchanges were made with INRB members to plan a field training in Uvira and Kalemie to implement the environmental sampling and collection of clinical samples. All preparations were made: training program, fine-tuning of procedures and protocols, development of data collection tools, etc.
Tentative dates were as follows: Uvira August 2012, Kalemie September 2012.

2013
Continuing insecurity and instability in the Uvira region in DRC also precluded missions for our local Congolese INRB partners. As a result, we decided to end our collaboration with the HGR-Uvira. The contract stopped in the end of 2013, all equipment and materials were donated to the hospital.

A follow-up of the activities in Mpulungu was done during a field visit May 2013. During this time, the field study to evaluate Cholera rapid diagnostic tests (RDT -Crystal Vc) was implemented. As we noticed an absence of cholera cases in Mpulungu, Zambia, later 2013, ITM decided to write out a similar additional study evaluating another RDT

product (SD-Bioline, Standard Diagnostics, Republic of South Korea) for the field site in Kalemie.

The training and implementation of the study was subcontracted to the INRB in Kinshasa, as ITM collaborators were not authorized (by ITM security policy) to travel this region. Fall 2013, a field visit in Kalemie was performed by a local INRB collaborator to provide training and capacity reinforcement to a member (laboratory technician) of the local "cellule cholera" in Kalemie.

WP 7.2 Microbiological screening of patients and surface water

Initial objectives of this work package:
    (1) To sample and to analyze lake surface water for microbiological contamination
    (2) To sample and analyze stool samples from patients suspected of cholera.
    (3) To evaluate the rapid diagnostic test for detection of *V. cholerae* in stool samples from patients suspected of cholera.

1. Environmental sampling

For reasons mentioned above, the environmental part of WP7.2 was only performed in Mpulungu, Zambia. Biweekly sampling of lake water, phyto-and zooplankton samples was started in August 2012. In total, 47 samplings were performed, other weeks were missed for various logistical reasons (*e.g.* no electricity, competing field missions of local staff, sick-leave of local staff). The last two months of 2014, no further sampling on the lake was done, as the Department of fisheries (DOF) responsible for sampling was committed to other field operations.

Fecal contamination rate of lake water was assessed and is presented in the figure below as colony forming units of indicator organisms (fecal coliforms) per 100 mL of lake water for the period August 2012 to November 2014 (Figure 53).

In addition to fecal contamination, water, phyto- and zooplankton samples as well as the three main fish species in Lake Tanganyika were assessed for presence of *V. cholerae.* A very limited number of *V. cholerae* were isolated from the environment (2 isolates from lake water, 1 isolate from *Stolothrissa tanganicae*).

Fig. 53. Fecal contamination of coastal and pelagic water, expressed as number of colony forming units (CFU) of fecal coliforms per 100 mL of lake water.

2. Stool samples from patient with suspected cholera

As implementation of the activities in DRC was hampered, we strengthened referral of samples from suspected cases to the national reference laboratory (INRB, Kinshasa, DRC) were they were further analyzed and confirmed.

In Zambia, starting August 2012, all cholera suspected patients visiting the Mpulungu Health Centre were sampled and screened. A small outbreak occurred in the second half of 2012. Figure 54 represents the number of cases per epidemiological week. After this period, only few sporadic cases presented at the Health Centre.

Fig. 54. Number of cholera case (Y axis) per week (X axis) in 2012, Mpulungu, Zambia

Evaluation of a rapid diagnostic test (RDT) – Crystal VC® field site Zambia.

-Full study protocol was written out and approved by the University of Zambia Biomedical Research Ethics Committee and by the Institute of Tropical Medicine institutional review board and the ethics committee of the University of Antwerp.

-Training and implementation of the study protocol was performed at site, but since May 2013 there were insufficient cholera cases to complete the study.

Evaluation of a rapid diagnostic test (RDT) – SD-Bioline in Kalemie, DRC

-Due to limited cholera cases in Zambia, an additional study protocol was designed by ITM. Standard Diagnostic, Inc. manufacturer of the SD- Cholera AgO1/O139 RDT was contacted and provided necessary RDTs.

-The study protocol was approved by the Ethics committee of the University of Kinshasa and by the Institute of Tropical Medicine institutional review board and the ethical committee of the University of Antwerp.

-Training and implementation was sub-contracted to the INRB, Kinshasa, as ITM staff was not allowed to travel in eastern DRC (including Kalemie).

-The study started in October 2013, after a field visit of an INRB collaborator.

-Between October 2013 and December 2013, 76 samples (of 440 required sample size) were included. As quality of samples and data could not be assured (problems of delay in transport, inadequate storage of samples, absence of critical data, delayed reporting), ITM decided to terminate the contract ahead of time, after the first phase.

WP 8 Microbiological confirmation

WP 8.1 Species identification and serogroups determination

As activities partially shifted to Zambia, these isolates were collected and shipped directly to ITM during the field visit in May 2013. At ITM they were confirmed by serology, stored at -70°C and further shipped to AP/HM (Marseille, France).

After this initial shipment, local staff were not able to organize any further shipment. Transport companies (FedEx and DHL) in Zambia (as in other African countries) no longer wanted to ship biological substances (as we experienced from other project, this concided with the onset of the Ebola outbreak). An overview of samples collected and transferred for further molecular analysis at AP/HM is presented in Table IV.

Table IV: Samples collected and transferred for further molecular analysis.

Year	Country	Site	Type	N samples	Identification
2011	DRC	Uvira	Environmental	2	*V. cholerae* non-O1
2012	Zambia	Mpulungu	Environmental	2	*V. cholerae* non-O1
	Zambia	Mpulungu	Clinical	27	*V. cholerae* O1, Inaba
2013	DRC	Katanga (Kalemie + Nyemba)	Clinical	39	*V. cholerae* O1, Inaba
2014	DRC	Katanga (Kalemie)	Clinical	34	*V. cholerae* O1, Inaba
	DRC	Sud Kivu (Uvira + Fizi)	Clinical	28	*V. cholerae* O1 (Inaba(3) ; Ogawa(25))

Concerning the DRC, samples were collected as part of the national surveillance program. They were sent to INRB for confirmation by serology and shipped from there to APH-Marseille.

WP 8.2 Isolation of strains and complementary analysis

A total of 531 *V. cholerae* isolates from recent cholera epidemics in the DRC, Zambia, Guinea, Ghana and Togo were subjected to MLVA. Analysis of the six VNTRs yielded 118 MLVA types. A MST was constructed using the combined MLVA data to assess the relationships between the isolates and the epidemic populations. On a continental scale, the MST revealed strong geographical clustering with isolates from the African

Great Lakes Region, including the DRC and Zambia, clustering together. Furthermore, the isolates collected in the West African countries of Guinea, Ghana and Togo formed a separate group (Figure 55).

Legend:
- DRC 2008
- DRC 2009A
- DRC 2009B
- DRC 2011
- DRC 2012
- DRC 2013
- DRC 2014
- Zambia 2012
- Togo 2010
- Togo 2011
- Togo 2012
- Ghana 2010
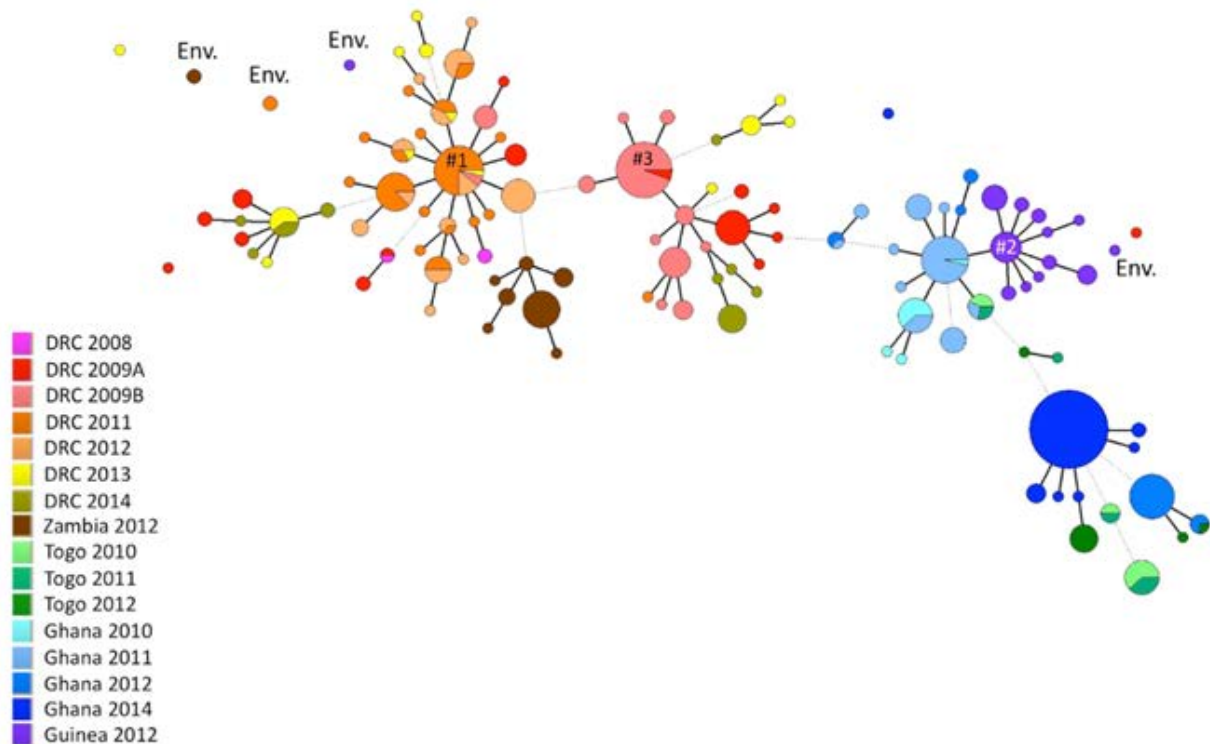- Ghana 2011
- Ghana 2012
- Ghana 2014
- Guinea 2012

Fig. 55. Minimum Spanning Tree based on the MLVA types of 531 *V. cholerae* isolates derived from several recent African epidemics. Each MLVA type is represented by a node, and the size of the nodes reflects the number of isolates with the indicated MLVA type. The relationships between isolate MLVA types are indicated by the type of connecting segments and the length of the segment between nodes. The solid lines indicate the most likely Single Locus Variant, and the dashed lines represent Double Locus Variants. The distance between the nodes represents the number of varying VNTRs. The colors reflect the distinct country and period of isolate origin (grouped by epidemic populations). Pie charts were used to indicate the distribution of strains isolated from different time periods or countries displaying an identical MLVA type. The predicted founder MLVA type (based on goeBURST analysis) of the DRC 2011/2012 (#1), Guinea 2012 (#2) and DRC late-2009 (#3) epidemic complexes are indicated. The isolates collected from environmental samples are labeled "Env.".

At a country-level scale our analyses revealed several distinct clonal groups, most notably (1) DRC 2011/2012, (2) DRC 2009, (3) Zambia 2012, (4) Ghana 2010/2011 and Guinea 2012, (5) Togo/Ghana 2012 and (6) Ghana 2014. In contrast, all six of the typed environmental isolates were found to be singletons, unrelated to the main clinical epidemic isolate clusters (Figure 55).

