PHD COURSE IN LIFE AND ENVIRONMENTAL SCIENCES

Report Form for PhD student annual evaluation (XXXVI and XXXVII cycles)

Title of PhD research: Theoretical models to investigate microbial dynamics and interactions in the marine ecosystem under present and future global change scenarios......

Name of PhD supervi	sor: AntonioDell'Anno&Roberto Danovaro
Research lab name:	

Cycle: [X] XXXVI [] XXXVII

PhD Curriculum::

[X] Marine biology and ecology[] Biomolecular Sciences[] Civil and environmental protection

DISVA instrumentation labs/infrastructure eventually involved in the project:

- [] Actea Mobile Laboratory
- [] Advanced Instrumentation lab
- [] Aquarium
- [] MassSpec lab
- []MaSBiC
- [] Simulation/informatics lab
- [] Other. Please, indicate:

ABSTRACT (1000 characters, including spaces):

Viruses are key agents of prokaryotic mortality in the global oceans, and by killing their hosts they play an important role in the functioning of the marine food webs and C and nutrient (particularly N and P) cycling. Virus-induced mortality is particularly relevant in deep-sea ecosystems where nearly all the prokaryotic C production is transformed into organic detritus, thus representing an additional important trophic resource for the metabolism of noninfected microbes. Therefore, the integration of the viral component into trophodynamic and biogeochemical models is of primary importance for an improved understanding of the function of the world's oceans. Climate change is progressively altering marine ecosystems including the deep seas, but the impacts of these changes on virus-host dynamics are still largely unknown.

In my research I will investigate prokaryote-virus interactions, and their potential responses to global climate changes with a special focus on the deep seas. I will analyse these interactions using advanced mathematical models and machine learning techniques to identify the main drivers and patterns influencing viral and prokaryotic distribution in marine ecosystems, and simulate distribution changes under different scenarios. I will develop theoretical models that can describe microbial dynamics and interactions, and forecast the

effects of temperature shift and changes in food supply under different scenarios on deep-sea ecosystem functioning. With these models I will try to assess the quantitative relevance of bottom-up (resource availability) and top-down (predatory-pressure, e.g. viruses) control in shaping microbial diversity in different marine ecosystem settings and climate change scenarios.

Part 1. Scientific case of the PhD Research (2 to 3 pages, including figures)

- BACKGROUND

Viruses are by far the most abundant 'life forms' in the world's oceans, exceeding prokaryotic abundance by at least one order of magnitude (Suttle 2005). Increasing evidence indicates that viral infection may be responsible for the high mortality of autotrophic and heterotrophic organisms in surface oceans, with cascading effects on carbon cycling and nutrient regeneration (Suttle 2007). Viral lysis of infected microbes transforms their cell contents and biomass into organic detritus (both dissolved and particulate), which can then be used again by non-infected prokaryotes. This process, which is particularly relevant in benthic deepsea ecosystems worldwide, supports prokaryotic heterotrophic production, but it also decreases the efficiency of the carbon transfer to higher trophic levels and influences the carbon budget of the oceans (Furhman 1999; Danovaro et al., 2008). Therefore, the integration of the viral component into trophodynamic and biogeochemical models is of primary importance for an improved understanding of the function of the world's oceans. Viruses can also have a major impact on prokaryotic diversity (Suttle 2007; Paul 2008; Sime-Ngando 2014). Understanding if viruses can be a driver of marine biodiversity and related key ecological processes is a key issue for the comprehension of ecosystem functioning. In marine ecosystems, information on the relationships between microbial diversity and ecosystem functioning is practically absent (Danovaro et al. 2017), as well as factors influencing such relationships. Theoretical studies based on the assumption that each microbial "species" is specialised on the exploitation of different organic compounds argued that a higher microbial diversity can promote higher levels of ecosystem performance (Loreau 2001). So far, only a single study, carried out in coastal areas reported a positive relationship between prokaryotic diversity and ecosystem functioning (Danovaro & Pusceddu 2007). Identifying the mechanistic explanations of the relationships between biodiversity and ecosystem functioning is crucial for a better understanding of the potential consequences of climate changes on the provisioning of ecosystem's good and services for human wellbeing. The oceans are changing rapidly, and seawater warming has been documented for different marine regions and also for the deep seas (Levitus et al. 2005). Biogeochemical models predict a global decrease of the oceanic primary productivity (Steinacher et al. 2010), and hence a reduction of organic C inputs to the deep seafloor, although such an effect can be also opposite on a regional scale (Smith et al. 2013). As viral replication and life cycle are closely linked with host metabolism, increases in temperature and changes in organic C availability will likely influence the interactions between viruses and the organisms they infect. However, how host-virus interactions respond to altered temperature regimes is unknown.