Our analysis showed that the DRC 2011 and DRC 2012 isolates grouped together as one discrete complex (Figure 55). During this two-year period in the DRC, cholera was caused by the extensive expansion and diversification from a single MLVA haplotype. The isolate found at the beginning of the 2011/2012 epidemic in Kisangani, Orientale

Province in March 2011 was haplotype #1, which was designated the founder of the DRC 2011/2012 complex. MLVA type #1 and a Single Locus Variant of this haplotype were the only types identified during the first week of the outbreak. This MLVA cluster then diversified in parallel with the spatiotemporal spread of the epidemic (Bompangue et al. 2012), as the most distant MLVA haplotypes within this cluster were identified in distant provinces in 2012. These findings correlate with an epidemiological report of the cholera epidemic that struck the DRC in 2011. This epidemic aggressively diffused from the onset point in Kisangani, Orientale Province across the country in less than 130 days. Strikingly, outbreaks followed the Congo River and quickly reached non-endemic zones in the West that had not experienced an epidemic for approximately 10 years (Bompangue et al. 2012). Kisangani and the Congo River are localized on the detailed map of the DRC (Fig 56, right).

Fig. 56. Detailed maps of Guinea and the DRC. Left: Guinea; the capital, Conakry, and Kaback Island are indicated along the Atlantic coast. Right: DRC; the cities (red squares) and provinces mentioned are indicated on the map. The Congo River and Lake Tanganyika are also shown (in blue).

Interestingly, the predicted founder of the 2011/2012 DRC epidemic, persisted in the country over the course of several years, as haplotype #1 was represented in isolates collected in the DRC during the 2009, 2011, 2012 and 2013 epidemics. Only one DRC isolate collected in 2011 did not belong to the major DRC 2011/2012 MLVA cluster. Instead, this haplotype was a Single Locus Variant of the major DRC 2009 cluster. In stark contrast, the *V. cholerae* non-O1 strain isolated at the same period from a water sample in Uvira, South Kivu was a genetically unrelated singleton (Figure 55, orange node labeled "Env."). Uvira is located on the northern shores of Lake Tanganyika as indicated in Figure 56 (right panel).

Overall, the panel of DRC 2009 isolates displayed a high level of genetic diversity. In fact, four DRC 2009 clinical isolates collected in February and March 2009 were

designated distantly related singletons (Figure 55). The MST was then analyzed in further detail considering the date of sample isolation. The isolates collected during the first half of the year were highly diverse (indicated as "DRC 2009 A" in light pink; Figure 55). In contrast, 63 of 66 isolates (95.5%) collected from July to November of 2009 in Katanga Province formed a tight clonal complex (indicated as "DRC 2009 B" in red; Figure 55). This bottleneck effect was concomitant with an epidemic rebound in Katanga Province after a complete lull in cholera transmission in May and June 2009. In July 2009, cholera first appeared in Kalemie, a city located on the shore of Lake Tanganyika, and then spread throughout the rest of the province (based on field investigations in the DRC; Renaud Piarroux). MLVA type #3 was designated the potential founder of this DRC 2009 B MLVA complex. Isolates of this haplotype were the first collected on February 5, 2009, and isolates harboring this haplotype were collected up to November 20, 2009. Notably, MLVA type #3 was also found in August 2009 in Uvira, a city located approximately 360 km north of Kalemie on the shore of Lake Tanganyika (all sites are labeled in the right panel of Figure 56). Therefore, we hypothesize that this strain likely persisted in the region following the early-2009 outbreaks and a subsequently gave rise to the late-2009 DRC epidemic.

Initial analysis of the DRC isolates collected in 2013 and 2014 appeared to be descendants of the strains already circulating in the country during the 2009 and 2011/2012 epidemics.

In Zambia, all clinical isolates from the 2012 epidemic formed a restricted clonal complex. Once more, the two non-O1 environmental isolates collected from the shores of Lake Tanganyika in Mpulungu, Northern Province were singletons (brown node labeled "Env.") unrelated to the clonal complex (Figure 55).

Likewise, the Guinea 2012 clinical isolates formed a solid clonal complex, with an MST of closely related derivative isolates that stemmed from the founder haplotype (MLVA type #2) (Figure 55). In a previous study, our group has shown that the Guinea 2012 epidemic appears to be due to the importation of a toxigenic clone from Sierra Leone. Using MLVA typing, we have demonstrated progressive genetic diversification of the strains from the founder type correlated with spatiotemporal epidemic spread (Rebaudet et al. 2014). The founding MLVA type was also the first and only MLVA type indentified during the onset of the epidemic, on Kaback Island (Figure 56, left), Guinea in February 2012 (Rebaudet et al. 2014). In contrast, the two Guinean environmental strains (dark violet nodes labeled "Env.") isolated from water samples at the site of the initial outbreak (Kaback) were unrelated singletons (Figure 55).

As *V. cholerae* is autochthonous in the coastal aquatic ecosystem, it has been widely presumed that cholera epidemics are triggered by environmental factors that promote the growth of local bacterial reservoirs (Colwell, 1996). However, our results support hypothesis H1, which supposes the absence of a perennial reservoir of *V. cholerae* in water bodies during inter-epidemics periods. All six of the environmental isolates collected from a range of countries were genetically unrelated singletons, of which only two environmental isolates were collected in the context of the CHOLTIC project. We acknowledge that additional environmental isolates of *V. cholerae* should be included in the panel to affirm the relationship (or lack of) between clinical outbreak strains and those found in water bodies. Indeed, to examine the diversity of environmental strains, efforts should be made to collect further samples. Nevertheless, our preliminary analysis of the environmental samples correlates with two recent reviews elucidating the environmental determinants of cholera outbreaks in Africa. These reviews by Rebaudet et al. (2013a,b) found that at least 76% of cholera cases in Sub-Saharan Africa occurred in non-coastal regions located over 100 km from the coast. From 2009 to 2011, annual incidence rates of cholera were three times higher in inland Africa compared with the coastal region. In fact, toxigenic *V. cholerae* isolates have only been recovered from the environment during an outbreak, when patient-derived contamination of water sources is expected (Rebaudet et al. 2013a,b).

Our findings are also consistent with the phylogenic assessment of an extensive panel of seventh pandemic *V. cholerae* isolates. Mutreja et al. (Mutreja et al., 2011) have revealed that a specific *V. cholerae* El Tor clonal lineage appears to be responsible for the current pandemic. Their study demonstrated that the seventh pandemic is monophyletic and originated from a single ancestral clone that has radiated globally in distinct waves (Mutreja et al., 2011). Lineages of the current pandemic appear to emerge, diversify and eventually become extinct (Mutreja et al., 2011). Notably, we observed a similar phenomenon at a smaller scale with the DRC 2009, DRC 2011/2012, Guinea 2012, Zambia 2012 and Ghana epidemics. As isolates from the African Great Lakes Region were not included in the seventh pandemic phylogeny, whole-genome sequencing and phylogenic analysis of this panel of African strains would provide even further insight into the mechanisms of cholera epidemics in the region.

Overall, this study shows that cholera is indeed a regional public health dilemma in Africa. With varying dynamics in each country, certain strains are able to persist in a given region over a period of several years and occasionally spread into non-endemic areas or neighboring countries. Indeed, several elements play a role in cholera epidemics including climate, geography, economy, hygiene, sanitation, access to potable water and population movement (Bompangue et al. 2012,

Rebaudet et al. 2012). Therefore, public health strategies should be optimized according to the dynamics and scale of cholera epidemics in each region. These findings also demonstrate the importance of monitoring the circulation of the bacterium among human populations, which appear to serve as the principal reservoir of toxigenic *V. cholerae*.

The majority of isolates assessed in this study were included in a publication in PLOS Neglected Tropical Disease (Moore et al. 2015, accepted).

WP 9 Interdisciplinary data integration

The objective of this WP is the integration of the Time Series (TS) of data collected or aggregated at a weekly time step in a synthetic analysis of their spatio-temporal interactions. Special attention is paid to the epidemiological TS as dependent variable with respect to the major environmental variables that have been identified a priori as having potential statistical association with cholera outbreaks among the meteorological, limnological, phyto-zoo-plankton and fisheries sets of variables.

Methods

Periods of analysis

Three types of datasets were used for the specific analyses:

(1) Cholera/RS Chl-a - 2002-2014. The longest TS (2002-2014) allow the analysis of the association between cholera cases aggregated (sum) per week and Health Zone (HZ) and the dynamic Chl-a concentration in the lake aggregated (p75) per week for the entire lake or per sub-regions of the lake (N=12) named "Eco-Regions" (ER) as determined previously (Bergamino et al., 2010) (Figure 57). Cholera from six (among seven) HZ located along the Congolese side of the lake were available for this longest period.

(2) Cholera / satellite Chl-a - 2008-2014**.** During the period 2008-2014, the same type of analysis was performed using finer spatial level because epidemiological data were available for Health Area (HA) of Kalemie and Uvira HZ. As the cholera cases can be assimilated to rare events at this spatial scale, we performed the analysis on aggregated statistics (sum) in two groups of HA: HA along the lake shoreline and inland HA. An analysis that takes in consideration the position of the cholera cases with respect to the lake ER is justified to demonstrate the existence of spatio-temporal interactions between epidemiology and limnology which is an essential argument supporting hypothesis H0.

(3) Data collected during the CHOLTIC operational period (2011-2014): Cholera, RS-Chl-a, meteorology, limnology, phytoplankton and fisheries. During this monitoring period (2011-2014) homogeneous data sampling was organized with collaboration of our local partners. The main time step is the week but the spatial representativeness is variable depending on the dataset. Epidemiological and satellite chl-a TS are described above. Other processed TS are the environmental and fisheries (chap 2, WP1) data from Uvira and Kalemie stations. The most significant local environmental and fisheries variables were selected for the 2011-2014 period (Table V). This choice

is justified by pair scatter-plots, expert knowledge and data-processing convenience. Nevertheless, the exhaustive dataset allows future analyses to be developed, especially regarding ecological modeling (WP10) and phytoplankton composition not significant in this preliminary approach.

All statistical analysis applied in this WP aims to take into account delayed responses, autocorrelation inducing a strong dependence between observations, non-stationarity and the rare event nature of the epidemiological-dependent variable.
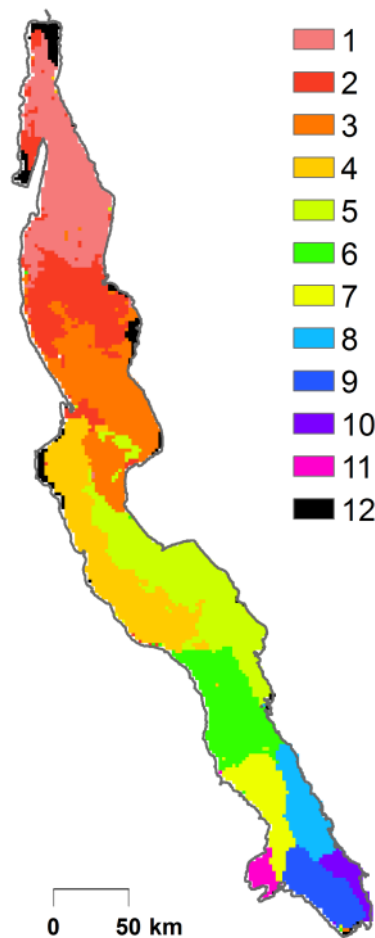

Fig. 57: Map of Eco-Regions (ER)

Table V. List of selected variables in the multivariate analysis of the 2011-2014 period.

Code	Variable	Unit	Variable - details
WN	Week number		
AirTemp.degC	Air T°	°C	Air T° measured by an automatic weather station every 15 minutes (average during the week)
RelHum.pc	Relative humidity	%	Rel. humidity measured every 15 minutes (average during the week)
Precip.mm	Rainfall	mm	Total rainfall during the week
SolarRad.W.m-2	Solar radiation	watt/ m²	Solar radiation measured every 15 minutes (average during the week)
Temp.surf.coast.deg C	Water T°	°C	Water T° at the coast at 0 m
Temp.surf.pel.degC	Water T°	°C	Water T° in pelagic water at 0 m
Temp.pel.80m.degC	Water T°	°C	Water T° in pelagic water at 80 m depth
–Chla.coast.mg.m-3	Chlorophyll a	µg/l	Chlorophyll a at the coast at 0 m
Chla.pel.mg.m-3	Chlorophyll a	µg/l	Chlorophyll a in pelagic water at 0 m
ChlaZmix-P-0	Chlorophyll a	µg/l	Chlorophyll a average in the mixolimnion zone
Turb -C	Turbidity	NTU	Turbidity at the coast at 0 m
Turb -P-0	Turbidity	NTU	Turbidity in pelagic water at 0 m
Catchperlight.Lates	Catch of *L.stappersii*	kg	Catch per standard unit of fishing boat (10 lights)-night
Catchperlight.Stolo	Catch of *Stolothrissa*	kg	Catch per standard unit of fishing boat (10 lights)-night
Catchperlight.Limno	Catch of *Limnothrissa*	kg	Catch per standard unit of fishing boat (10 lights)-night
RS-Chla ER 1 to 12	Chlorophyll a mean	µg/l	Data from remote sensing analysis in ER 1 to 12 (3d quartile during week)
RS-dayLSWT ER 1 to 12	Day Lake Surface Water Temperature	°C	Data from remote sensing analysis in ER 1 to 12 (median during week)
Cholera cases HZ	N cholera cases in Uvira HZ	N	Aggregated number of cholera cases for all health area (HA) of Uvira Health Zone (HZ)
Cholera cases – coastal HA's	N cholera cases in coastal Uvira HZ	N	Aggregated number of cholera cases for health area (HA) of Uvira Health Zone (HZ) along the lake
Cholera cases – inland HA's	N inland cholera cases Uvira HZ	N	Aggregated number of cholera cases for inland health area (HA) of Uvira Health Zone (HZ)

Pre-processing, Missing Data Filling

For the three periods, the Missing Data Filling (MDF) is the first processing step. Regarding RS-Chl-a LSWT, spatio-temporally (lake, ER and week) aggregated TS are nearly complete but noisy due to low cloudmasking constraints and spatio-temporal aggregation. Time-gaps are short and exceptional for the same reason. They were filled using linear interpolation. The few missing data in the selected environmental variable measured locally at permanent stations were also filled using linear interpolation. This method was chosen based on several tests consisting in the comparison of the original TS with the reconstructed TS using several filling algorithms (mean between k^{th} order neighbors and linear interpolation).

Binomial analysis

This analysis was applied both to the 2002-2014 and 2008-2013 TS. It consists of thresholding each variable (RS-Chl-a and cholera cases) to turn it into a binomial variable followed by a Chi-Square test of Independence based on bivariate contingency tables. In the present report, we will call:

'HH' the co-occurrence of high values of both Chl-a concentrations and cholera cases

'LL' the co-occurrence of low values of both Chl-a concentrations and cholera cases

'LH' the occurrence of low values of cholera cases with high values of Chl-a concentrations

'HL' the occurrence of high values of cholera cases with low values of Chl-a concentrations

This method takes into account delayed association in time (lag) and allows the processing of rare events and non-stationary variables with temporal autocorrelation. The sensitivity to the threshold values is also taken in consideration.

For each ER, seven tables were generated. These tables (example: Figure 58) contain as many rows as the number of threshold levels for epidemiology and as many columns as the number of threshold levels for Chl-a concentrations. The list of thresholds levels corresponds to the following percentiles: 20, 30, 40, 50, 60, 70, 75, 80, 85, 90, 95 and 97.5. For each threshold combination, the contingency table was built and the following statistics are reported: Chi-square statistic, Chi-square p-value, ratio LL/HH, ratio ([HH+LL]/[LH+HL]), Kappa Index of Agreement (KIA) and the lower and the higher boundary of the confidence interval (Δ= 5 %) of KIA. "NA" is reported in the tables if Chi-square test fails the minimum number of observations in the 4 cells of the corresponding contingency table is lower than 5.

These tables were then used to identify the thresholds combination producing the most significant positive association between cholera and Chl-a. For these significant associations, we created a time plot (see example in Figure 60) to identify which

weeks of the year were concerned by the association between cholera and Chl-a concentration (HH and LL co-occurrences).

The same analysis has been performed on negatively lagged TS of Chl-a to analyze the variation of the association considering a delayed response of cholera with regard to Chl-a. Moreover, this delayed association has also been analyzed using Kendall-TAU statistics.

As the epidemiological TS of Kalemie shows a rupture producing a mean and a variance reduction that has been confirmed using the At-Most-One-Change (AMOC) point method, this longest TS of Kalemie has been split in two periods [2002-2007] and [2008-2014] and the former analysis was thus separately applied to each sub-period.

ARIMA - remainders analysis

ARIMA (Auto-Regressive integrated Moving Average) model has been parametrized to address non-stationarity (due to seasonality) and autocorrelation. This method produces modeled TS in function of time and remainders that were confirmed as white noise using the Ljung–Box version of the Portmanteau goodness-of-fit test. The co-variation between those remainders taking into account temporal lag was then investigated to reveal possible interactions between remainders of independent variables (environment and fisheries). The remainders analysis is of great significance because their magnitude is generally very large with respect to the variance of the modeled TS.

ARIMA - modeled TS analysis

Modeled cholera and environmental TS produced by the ARIMA processing were visually analyzed and compared. This comparison focuses on the amplitude and phase of the main cyclic components and their inter-annual variations. A comparison of temporal patterns is done with known ecological changes at Lake Tanganyika or with cholera outbreaks in other regions.

RESULTS

The results of the various statistical analyses to identify possible significant links between environmental parameters and cholera epidemics along Lake Tanganyika are presented below.

Binomial analysis: contingency table crossing binomial variables

We firstly investigated Uvira 2002-2014 TS, Kalemie 2002-2007 TS and Kalemie 2008-2014 TS by crossing cholera cases in HZ and RS-Chl-a in each ER. Figure 58 gives an example of this analysis for Uvira HZ compared to RS-Chl-a concentrations in two ERs of the lake. On the left, ER1 near Uvira shows some significant p-value corresponding to positive association (positive Kappa). This is a typical result for an ER near Uvira (ER1, ER2 and ER3). On the right, ER7 far from Uvira shows a minimum p-value corresponding to negative association (negative Kappa). The more frequent small p-values associated to positive KIA in tables corresponding to near ER attest for a lower sensitivity to the threshold level. This also supports the hypothesis of a quite strong spatio-temporal interaction between temporal behavior of RS-Chl-a concentrations and epidemics.

Figure 59 shows the contingency tables (using the thresholds identified in the previous step at Figure 58) for two statistically significant examples regarding Uvira among the large number of contingency analysis performed.

The HH and LL co-occurrences time-plots corresponding to Uvira and ER1 are reported in Figure 60. This example is very representative of the majority of the temporal signals we obtained. HH co-occurrences generally happen during two main periods: (1) during the wet season between early December (week 48) to end of March (week 12) and (2) during the dry season from early June (week 24) to mid-October (week 42). In between a period with predominating "latent" (< P75) cholera corresponding to low Chl-a concentrations (LL co-occurrences) is observed, which could be interpreted as contamination possibly linked with rains in this period but little with phyto-planktonic blooms.

Interestingly, both periods of HH co-occurrence seem to be related to periods of seasonal changes:

- In May/June, northeast winds shift to southeast winds corresponding to the end of wet season and beginning of the dry season
- In November/December, a few weeks after seasonal changes when SE winds shift to NE corresponding to the beginning of the wet season.

In the south, a phytoplanktonic bloom was observed each year (Figure 23). This apparently induced an increase catch of *Limnothrissa* in May/June 2012 and *Stolothrissa* migration in May to September 2013 (Figure 60). This migration starting in May and related to the planktonic bloom correspond to previous observations in the south during the CLIMFISH project (Plisnier et al 2009). During a CHOLTIC mission, Mutondwe Island villagers indicated that those moments often corresponded to gathering of fishermen, such as at Mutondwe Island, where cholera outbreaks often start. After September, similar events seemed to take place, although more historical environmental and cholera data are required to verify this observation.

The weekly inter-annual variability is influenced by the constraint on high KIA value used (linked to HH+LL frequency and HH/LL ratio constraints) that controls the selected thresholds. The maximum KIA value is nevertheless quite low in the majority of the cases.