- SCIENTIFIC AIMS

- 1. To identify patterns and drivers influencing viral and prokaryotic distribution in marine ecosystems.
- 2. To develop models that describe viral and prokaryotic global distribution, and simulate distribution changes under different environmental scenarios.
- 3. To investigate the potential effects of temperature and C availability on prokaryotic metabolism.
- 4. To investigate temporal changes of prokaryote-virus interactions in relation with changes of temperature and trophic availability.
- 5. To investigate the relevance of bottom-up (resource availability) and top-down (predatory-pressure, e.g. viruses) control in shaping microbial diversity under different marine ecosystem settings and climate change scenarios.
- 6. To develop theoretical/mechanistic models that describe microbial food-web dynamics and forecast their potential responses under future global change scenarios.

- WORKPLAN AND RESEARCH ACTIVITIES

WP 1. Objective.

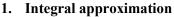
Identification of patterns and drivers of prokaryotes and viruses in marine systems and development of models to estimate their spatial distribution at a global scale and their responses under different global change scenarios.

Methods

- Metanalysis of the literature.
- Data collection
 - o data on prokaryotes and viruses from published papers;
 - temperature, pH, nutrients and primary production data from Copernicus databases, considering different ecosystems (sea water, freshwater, soil);
 - o carbon flux data from Oregon database.
- Database construction and Data Quality
 - Data Normalization, to make data coherent and comparable among all records and all sources;
 - Outlier Detection;
 - *import of all data cleaned in a single Database (Datawarehouse), ready for reporting and analytics.*
 - Exploratory Data Analysis:
 - o Collinearity Detection using Pearson Correlation Coefficient among all covariates;
 - *Multi-Collinearity Detection using Variance Inflation Factor among all covariates;*
 - Descriptive analysis, making different graphs (e.g., violin-plots) to analyse the microbial distribution in different environments and at different depths.
- Regression Analysis:
 - Linear Regressions (viral and prokaryotic abundances or production Vs each driver, also considering each ecosystem separately);
 - Generalized Linear Regressions (viral and prokaryotic abundances or production Vs all drivers, also considering each ecosystem separately);
 - *Model selection (stepAIC).*

Expected/Obtained Results.

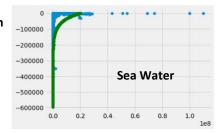
Depth



The first step of the stock analysis was the computation of different integrals which approximate the distribution of microbial (including prokaryotes and their viruses) abundances and productions in the water columns and in the sediments. Figure 1 indicates that prokaryotic abundance decreases exponentially with water column depth.

2. Global estimate

I estimated global microbial abundances and productions combining previous one-dimensional integrals with volume and surface approximations of oceans, sediments and different soil types (Figure 2).



Prokaryotic Abundance

Fig. 1. Scatterplot of prokaryotic abundances vs depths (in cm, sea water).

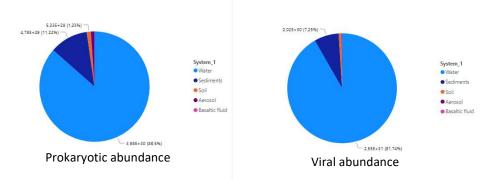


Fig. 2. Pie plot of prokaryotic and viral abundances in different environments

3. Violin/Box Plot

To investigate in more detail which are the systems hosting the higher prokaryotic and viral abundances I added geographical and environmental labels in the dataset to create violin plots (Figure 3).

4. Univariate analysis for pattern investigation

The second part of this WP was dedicated to the analysis of the relationships between microbial abundances (or productions) and environmental/trophic drivers. According to Wigington et al. (2016), I used a Log₁₀

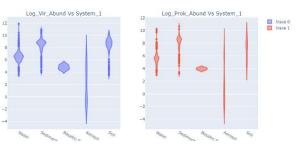


Fig. 3. Violin plots of prokaryotic and viral abundances in different environments (logarithmic scale).

transformation of the abundance/production data, obtaining, at the same time, a reduction of the biases due to outliers. Then I investigated univariate patterns, using linear regressions with only one driver and one response variable considered each time. In Figure 4 some of these relationships are shown.