The same approach has been applied using negatively lagged RS-Chl-a. Figure 61 illustrates the variation of KIA statistics from lag 0 to 20 weeks. Regarding the thresholds combination that minimizes Chi-square p-value and maximizes KIA it remains very stable until a lag of 10 weeks. Even if the values are generally significant, both statistics remain very low and their variations are smaller than the confidence intervals. From the first methodological approach we cannot thus demonstrate any statistically significant delayed response of cholera cases with regard to the RS-Chl-a concentration. Kendall-TAU analysis in function of the lag results in the same conclusion.
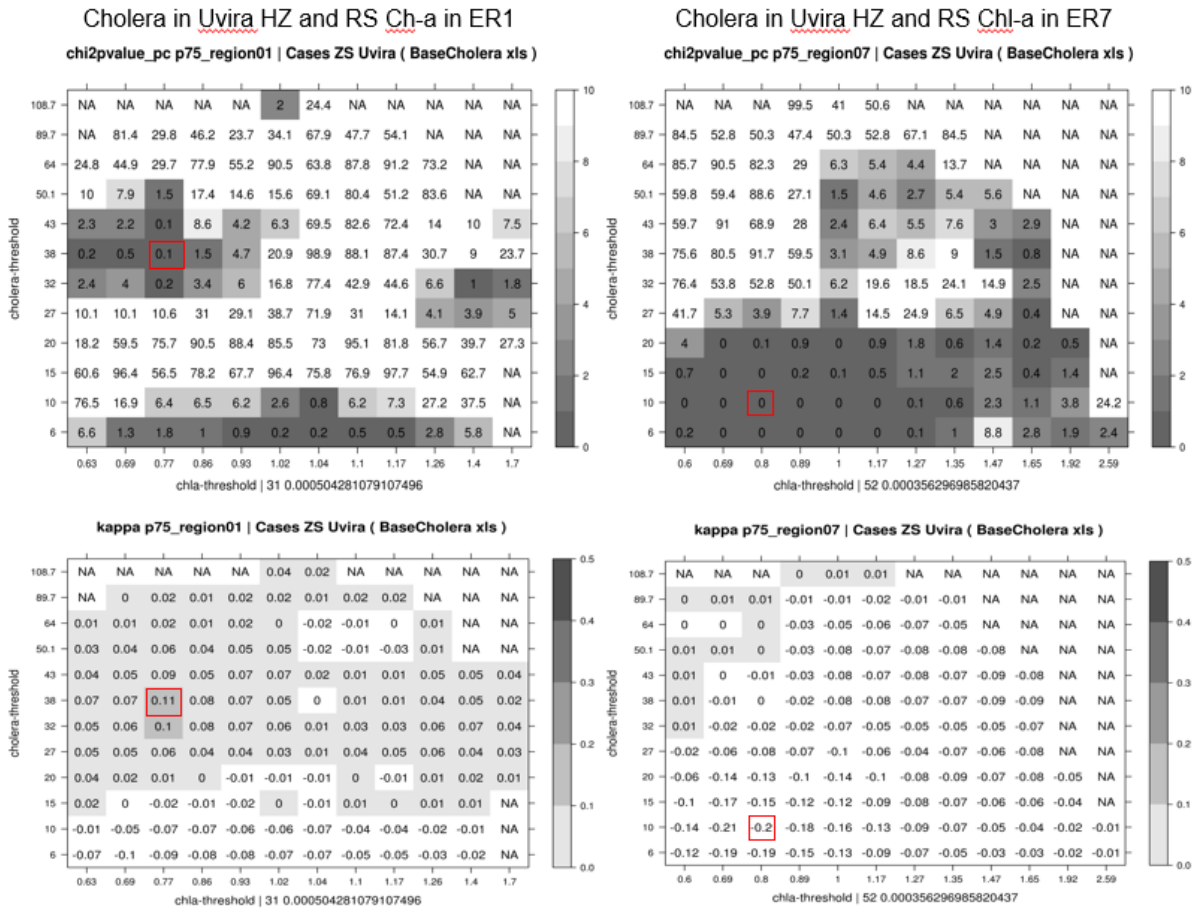
Fig. 58. Top: p-value of Chi-square statistic computed using different thresholds on RS-Chl-a (X) and cholera cases in Uvira HZ (Y) – Darker the cell, the smaller the risk to reject the independence hypothesis. Bottom: Kappa Index (KIA). X (chl-a in µg/l) and Y (N cholera cases).
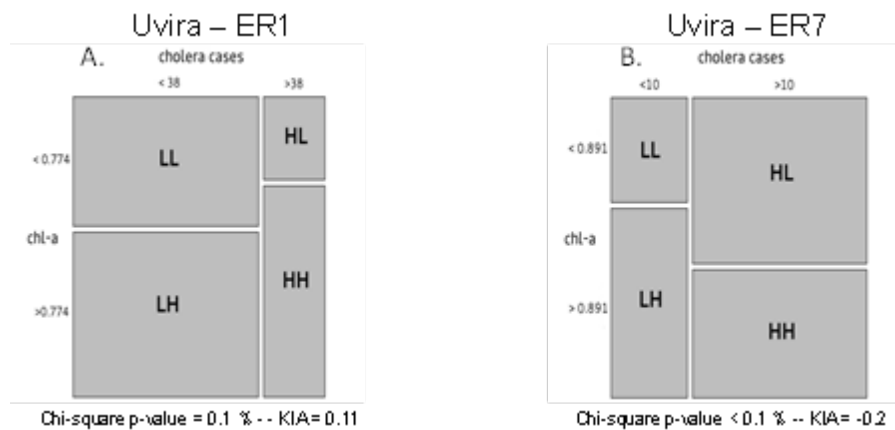

Fig. 59. Contingency tables of combination of threshold values on cholera cases and RS Chl-a concentration that minimizes the p-value of the computed Chi-square statistics and maximizing the KIA value. (A) Cholera in Uvira HZ (threshold = P75 = 38 cases) and RS-Chl-a in ER1 (threshold = P40 = 0.774 mg/m³). (B) Cholera in Uvira HZ (threshold = P40 = 10 cases) and RS-Chl-a in ER7 (threshold = P30 = 0.891 mg/m³).
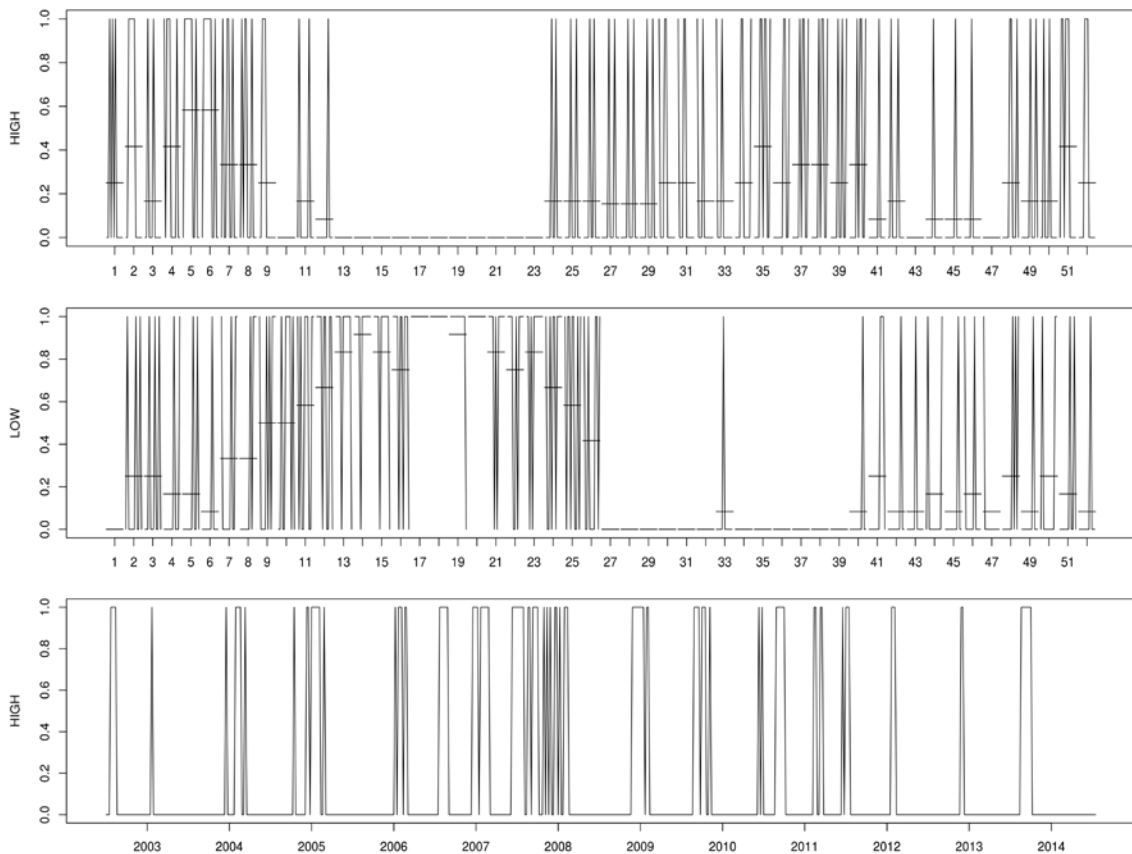
Fig. 60. Week-plots showing the mean-annual (horizontal bar) per week and weekly inter-annual variation of HH co-occurrences (top) and LL co-occurrences (middle) of cholera cases and RS-Chl-a concentration for the couple of threshold values presented in Figure 58 left and Figure 59 left. Bottom: TS of HH co-occurrences (1 on the plot) for the same case.
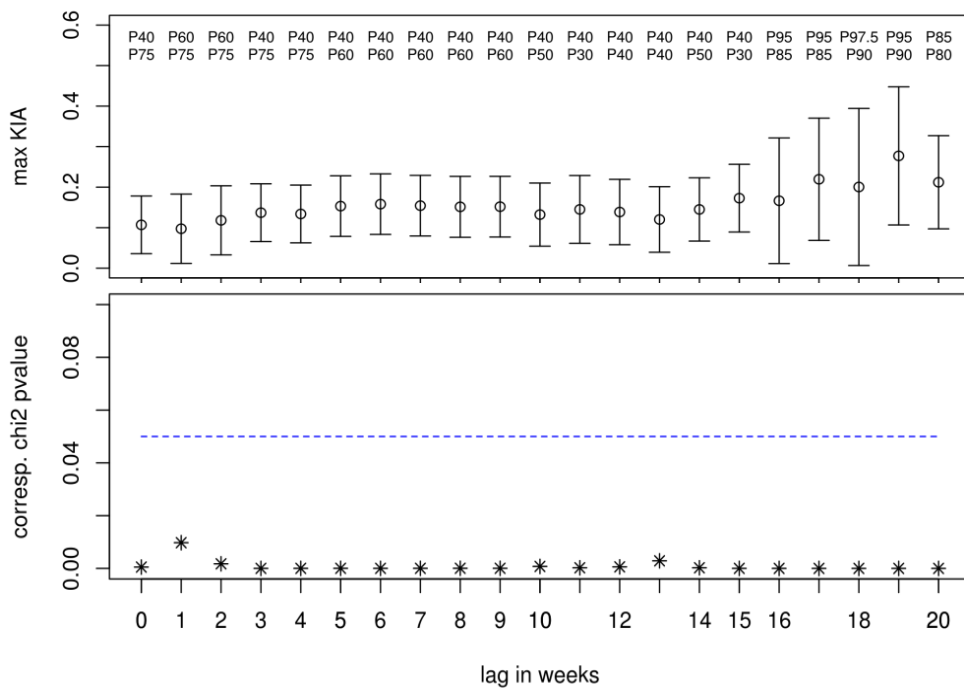

Fig. 61. Analysis of hypothetical delayed association of cholera cases and RS-Chl-a based on TS thresholding. Uvira HZ and RS-Chl-a in ER1.

ARIMA - remainders analysis

The link between remainders of modeled TS using the ARIMA obtained for epidemiological and main environmental variables at Uvira between 2011 and 2014 is reported below.

Leaving out autocorrelation correction at this stage, high Pearson correlation values between selected and unlagged TS were observed (Figure 62A). This is especially the case between meteorology and limnology. RS.chl-a was also highly correlated with RS.dayLSWT and both are generally well correlated to some meteorological and limnological variables, especially RelHum and Temp.surf, and to a lower extent Precip, SolarRad and TranspSD among others. However, fish catches and cholera cases were not significantly correlated to the other variables.

The high correlations values identified are partially linked to synchronous seasonal cycles. Furthermore, they are biased by the non-independence between observations due to temporal autocorrelation and by small numbers in the cases of fish catches and cholera cases. Figure 62B illustrates the cross-correlations between remainders without time lag. All coefficient values dropped drastically except for the relationship between AirTemp and RelHum or SolarRad and between Chla.pel and ChlaZmix. Regarding the correlation between remainders of cholera cases and other variables, values remains non-significant. This does not match well with the first analysis crossing cholera and RS-Chl a that showed weak but significant association. This is probably due to similar seasonal cycles represented in the ARIMA model and thus eliminated in the remainders. We thus investigated correlations variation between remainders in function with various temporally lagged (-Δt) environmental TS (Figure 63 and Table VI) to investigate possible unseasonal (remainders) delayed response of cholera cases. Remainders at t-Δt of some environmental variables can be considered as significant predictors of cholera at time t even with low correlation values.
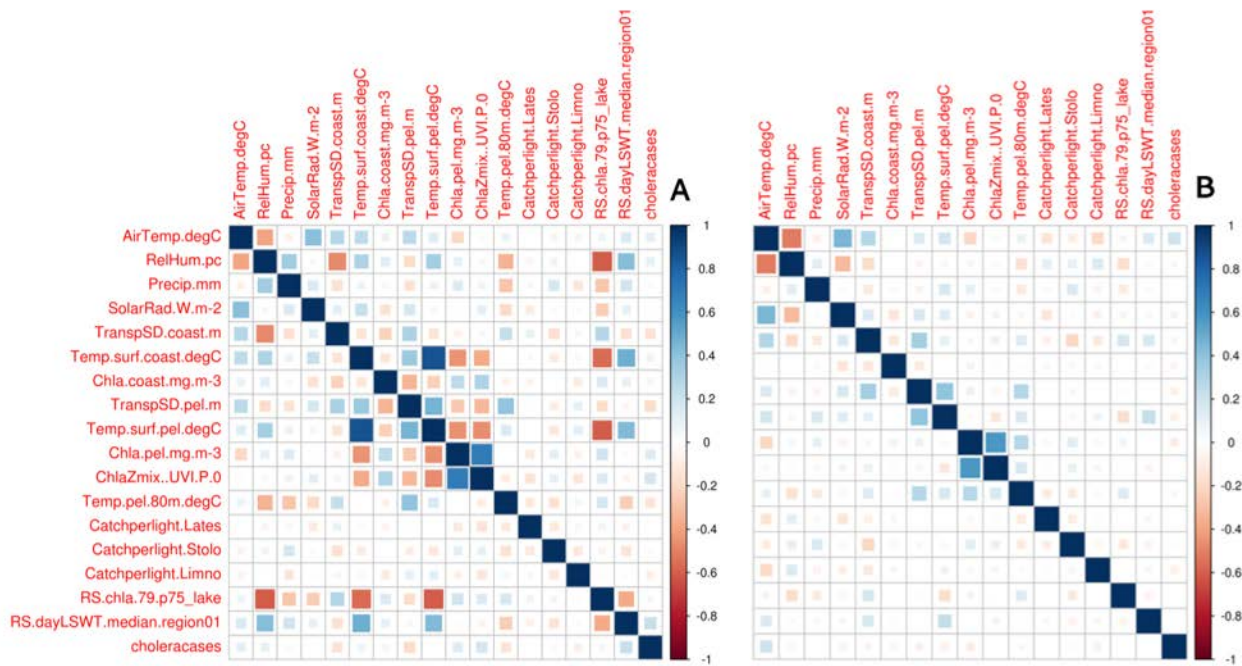
Fig. 62. (A) Pearson correlations between TS of selected variables (A) and ARIMA model remainders of selected variables (B) at Uvira between 2011 and 2014.

Temperatures (Air, Surface pelagic, RS.dayLSWT) are short time (respective Δt = 7, 2 and 6 weeks) predictors, but they reveal incoherent association signs (respectively negative, positive and negative). The sign incoherence can be questioned. TranspSD.coast is also a short-term predictor (Δt = 8 weeks): the increase of TranspSD.coast remainder is associated to an increase of the cholera number of cases (positive correlation).

Precipitation represents short, middle and long term predictors. Precipitation is positively or negatively associated with cholera cases depending on the delay (positive association for Δt = 6 and 13-14 weeks and negative association for Δt = 11 weeks). Chla.pel was positively associated with cholera cases for Δt = 7 and 16 weeks, the rhythmic nature of planktonic blooms related to periodicities of internal waves could however induce spurious correlations as such a very long delay between planktonic blooms linked to cholera cases seems unlikely.

Similar lags are observed regarding association between remainders of some variables with remainders of cholera:

- 6-7-8 weeks for negative association of air temperature and the lake surface temperature (from RS).
- 6-7-8 weeks also for positive association between choleras cases and precipitation, water transparency and planktonic abundance.

- Cyclic association pattern for precipitation of 6 (+), 11 (-) and 13-14 (+) weeks in phase with water transparency and phytoplankton abundance in the pelagic area (6-7 weeks).
- Phase opposition between air temperature and cholera cases (6-7 weeks).

The 13-14 weeks lagged positive association of precipitation with cholera cases that is 3 or 4 weeks postponed with respect to Chla-coast and ChlaZmix (respectively Δt = 17 and 16 weeks)
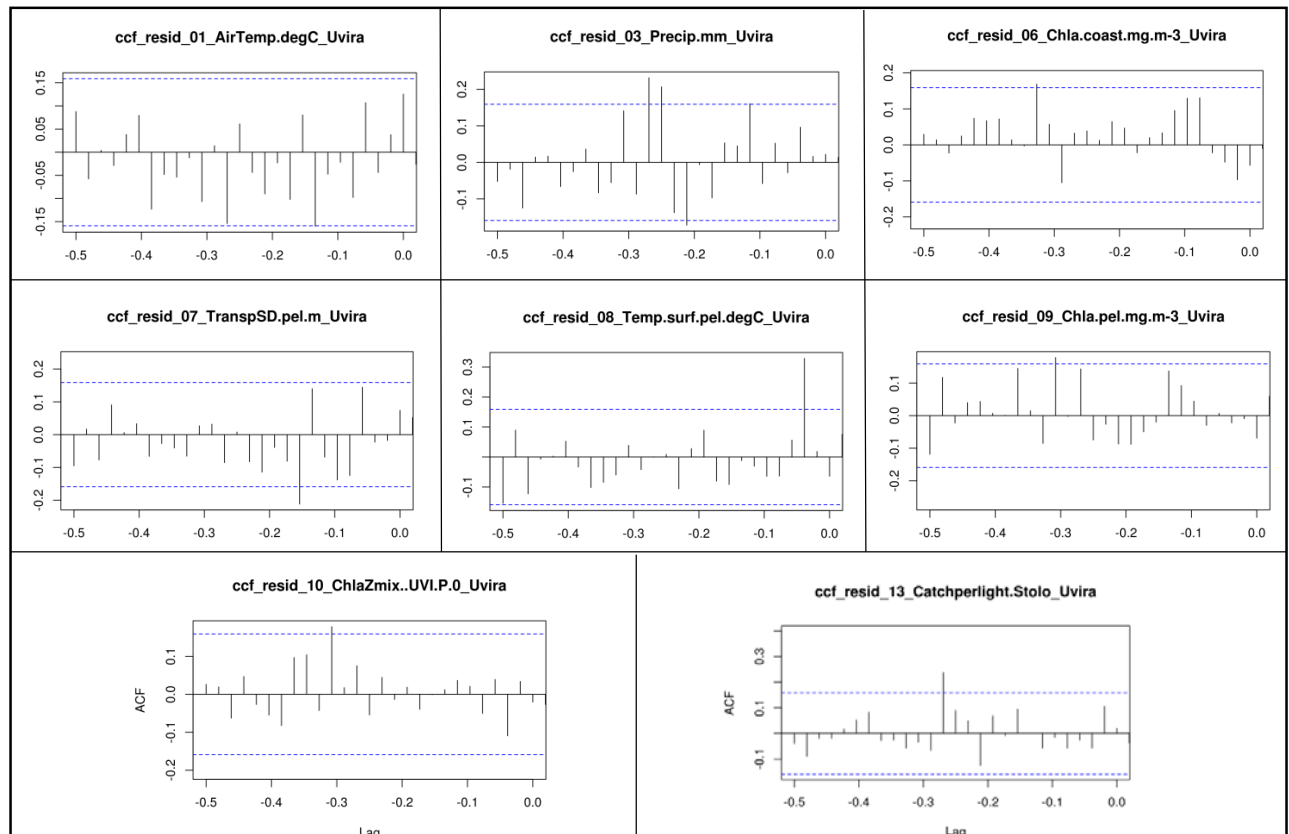

Fig. 63. Pearson correlations between TS remainders of selected environmental variables and remainders of cholera cases after ARIMA modeling in Uvira (2011-2014) for different temporal lags (in year). Negative lags for environmental variable remainders are possible predictors of remainder of cholera cases at time t considering expected ARIMA model number.

Table VI. Synthesis of significant correlation for negative lags between remainders (ARIMA model) of the x-variables (environmental) and cholera cases in Uvira between 2011 and 2014.

Variables remainders	dt (weeks)	Prediction delay	Sign of r
Surface temperature in pelagic zone (Temp.surf.pel)	2	Very short term	+
Surface lake temperature (RS.dayLSWT)	6	Short term	-
Air temperature (AirTemp)	7	Short term	-
Water transparency in pelagic area (TranspSD.pel)	8	Short term	+
Precipitations	6 11 13-14	Short, middle or long term	+ - +
Chl a in pelagic water (Chla.pel)	7 16	Short or long term	+ +

ARIMA - modeled TS analysis

As a complement of the remainders analysis reported above, we discuss here interactions between cholera and environmental variables with particular attention to the seasonal and autoregressive components of the TS.

The modeled Cholera TS (Figure 64) often displays two periods of epidemic peaks that were also well identified via binomial analysis:

(1) **Wet season peaks**

Peaks of cholera cases at Uvira HZ associated with peaks of chl-a were observed during the wet season. This was observed during the first two CHOLTIC monitoring annual cycles (between December and March). Interannual climatic variability might explain differences in timing of patterns from year to year.

(2) **End of wet and dry season peaks**

Peaks of cholera cases at Uvira are sometimes observed after mid-year (Figs. 60 and 64). Those are generally characterized by less frequent cholera cases although a strong cholera outbreak was noted in April 2014. This is a good example of peaks in the transition month at the end of the rain season (Nb: all planktonic blooms may not be detected via remote sensing due to cloudiness).
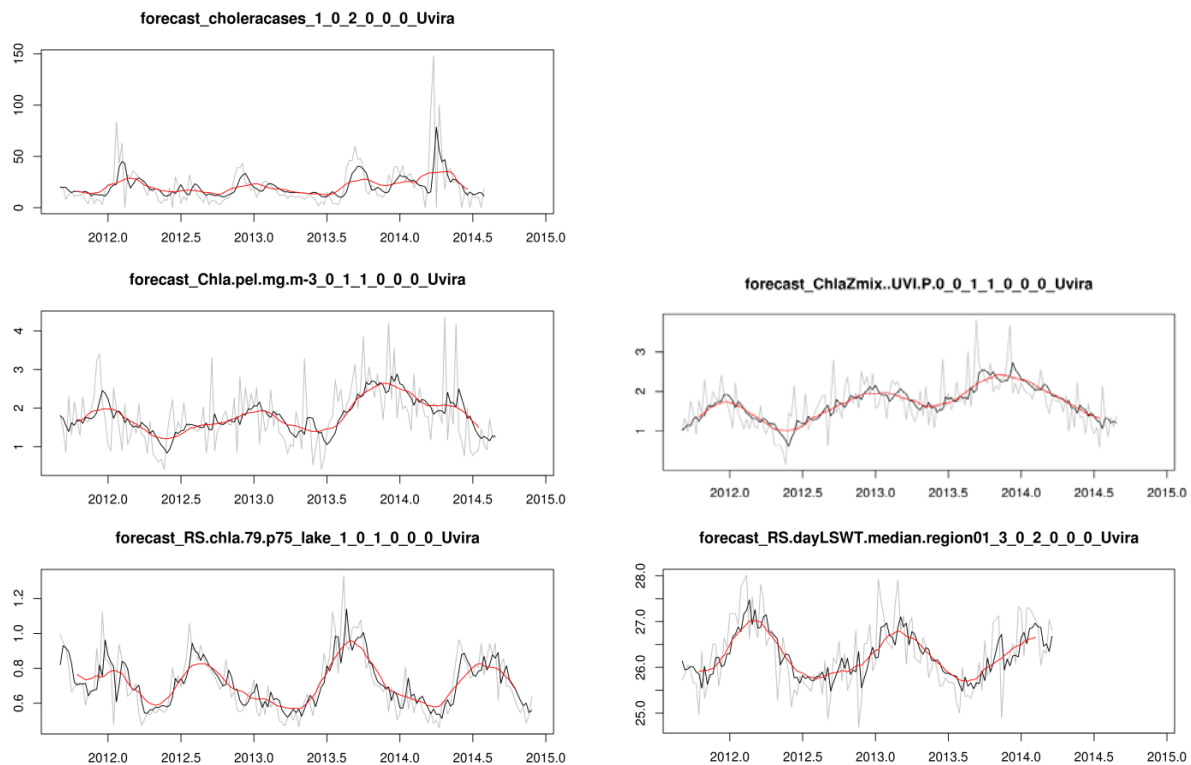
Fig. 64. Grey curve: observed TS of cholera cases or environmental variables. Black curve: TS of cholera cases or environmental variables produced using ARIMA model. Red curve: trimester smoothing of the black curve.