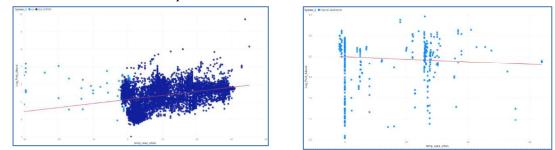


Fig. 4. Scatter plots with regression line for prokaryotic abundances (in logarithmic scale) with respect to temperature in marine water column (a) and sediment (b).

5. Correlation analysis

The presence of interesting univariate relationships among microbial data and some environmental drivers, guided me to conclude that a multivariate model could be meaningful, but considering separately the different systems investigated (Seawater, marine sediments, freshwater, and soil). The first step for each multivariate analysis is the correlation analysis, indeed, one of the main assumptions for a regression is the lack of perfect multicollinearity. To detect the presence of correlation we used the Pearson Correlation Coefficient calculated for each couple of drivers.

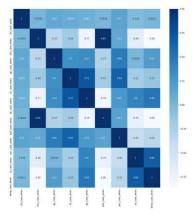


Fig. 5. Heatmap with Pearson Correlation Coefficients for drivers

MultiCollinearity Detenction		feature	VIF
The Pearson Coefficient measure the multicollinearity between couple	0	chl_case_when	1.038790
of variables (also called correlation), to investigate multicollinearity, i.e.		o2_case_when	33.010397
if one variable can be approximated as linear combinations of the others, I used another technique called VIF (Variance Inflation Factors). If the		pp_case_when	2.328068
VIF value is greater than 10, the corresponding variable should be		ph_case_when	3146.037544
dropped from the total model (Figure 6). As result of all these analyses,	4	poc_case_when	6.164204
I excluded, for sea water, NO ₃ and PO ₄ concentrations because of high		so_case_when	2995.398077
Pearson Correlation Coefficient with salinity, pH and O ₂ concentrations.		temp case when	10 790095

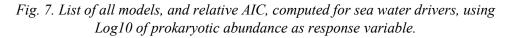
7. Model Selection

6. MultiCollinearity Detenction

With all the assumptions satisfied, I started to select the best

model explaining the distribution of microbial abundances and productions in relation to the drivers selected before. I used the stepAIC technique (Akaike Information Criterion), on different covariates' combinations to find the best model, and then I repeated the procedure on different types of regression (Linear Regression, Poisson Regression, Negative Binomial) with the aim to compare the results (Figure 7).

predictors	aic 🖵	r2 💌
['pp_case_when', 'chl_case_when', 'poc_case_when', 'temp_case_when']	1059,358	0,473677
['pp_case_when', 'poc_case_when', 'temp_case_when']	1082,321	0,469793
['chl_case_when', 'poc_case_when', 'temp_case_when']	1358,266	0,424898
['poc_case_when', 'temp_case_when']	1376,247	0,421503
['pp_case_when', 'chl_case_when', 'temp_case_when']	1637,606	0,375578
['pp_case_when', 'temp_case_when']	1638,735	0,375002
['pp_case_when', 'chl_case_when', 'poc_case_when']	1876,92	0,329974
['pp_case_when', 'poc_case_when']	1887,123	0,327561
['pp_case_when', 'chl_case_when']	2001,58	0,304504
['pp_case_when']	2004,016	0,303595
['chl_case_when', 'poc_case_when']	2641,894	0,160145
['poc_case_when']	2644,662	0,158964
['chl_case_when', 'temp_case_when']	2713 <i>,</i> 094	0,142345
['temp_case_when']	2714,225	0,141554
['chl_case_when']	3229,76	0,00078
0	3230,408	0



8. Next steps

The regression analysis has allowed highlighting the main drivers that explain the distribution of viruses and prokaryotes in all ecosystems of the biosphere considered, and the results obtained are

consistent with the studies carried out so far. Moreover, some of the main drivers are common to several systems (e.g. temperature). The last step of this work would be to use CMPI5 models with different RCP scenarios to forecast how the drivers identified before can be modified due to climate change. I will use these new drivers' values as inputs for the best models, found in the model selection, with the aim to assess changes in viral-host interactions over time. The idea is to obtain 3D maps that can describe changes in microbial dynamics in relation to climate-induced changes of environmental conditions (Figure 8).

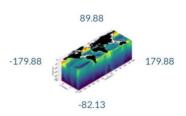


Fig. 8. 3D maps with microbial distributions.

Fig. 6. Variance Inflation Factors in sea water.

WP 2. Objective.

The aim of this WP is to verify whether the global temperature increase can modify virus-host interactions, potentially exacerbating the impact of viral infections and then altering C and nutrient cycling.

Methods.

- Metanalysis of the literature.
- Data collection
 - Data on prokaryotic metabolism and temperature from published papers.
 - Unpublished data on temperature and prokaryotic metabolism collected in different marine ecosystems.
 - Unpublished data on prokaryotic and viral abundance, prokaryotic metabolism and viral production collected over a decade at ca. 3000 m depth in the Western Mediterranean Sea
- Database construction and Data Quality
 - Data Normalization, to make data coherent and comparable among all records and all sources.
 - *Outlier Detection.*
 - Import of all data cleaned in a single Database, ready for reporting and analytics.
 - Regression Analysis:
 - Linear regression analysis between prokaryotic metabolism and temperature in different deep-sea ecosystems (Arrhenius equation).
 - Comparison of the regression outputs.
- Differential Analysis:
 - Development of a mechanistic model to describe microbial dynamics in the deep sea, using Generalized Lotka-Volterra models.
 - Development of a temperature-dependent mechanistic model to describe microbial dynamics in the deep sea, starting from Generalized Lotka-Volterra models.

Expected/Obtained Results.

1. GLV Time-dependent

As first step of this WP, I started from an analysis of the bibliography and then I adapted the existing Generalized Lotka-Volterra models to our system, adding new terms to the differential equations, especially for the part describing the carbon dynamics. Below the theoretical form of the model developed:

Then I estimated the model parameters (i.e., coefficients) from our dataset containing information on prokaryotic and viral abundances and metabolism. Finally, I ran the mechanistic models, one for each sampling year (Figures 9 and 10), to compare the real data with the computational ones. Indeed, the dynamics in prey-predator systems is periodic and oscillatory, but our measures are static, then

the typical statistical techniques, as the regressions, are not enough powerful to catch the relationships virus-host changing in time and strictly related to the resource (i.e., carbon) contribution.

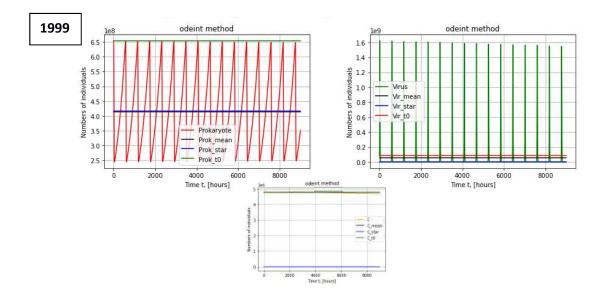


Fig. 9. Computational solutions and steady states of the dynamical system (1) calculated for the year 1999, compared to real data, considering prokaryotic and viral abundances, and carbon as resource.

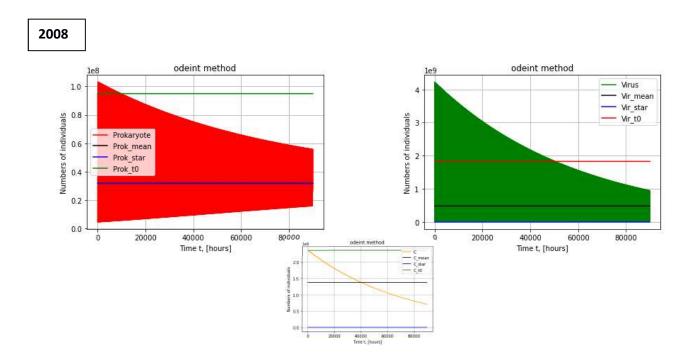


Fig. 10. Computational solutions and steady states of the dynamical system (1) calculated for the year 2008, compared to real data, considering prokaryotic and viral abundances, and Carbon as resource.