Modeled cholera TS pattern shows, after a first peak, a damped wave of secondary peaks. This pattern is comparable to various other lacustrine variables TS of the CHOLTIC project (i.e., Figures 3, 5, 6 and 7) corresponding to previous observations (Plisnier & Coenen, 2001) and indicate probable links with planktonic organisms. During plankton blooms the alkaline nature of Lake Tanganyika (average pH of 9.2) is still increased and pH may reach values >9.5 or even >10. This may favor *Vibrio cholera as well as* increased organic conditions.

It is possible however that the observed cholera pattern reflect an intrinsic pattern evolution (Sanches et al., 2011), which could be similar but not directly related to the environmental TS. Bacteriological analysis of a great number of samples would be the only way to investigate this more firmly.

In some way, the temporal pattern of modeled TS for some environment variables shows some similarities with reconstructed cholera TS. For instance, Chla.pel and ChlaZmix models are in agreement with modeled TS of cholera (increased peak values after mid-2013). Nevertheless, for all TS, the variance of the modeled TS is generally much lower than the variance of remainders that have surely more significant effects on the importance of the outbreaks as indicated above.

It is interesting to note that patterns of cholera peaks and environmental pattern may be linked to the seasonality with some highest peaks after season changes. However during the rains, the number of cases seems higher. This could be the effect of rainfalls on higher epidemics diffusion. But with the data set used in CHOLTIC we were not able to demonstrate statistically significant links between epidemiology and rain.

Conclusions

The main objective of CHOLTIC was to test possible links between environmental data and cholera outbreaks in the Lake Tanganyika area. This is formulated by the following hypothesis:

H0 hypothesis: *"Environmental factors impact cholera outbreaks, the lake being a possible reservoir of V. cholerae at the origin of the outbreaks".*

If this hypothesis can be verified, the lake being very sensible to climatic conditions (Coulter, 1991; Plisnier et al. 1999), the changes of climate may explain why *V. cholerae* has found favorable conditions to persist and regularly reach infectious level of abundance.

From the CHOLTIC results presented in this report and the above analysis, several elements support the H0 hypothesis:

- Spatio-temporal interaction between cholera outbreaks in Uvira HZ and RS-Chl-a concentration in the lake.

- Damped oscillation of cholera cases associated with both seasons was observed from the ARIMA model of cholera TS. This appears to be a similar type of pattern as those observed in various lacustrine related variables (T°, chl-a, clupeid catches). However, during the very short project period, intermediate outbreaks may also appear in the TS with less clear links to the seasonal evolution of environmental variables.

- Relationship between remainders of ARIMA model for cholera cases and the main environmental variables, show that abnormally high pelagic surface temperature seems to have a strong and positive two-week delayed effect on cholera cases with regard to the expected number of cases.

- Epidemiology of various health areas in the DRC (chap 2, WP 6) indicates that cholera epidemics develop earlier in lacustrine zones, thereby confirming the probably lacustrine origin of cholera outbreaks in the region (Bompangue et al. 2012).

Historically, a close relationship has been shown with access to lake and drinking lake water by local populations during the 1992 cholera outbreak in Burundi (Birmingham et al. 1997).

The results allow considering early warning methods development. However, at the present stage, the information is not statistically significant enough as a basis for an operational cholera warning. The present results indicate patterns and correlation with the environment that need (1) to be confirmed with longer time series and (2) to be confirmed by intensive bacteriological sampling and analysis to check the strains into the lake and identify links with plankton or other lacustrine organisms.

Peaks of cholera and chl-a during the dry seasons are particularly interesting as they show that rainfall, often quoted in relation to cholera epidemics, should not drive all attention in cholera epidemics studies. Rainfalls are linked to a variety of other meteorological and environmental changes. A particular attention to first cholera cases and conditions of infections should be one of the priorities for further investigation as development of cholera may be influenced by a variety of factors, including rainfalls, which could make more difficult the detection of very initial outbreak conditions.

The probable decrease of primary production at the lake (O'Reilly et al. 2003) on average is only one of the aspect of climate change at Lake Tanganyika were several other changes have been observed: thermic stratification change (stability increase), decreased depth of the oxic layer (Van Bocxlaer et al. 2012), changes in planktonic species composition (Cocquyt & Vyverman 2005) may be related to conditions favorable for the establishment of *Vibrio cholerae* in the lake after probable initial infection by a traveler to the lake (Malengreau et al. 1979).

Other alternative hypotheses are the following:

H1 hypothesis: *"Environmental factors don't impact cholera outbreaks, specifically the lake is not a reservoir of V. cholerae at the origin of the outbreaks"*

H2 hypothesis: *"Humans, their behavior and their displacements are the main sources as well as transmissions vectors associated with cholera outbreaks".*

The present result may not reject H1 although initial results appear to weight more toward the first H0 hypothesis.

Concerning H2, in the frame of CHOLTIC, it was not foreseen to check this hypothesis but human behavior and activities are known elsewhere (Nelson et al. 2009) to play a role in cholera outbreaks. In the frame of CHOLTIC, genetics of cholera strains (cfr chap 2, WP8) has shown results mainly linked with anthropic infections although this is related also to available samples that were mainly originating from humans. The number of environmental samples need to be increased in the future as much as possible, which is often a challenge in the region for logistical and sometimes security reasons.

It is concluded that the H0 hypothesis concerning environmental relationships with cholera is supported by several statistical and ecological relationships at Lake Tanganyika. However, we can not exclude that humans have an important impact. Lacustrine and meteorological conditions could favor the development of *Vibrio cholerae*, particularly as the lake presents a high average pH particularly during periods of planktonic blooms. In the last 40 years, various ecological changes have taken place in the lake, largely caused by climate changes (particularly air temperature and also probable wind decrease). Changes in thermic stratification, oxygenation and primary production patterns of Lake Tanganyika may have been favorable for reservoir(s) organism(s) and/or transmission of *Vibrio cholerae* in the lake area.

3. POLICY SUPPORT

The CHOLTIC project has brought to attention, of various stakeholders including authorities and local researchers, the possible links between lacustrine environment and the cholera outbreak, which without doubt will be beneficial for increased attention and progress in the understanding of cholera outbreaks and adaptations.

Various teams in Africa and Europe have worked together toward the goal of CHOLTIC. This experience will be valuable to evaluate and develop further collaboration. Donors and funding attention will grow in parallel to research reporting in publications already published and other in preparation from the project's results.

As a support for the decision, the following recommendations are proposed:

Local authorities

-Increase population awareness on the need to use safe and clean water for drinking
- develop better infrastructure of health for areas which are more isolated and people have are at higher risks
-support researchers in the health and environment field with appropriate training, material and facilities to develop research to progress into the knowledge of cholera fundamental as well as applied researches.

International organization and donors:

-Increased attention on the cholera disease. The recent advancements in researches and the need to achieve an understanding at a regional level including various areas of endemicity in Africa, particularly around the Great African Lakes should be made more known. International organization and donors are required to support not only adaptation to fight cholera outbreaks (such as better drinking water networks) but also multidisciplinary research. This involves by definition a great number of disciplines in the field of health and environment. Understanding and forecasting cholera requires this effort to try to decrease and as much as possible eliminate Cholera.

Researchers:

-Develop multidisciplinary collaboration efforts taking into account the acquired experience to overcome any difficulties to reach the objectives.

-Sensitive all stakeholders on the need to document patients history at a fine geographical scale and build longest possible time series of the best possible quality. Investigation on early cases of the outbreaks is particularly important.

-In the case of Lake Tanganyika, the coastal areas often displaying important phytoplanktonic blooms requires further research and bacteriological analysis in the future.

-A great number of analysis on environmental samples would be usefull to be able to identify the organisms suspected to be reservoirs of *V.cholerae*.

-A study on diffusion of cholera at a detailed geographical scale would be usefull.

4 DISSEMINATION AND VALORISATION

4.1 Outcomes

CHOLTIC has enabled the identification of various environmental parameters significantly correlated with cholera outbreaks. This will increase more direct specific cholera research on lacustrine environments in close collaboration with local researchers, authorities in various fields (environment, health). Challenges of carrying such studies in a remote environment have been useful to identify ways and methods to improve research into this field while various procedures such as for remote sensing processing, bacteriology or ecological modeling have been progressing to know better the lake conditions and applicable methods to study the cholera topics.

The project specific outcomes include:

- a multidisciplinary database in limnology, plankton, fisheries, epidemiology, bacteriology methods, genetic as research tools for cholera outbreaks investigations;

- an improved eco-hydro model to understand particularly the climate impact on the lake hydrodynamics and ecology as a tool to understand triggering condition of cholera outbreaks;

- statistical relationships correlating environmental variables and cholera outbreaks. The identification of types, locations and timing of phytoplanktonic blooms (inducing zooplankton blooms in the trophic chain and sometimes fish and fishermen migrations) is in full agreement with statistical analysis on the association between cholera epidemics and chl-a in lake waters estimated from RS.

- various publications already published in peer reviewed international journals and other in preparation.

4.2 Valorisation

CHOLTIC project results have contributed to understand better the cholera outbreaks and their possible links to environmental data in the region of Lake Tanganyika.

Reports, publications, reinforcement of capacities, meetings with various groups of stakeholders have largely allowed to valorize many of the project results already, while

acquired result will be useful for future investigation of cholera-environment relationships.

Local climate, limnology (field and RS measurements), fisheries, plankton, hydrodynamic and ecological modeling, bacteriology and genetics had not been previously investigated in relation to cholera epidemics in a great African lake. An increased attention on cholera outbreaks and relationships of *V.cholerae* with the African Great Lakes will results from the CHOLTIC project.

5 PUBLICATIONS

Peer review publications

Bompangue Nkoko D., Giraudoux P., Plisnier P.-D., Mutombo Tinda A., Piarroux M., Sudre B., Horion S., Muyembe Tamfum J.-J., Kebela Ilunga B. and Piarroux R. 2011. Dynamics of Cholera Outbreaks in Great Lakes Region of Africa, 1978–2008. Emerging Infectious Diseases 17, 11, 2026-2034

Cocquyt C., de Haan M., Jahn R. and Hinz F. 2012. *Nitzschia epiphytica*, *N. epiphyticoides*, and *N. pseudepiphytica* (Bacillariophyta), three small diatoms from East and Central Africa. Phycologia 51, 2, 126-134.

Kraemer B., Hook S., Huttula T., Kotilainen P., M O'Reilly C., Peltonen A., Plisnier P.-D., Sarvala J., Tamatamah R., Vadeboncoeur Y., Wehrl B. and McIntyre P. B. Spatiotemporal variation in long-term temperature trends in Lake Tanganyika. PLOS ONE, accepted.

Moore S., Miwanda B., Yao Sadji A., Thefenne H., Jeddi F., Rebaudet S., de Boeck H., Bidjada B., Depina J.-J., Bompangue D., Aruna Abedi A., Koivogui L., Keita S., Garnotel E., Plisnier P.-D., Ruimy R., Thomson N., Muyembe J.-J. and Piarroux R.. 2015 Relationship between Distinct African Cholera Epidemics Revealed via MLVA Haplotyping of 337 *Vibrio cholerae* Isolates. PLoS Negl Trop Dis. 2015 (PNTD-D-15-00132R1).

Naithani J., Plisnier P.-D. and Deleersnijder E. 2011. Possible effects of global climate change on the ecosystem of Lake Tanganyika. Hydrobiologia 671: 147-163.

Naithani J., Plisnier P.-D. and Deleersnijder E., 2012. Modelling the ecohydrodynamics of Lake Tanganyika. In: Bengtsson L., Herschy R.W. and Fairbridge R.W. (Eds), Encyclopedia of Lakes and Reservoirs. Springer, New York.

Sharma, S. and al. (Plisnier, P-D), 2015. A global database of lake surface temperatures collected by in situ and satellite methods from 1985–2009. Scientific Data 2, 150008

Thiery W., Martynov A., Darchambeau F., Descy J.-P., Plisnier P.-D., Sushama L. and van Lipzig N. P. M. 2014. Understanding the performance of the FLake model over the African Great Lakes. Geoscientific Model Development 7, 317–337

Van Bocxlaer B., Schultheiß R., Plisnier P.-D. and Albrecht C. 2012. Does the decline of gastropods in deep water herald ecosystem change in Lakes Malawi and Tanganyika? Freshwater Biology 57, 8, 1733-1744.

Other publications

Cocquyt C. and Plisnier P.-D. (2013) Fytoplankton uit het Tanganyika-meer en Cholera. Diatomededelingen 37, 27-33.

Maréchal C., Cawoy V., Cocquyt C., Dauby G., Dessein S., Douglas-Hamilton I., Dupain J., Fischer E., Fouth Obang D., Groom Q., Henschel P., Jeffrey K.J., Korte L., Lewis S.L., Luhunu S., Maisels F., Melletti M., Ngoufo R., Ntore S., Palla F., Scholte P., Sonké B., Stevart T., Stoffelen P., Van den Broeck D., Walters G. & Williamson E.A. 2014. Biodiversity conservation and management. In: de Wasseige C., Flynn J., Louppe D.,

Hiol Hiol F. and Mayaux Ph. (eds) The forests of the Congo Basin – State of the Forest 2013. Weyrich édition, Neufchâteau Belgium: 67-96.

Plisnier P.-D., Bompangue D., Piarroux R., Cocquyt C., Jacobs J., De Boeck H., Giraudoux P., Deleersnijder E., Naithani, Hanert E., Muyembe J.-J., Miwanda B., Nshombo V., Poncelet N. and Cornet Y. 2011. CHOLTIC: Epidémies de Cholera au lac Tanganyika induites par les changements climatiques? / Cholera-epidemieën in het Tanganyikameer veroorzaakt door klimaatveranderingen? Science Connection 35, 38-42.

Plisnier P-D 2012. Lake Tanganyika Regional Integrated Environmental Monitoring Programme - Implementation Strategy. Partnership Interventions for the Implementation of the Strategic Action Programme for Lake Tanganyika. UNDP-GEF, 91 pp.

Verschuren D., Plisnier P.-D., Cocquyt C., Hughes H., Lebrun J.; Gelorini V., Rumes B., Mahy G. and André L. 2012. Climatic and anthropogenic impacts on African Ecosystems. Final report. Brussels, Belgian Sciences Policy 2011, 93 pp.

6 ACKNOWLEDGMENTS

7   REFERENCES

Ahmad Z., Franz B.A., McClain C.R., Kwiatkowska E.J., Werdell J., Shettle E.P. and Holben B.N., 2010. New aerosol models for the retrieval of aerosol optical thickness and normalized water-leaving radiances from the SeaWiFS and MODIS sensors over coastal regions and open oceans. Applied Optics 49, 29, 5545-5560.

Ali M., Emch M., Donnay, J.P., Yunus M. and Sack R.B., 2002. Identifying environmental risk factors for endemic cholera: a raster GIS approach. Health & Place 8, 201-210.

Bergamino N., Horion S., Stenuite S., Cornet Y., Loiselle S., Plisnier P.-D. and Descy J-P., 2010 Spatio-temporal dynamics of phytoplankton and primary production in Lake Tanganyika using a MODIS based bio-optical time series. Remote Sensing of Environment 114, 4, 772-780.

Birmingham M.E., Lee L.A., Ndayimirije N., Nkurikiye S., Hersh B.S., Wells J.G. and Deming M.S., 1997. Epidemic cholera in Burundi: patterns of transmission in the Great Rift Valley Lake region. The Lancet 349, 9057, 981-985.

Bompangue D., Giraudoux P., Handschumacher P., Piarroux M., Sudre B., Ekwanzala M., Kebela I. and Piarroux R., 2008. Lakes as source of cholera outbreaks, Democratic Republic of Congo. Emerging Infectious Diseases 14, 5, 798-800.

Bompangue D., Giraudoux P., Piarroux M., Mutombe G., Shamavu R., Sudre B., Mutombo A., Mondonge V. and Piarroux R., 2009. Cholera epidemics, war and disasters around Goma and Lake Kivu: an eight-year survey. PLoS Neglected Tropical Diseases, May, doi/10.1371/journal.pntd.0000436.

Bompangue D.,Giraudoux P., Plisnier P.-D., Mutombo A., Piarroux M., Sudre B., Horion S., Muyembe J.J., Kebela B. and Piarroux R., 2011. Dynamics of Cholera Outbreaks in Great Lakes Region of Africa, 1978–2008. Emerging Infectious Diseases 17, 11, 2026-2034.

Bompangue D, Vesenbeckh S.M., Giraudoux P., Castro M., Muyembe JJ, Ilunga B.K. and Murray M., 2012. Cholera ante portas – The re-emergence of cholera in Kinshasa after a ten-year hiatus. PLoS Currents Disasters 1, 1-12.

Campbell J.W. and O'Reilly J.E., 2006. Metrics for Quantifying the Uncertainty in a Chlorophyll Algorithm: Explicit equations and examples using the OC4.v4 algorithm and NOMAD data. Papers from the Ocean Color Bio-optical Algorithm Mini-Workshop, 27-29 Sept. 2005, 748. http://oceancolor.gsfc.nasa.gov/MEETINGS/OCBAM/docs/PerformanceMetrics.pdf   (latest consultation 26/5/2015)

Chitamwebwa D., 1999. Meromixis, stratification and internal waves in Kigoma waters of Lake Tanganyika. Hydrobiologia 407, 59–64.

Cleveland R.B., Cleveland W.S., McRae J.E. and Terpenning I., 1990. A seasonal-trend decomposition procedure based on Loess. Journal of Official Statistics 6, 1, 3-73.

Cocquyt C. and Vyverman W., 2005. Phytoplankton in Lake Tanganyika: a comparison of community composition and biomass off Kigoma with previous studies 27 years ago. Journal of Great Lakes Research 31, 535-546.

Colwell, R., 1996. Global climate and infectious disease: The cholera paradigm. Science 274, 2025-2031.

Coulter G.W., 1968. Hydrological processes and primary production in Lake Tanganyika. Proceedings 11th Conference Great Lakes Research, 609-626.

Coulter G.W. (ed.), 1991. Lake Tanganyika and its Life. Oxford University Press, London, Oxford and New York.

Dussart B.H., 1980. Les crustacés copépodes d'Afrique, catalogue et biogeographie. Hydrobiologia, 72, 165-170.

Dussart B.H. and Defaye D, 2001. Introduction to the Copepoda. (2nd Edition) In: Dumont, H. (Ed.), Guide to the iIdentitification of the microinvertebrates of the Continental Waters of the world. Backhuys Publishers, Leiden.

Feldman G.C. and McClain C.R., 2014. Ocean Color Web, MODIS Aqua and Terra v6. NASA Goddard Space Flight Center. http://oceancolor.gsfc.nasa.gov/ (access on December 2014).

François C., Brisson A., Le Borgne P. and Marsouin A., 2002. Definition of a radiosounding database for sea surface brightness temperature simulations - Application to sea surface temperature retrieval algorithm determination. Remote Sensing of Environment, 81, 309-326.

Franz B.A., Kwiatkowska E.J., Meister G. and McClain C.R., 2007. Utility of MODIS-Terra for Ocean Color Applications. Proceedings SPIE 6677, 14 pp.

Gordon H.R. and Wang M., 1994. Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: a preliminary algorithm. Applied Optics 33, 3, 443-452.

Griffith, D.C., Kelly-Hope L.A. and Miller M.A., 2006. Review of reported cholera outbreaks worldwide, 1995–2005. American Journal of Tropical Medicine and Hygiene 75, 5, 973-977.

Haney J.F., 1988. Diel patterns of zooplankton behavior. Bulletin of Marine Science 43, 583-603.

Hashizume M, Armstrong B., Hagat S., Wagatsuma Y., Faruque A.S.G., Hayashi T. and Sack D.A., 2008. The effect of rainfall on the incidence of cholera in Bengladesh. Epidemiology 19, 1, 103-110.

Hecky R.E. and Kling H., 1981. The phytoplankton and protozooplankton of the euphotic zone of Lake Tanganyika: Species composition, biomass, chlorophyll content, and spatio-temporal distribution. Limnology and Oceanography 26, 548-564.

Hecky R. E. and Kling H., 1987. Phytoplankton ecology of the great lakes in the rift valleys of Central Africa. Archiv für Hydrobiologie -– Beiheft Ergebnisse der Limnologie 25, 197-228.

Horion S., Bergamino N., Stenuite S., Descy J.-P., Plisnier P.-D., Loiselle S.A. and Cornet Y., 2010 Optimized extraction of daily bio-optical time series derived from MODIS/Aqua imagery for Lake Tanganyika, Africa, Remote Sensing of Environment 114, 4, 781-791.

Isumbisho M., Sarmento H., Kaningini B., Micha J.-C.and Descy J.-P., 2006. Zooplankton of Lake Kivu, East Africa, half a century after the Tanganyika sardine introduction. Journal of Plankton Research 28, 971–989.

Ivlev V.S. 1945. The biological productivity of waters. Usp. sovrem. Biol. 19, 98-120.

Jørgensen S.E.and Bendoricchio G., 2001. Fundamentals of Ecological Modelling. Elsevier, Oxford.

Knox A., Bertuzzo E., Mari L., Odermatt D., Verrecchia E. and Rinaldo A., 2014. Optimizing a remotely sensed proxy for plankton biomass in Lake Kivu. International Journal of Remote Sensing 35, 13, 5219-5238.