As it can observed from the graph, the system velocity increases, as highlighted by the period of the oscillations that decreases because the metabolic parameters of the model increase, but this behaviour is bound to a decreasing macro trend that follows the carbon dynamics.

2. Regression analysis of metabolic parameters

Because the temperature seems to have an important role in the different scenarios computed for each sampling year, we tried to develop a temperature dependent model (usual GLV is time dependent). Then as first step I tried to express all the parameters of the GLVs model with respect to temperature, using Linear or Log-Linear models, depending on the distributions of the points (Figure 11).

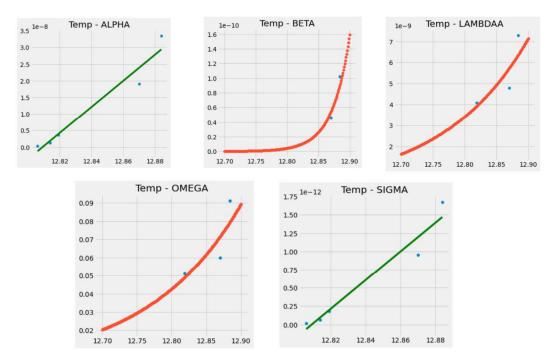


Fig. 11. Each graph represents a Regression (Linear or Log-Linear) of a parameter of the GLV model (1) with respect to temperature.

3. Steady States Analysis

As second step we studied the dynamics of the GLV steady states (i.e., solutions static in time) in relation to temperature, indeed each point of the scatterplots below is a solution of the model (1) computed fixing a temperature value and solving the system with the derivative in time equal to zero (Figure 12).

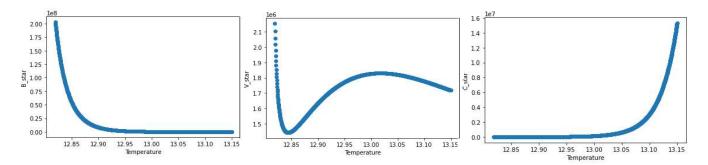


Fig. 12. Each graph represents the steady state for each function of the system (1), B* for Prokaryotes, V* for viruses, C* for carbon, all these steady states are calculated for different fixed temperature values.

4. GLV Temperature-dependent

Finally, we combine all the previous information to create a GLV temperature-dependent model, where the derivative are now with respect to temperature and no longer with respect to time (Figure 13).

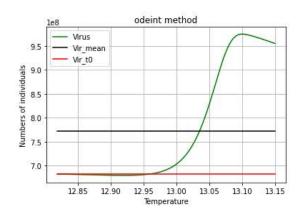


Fig. 13. This scatterplot represents the dynamics of the computational viral abundance calculated from the GLV temperature-dependent.

5. Next steps.

The mechanistic model, parametric with respect to temperature, has highlighted the importance of this latter in explaining the dynamics between viruses and prokaryotes, because it influences the metabolic parameters that govern the microbial dynamics. The last step of my work would be to use CMPI5 models with different RCP scenarios to forecast temperature changes due to climate change in the deep-sea ecosystem investigated (ca. 3000 m depth in the Western Mediterranean Sea). I will use these parameter values as the input to our new model to study its dynamics and simulate the variations in viral and prokaryotic abundances due to temperature changes.

WP 3. Objective.

To investigate the relevance of bottom-up (resource availability) and top-down (predatory-pressure, e.g. viruses) controls in shaping microbial diversity in different marine ecosystem settings and climate change scenarios.

Methods

- Data collection:
 - Published and unpublished data on trophic availability, viral infections and prokaryotic diversity in different deep-sea marine ecosystems
 - Published data and data in specific repository (e.g. Copernicus) of different environmental variables (e.g., temperature, salinity, oxygen concentrations)
- Database construction and Data Quality
 - Data Normalization, to make data coherent and comparable among all records and all sources.
 - Outlier Detection.
 - *Import of all data cleaned in a single Database, ready for reporting and analytics.*
- Regression Analysis:
 - Linear Regressions, Generalized Linear Regressions, Generalized Linear Mixed Models and Random Forests to evaluate which abiotic and biotic factors or

their interactions can be relevant to influence microbial biodiversity in the deep sea.

- Differential Analysis:
 - Development of a mechanistic model to test whether viral-induced mortality can explain prokaryotic diversity using Generalized Lotka-Volterra models.