Lagne Férandez M.A., Bauernfeind A., Jiménez J.B., Gil C.L., El Omeiri N. and Guibert D.H., 2009. Influence of temperature and rainfall on the evolution of cholera epidemics in Lusaka, Zambia 2003-2006: analysis of a time series. Transactions of the Royal Society of Tropical Medicine and Hygiene 103, 2,137-143.

Lipp E.K., Huq A. and Colwell R.R., 2002. Effects of global climate on infectious disease: the cholera model. Clinical microbiology reviews 15, 4, 757-770.

MacCallum S. and Merchant C.J., 2013. ATSR Reprocessing for Climate Lake Surface Water Temperature: ARC-Lake, University of Edinburgh. http://datashare.is.ed.ac.uk/handle/10283/88 (last access on 26 May 2015).

Malengreau M., Gillieaux M., De Feyter M. and Wittman L., 1979. A propos de l' épidémie de choléra à l'est du Zaïre en 1978. Annales de la Société belge de médecine tropicale 59, 401-412.

Marchant R. Mumbi C., Behera S. and Yamagata T., 2007. The Indian Ocean dipole - the unsung driver of climatic variability in East Africa. African Journal of Ecology 45, 1, 4-16.

McCarthy S.C., Gould R.W.Jr., Richman J., Kearney C. and Lawson A., 2012. Impact of Aerosol Model Selection on Water-Leaving Radiance Retrievals from Satellite Ocean Color Imagery. Remote Sensing 4, 3638-3665.

McClain E.P., Pichel W.G. and Walton C.C., 1985. Comparative performance of AVHRR-based multichannel sea surface temperatures. Journal of Geophysical Research: Oceans 90, 11587-11601.

Ministère Congolais de l'Environnement, des Eaux et Forêts et de la Conservation de la Nature, 2005. Diagnostic du secteur forestier en RDC et problématique de la gestion. Forum national de consultation des acteurs du secteur forestier de la RDC. Kinshasa, 1-18.

Ministère de la santé publique de la RD Congo 2011. Surveillance intégrée de la maladie et ripostes. Kinshasa, 209 pp.

Mutreja A., Kim D.W., Thomson N.R., Connor T.R., Lee J.H., Kariuki S., Croucher N.J., Choi S.Y., Harris S.R., Lebens M., Niyogi S.K., Kim R.J., Ramamurthy T., Chun J., Wood J.L.N., Clemens J.D., Czerkinsky C., Nair G.B., Holmgren J., Parkhill J. and Dougan G., 2011.Evidence for several waves of global transmission in the seventh cholera pandemic. Nature 477, 462- 564.

Naithani J., Darchambeau F., Deleersnijder E., Descy J.-P. and Wolanski E., 2007a. Study of the nutrient and plankton dynamics in Lake Tanganyika using a reduced-gravity model. Ecological Modelling 200, 225–233.

Naithani J., Deleersnijder E. and Plisnier P.-D., 2002 Origin of intraseasonal variability in Lake Tanganyika. Geophysical Research Letters 29, 23, DOI: 10.1029/ 2002GL015843.

Naithani J., Deleersnijder E. and Plisnier P.-D., 2003. Analysis of wind-induced thermocline oscillations of Lake Tanganyika. Environmental Fluid Mechanics 3, 23-39.

Naithan J., and Deleersnijder E., 2004. Are there internal Kelvin waves in Lake Tanganyika? Geophysical research letters 31, 6, DOI: 10.1029/ 2003GL019156.

Naithani J., Plisnier P.-D. and Deleersnijder E., 2007b. A simple model of the eco-hydrodynamics of the epilimnion of Lake Tanganyika. Freshwater Biology 52, 2087–2100.

Naithani J., Plisnier P.-D. and Deleersnijder E., 2011. Possible effects of global climate change on the ecosystem of Lake Tanganyika. Hydrobiologia 671, 147–163.

Nelson E. J., Harris J. B., Morris J. G., Calderwood S. B. and Camilli A. 2009. Cholera transmission: the host, pathogen and bacteriophage dynamic. Nature Reviews Microbiology, 7(10), 693-702.

O'Reilly C.M, Alin S.R., Plisnier P.-D., Cohen A.S. and McKee B.A., 2003. Climate change decreases aquatic ecosystem productivity of Lake Tanganyika, Africa. Nature 424: 766-768.

Parsons T.R., Le Brasseur R.J. and Fulton J.D., 1967. Some observations on the dependence of zooplankton grazing on cell size and concentration of phytoplankton blooms. Journal of the Oceanographical Society of Japan 23, 10-17.

Piarroux R., Bompangue D., Oger P, Haaser F., Boinet A. and Vandevelde T. 2009. From research to field action: example of the fight against cholera in the Democratic Republic of Congo.Field Action Sci.Rep., 1, 1-9.

Plisnier P.-D., Chitamwebwa D., Mwape L., Tshibangu K., Langenberg V. and Coenen E., 1999. Limnological annual cycle inferred from physical-chemical fluctuations at three stations of Lake Tanganyika. Hydrobiologia 407, 45-58.

Plisnier P.-D. and Coenen E., 2001. Pulsed and dampened annual limnological fluctuations in Lake Tanganyika. In: Munawar M. and Hecky R. (Eds), The Great Lakes of the World (GLOW): Food-web, health and integrity. Ecovision World Monograph Series, Leiden, The Netherlands: 81-94.

Plisnier P.-D, Mgana H., Kimirei I., Chande A., Makasa L., Chimanga J., Zulu F., Cocquyt C., Horion S., Bergamino N., Naithani J., Deleersnijder E., André L., Descy J.-P. and Cornet Y., 2009. Limnological variability and pelagic fish abundance (*Stolothrissa tanganicae* and *Lates stappersii*) in Lake Tanganyika. Hydrobiologia 625, 117-134.

Price J.F., 1979. On the scaling of stress-driven entrain- ment experiments. Journal of Fluid Mechanics 90, 509–529.

Rebaudet S., Sudre B., Faucher B. and, Piarroux R., 2013a. Environmental determinants of cholera outbreaks in inland Africa: A Systematic Review of Main Transmission Foci and Propagation Routes. Journal of Infectious Diseases 208, Suppl. 1, S46–S54.

Rebaudet S., Sudre B., Faucher B.and Piarroux R., 2013b. Cholera in coastal Africa: a systematic review of its heterogeneous environmental determinants. Journal of Infectious Diseases, 208 Suppl. 1, S98–S106.

Rebaudet S., Mengel M.A., Koivogui L., Moore S., Mutreja A., Kande Y., Yattara O., Sarr Keita V., Njanpop-Lafoucade B.-M., Fournier P.E., Garnotel E., Keita S. and Piarroux R., 2014. Deciphering the Origin of the 2012 Cholera Epidemic in Guinea by Integrating Epidemiological and Molecular Analyses. PLoS Neglected Tropical Diseases, June, doi: 10.1371/journal.pntd.0002898.

Sanches R.P., Ferreira C.P., & Kraenkel R.A., 2011. The role of immunity and seasonality in cholera epidemics. Bulletin of mathematical biology, 73, 12, 2916-2931.

Sarvala J., Salonen K., Jarvinen M., Aro E., Huttula T., Kotilainen P., Kurki H., Langenberg V., Mannini P., Peltonen A., Plisnier P.-D., Vuorinen I., Molsa H. and Lindqvist O.V., 1999. Trophic structure of Lake Tanganyika: carbon flows in the pelagic food web. Hydrobiologia 407, 149–173.

Shapiro, R. L., Otieno, M. R., Adcock, P. M., Phillips-Howard, P. A., Hawley, W. A., Kumar, L., and Slutsker, L. 1999. Transmission of epidemic *Vibrio cholerae* O1 in rural western Kenya

associated with drinking water from Lake Victoria: an environmental reservoir for cholera ? The American journal of tropical medicine and hygiene, 60(2), 271-276.

Uthermöhl H. 1931. Neue Wege in der quantitativen Erfassung des Planktonts. Verhandlungen der Internationalen Vereinigung für Theoretische und Angewandte. Limnolpgie 5, 567-595.

Verburg P., Hecky R. E. and Kling H., 2003. Ecological consequences of a century of warming in Lake Tanganyika. Science 301, 5632, 505-507.

Walton C.C., Pichel W.G., Sapper J.F. and May D.A., 1998. The development and operational application of nonlinear algorithms for the measurement of sea surface temperatures with the NOAA polar-orbiting environmental satellites. Journal of Geophysical Research: Oceans 103, 27999-28012.

ACRONYMS

AATSR	Advanced Along-Track Scanning Radiometer
aer_opt	Aerosol option for atmospheric correction in SeaDAS
Africover	UN project of geographical information on Africa using satellites (land use, climate conditions, natural resources).
AP-HM	Assistance Publique-Hôpitaux de Marseille
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
AVHRR	Advanced Very High Resolution Radiometer
BELSPO	Belgian Science Policy
BT	Brightness temperature
Chl-a	Chlorophyll a
CHOLTIC	Cholera outbreaks at Lake Tanganyika induced by Climate Change?
CLIMFISH	Impact du changement climatique sur le développement soutenu des pêches au lac Tanganyika (BELSPO project)
CLIMLAKE	Climate variability as recorded in Lake Tanganyika (BELSPO project)
CRH	Centre de Recherches Hydrobiologiques (Uvira-D.R.Congo)
Crystal VC	Brand name of rapid diagnostic test for detection of *Vibrio cholerae*
DOF	Department of Fisheries (Zambia)
DRC	Democratic Republic of the Congo
EAGLES	East African Great Lake Ecosystem Sensitivity to changes (BELSPO project)
ENVISAT	ENVIronment SATellite from European Space Agency
ERDAS	Earth Resources Data Analysis System
GIS	Geographic Information system
Globcover	Project whose goal is to develop the sharpest map of Earth's global land cover.
HGR	Hôpital Général de Référence (Uvira)
HPLC	High Performance Liquid Chromatography
INRB	Institut National de la Recherche Biomédicale (Kinshasa)
ITM	Institute of Tropical Medicine (Antwerpen)
K490	Light attenuation coefficient in the water column at 490 nm
Landsat ETM+	Landsat Enhanced Thematic Mapper Plus
Landsat OLI	Landsat 8 Operational Land Imager
Landsat TM	Landsat Thematic Mapper
LSWT	Lake Surface Water Temperature
MERIS	MEdium Resolution Imaging Spectrometer on board ESA ENVISAT
MLVA	Multi-locus VNTR Analysis
MODIS	Moderate Resolution Imaging Radiometer on board Aqua and Terra satellites
MOH	Ministry of Health (Zambia)
MOU	Memorandum of Understanding
MUMM	Management Unit of the North Sea Mathematical Models
NDVI	Normalized Difference Vegetation index

NIR	Near infrared
NOAA.	National Oceanographic and Atmospheric Administration
OBPG	Ocean Biology Processing Group
OC	Ocean Color
PFGE	Pulsed-Field Gel Electrophoresis
RMCA	Royal Museum for Central Africa
Rrs	Remote sensing reflectance
RS	Remote Sensing
RTM	Radiative Transfer Model
SeaDAS	Software suite developed by NASA for the processing visualization, analysis and quality control of ocean color data
SST	Sea Surface Temperature
TS	Time Series
UCL	Université Catholique de Louvain
UFC	Université de Franche Comté (Besançon, France)
ULG	Université de Liège
USGS	United States Geological Survey
UTC	Coordinated Universal Time
VNTR	Cariable number of tandem repeats
WP	Work Package

ANNEX

Meeting notes and publications are available on request in separated pdf files