Expected/Obtained Results.

Improving the understanding of factors influencing microbial diversity in the deep sea and testing the killing the winner hypothesis (Winter et al., 2010).

- REFERENCES

- 1. Bratbak, G., Thingstad, T. F. (1985). Phytoplankton-bacteria interactions: an apparent paradox? Analysis of a model system with both competition and commensalism. Mar. Ecol. Prog. Ser. 25
- 2. Danovaro R, Pusceddu A. (2007). Biodiversity and Ecosystem Functioning in Coastal Lagoons: Does Microbial diversity Play Any Role? Est Coast Shelf Sci 75:4-12
- 3. Danovaro R, Dell'Anno A, et al. (2008). Major viral impact on the functioning of benthic deep-sea ecosystems. Nature 454:1084-1087
- 4. Danovaro R, Corinaldesi C, Dell'Anno A, Snelgrove PVR. (2017). The deep-sea under global change. Curr Biol 27:R431–R510
- 5. Danovaro R, Dell'Anno A, et al. (2016). Virus-mediated archaeal hecatomb in the deep seafloor. Sci Adv 2:e1600492
- 6. Dell'Anno A, Danovaro R. (2005). Extracellular DNA plays a key role in deep-sea ecosystem functioning. Science 309:2179
- 7. Dell'Anno A, et al. (2015). Virus decomposition provides an important contribution to benthic deepsea ecosystem functioning. Proc Nat Acad Sci USA 112:E2014-E2019
- 8. Kirchman, David L., Xosé Anxelu G. Morán, and Hugh Ducklow. (2009). "Microbial growth in the polar oceans—role of temperature and potential impact of climate change." Nature Reviews Microbiology 7: 451-459.
- 9. Levitus S, et al. (2005). Warming of the world ocean, 1955–2003. Geophys Res Lett 32:L02604
- 10. Lønborg, Christian, et al. (2016) "Depth dependent relationships between temperature and ocean heterotrophic prokaryotic production." Frontiers in Marine Science 3: 90.
- 11. López-Urrutia, Ángel, and Xosé Anxelu G. Morán. (2007). "Resource limitation of bacterial production distorts the temperature dependence of oceanic carbon cycling." Ecology 88: 817-822.
- 12. Loreau M. (2001). Microbial diversity, producer-decomposer interactions and ecosystem processes: a theoretical model. Proc Biol Sci 268:303–309
- 13. McClain, Craig R., et al. (2012) "Energetics of life on the deep seafloor." Proceedings of the National Academy of Sciences 109: 15366-15371.
- 14. Mora C., Wei C.-L. Rollo, Amaro T., Baco A.R. et al. (2013). Biotic and Human Vulnerability to Projected Changes in Ocean Biochemistry over the 21st Century. PLos Biol. 11(10)
- 15. Morán, Xosé Anxelu G., et al. (2017) "Temperature regulation of marine heterotrophic prokaryotes increases latitudinally as a breach between bottom-up and top-down controls." Global change biology 23: 3956-3964.
- Suttle CA. (2007). Marine viruses—major players in the global ecosystem. Nature Rev Microbiol 5:801-812
- 17. Thingstad, T. F., Pengerud, B. (1985). Fate and effect of allochthonous organic material in aquatic microbial eco- systems. An analysis based on chemostat theory. Mar. Ecol. Prog. Ser. 21
- 18. Thingstad T.F., Sakshaug E. (1990). Control of phytoplankton growth in nutrient recycling ecosystems. Theory and terminology. Mar. Ecol. Prog. Ser. 63:261-272.

- 19. Thingstad, T. F. (1997). Accumulation of degradable DOC in surface waters: Is it caused by malfunctioning microbial loop? Limnol. Oceanogr. 42
- 20. Thingstad T.F., Lignell R. (1997). Theoretical models for the control of bacterial growth and abundance, diversity and carbon demand. Aquat Microb Ecol. 13:19–27.
- 21. Thingstad T.F (1998). A theoretical approach to structuring mechanism in the pelagic food web. Hydrobiologia 363:59-72.
- 22. Thingstad T.F. (2000). Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. Limnol Oceanogr. 45(6)
- 23. Tittensor D.P., Mora C., Jetz W., Lotze H.K., Ricard D., Vanden Berger E., Worm B. (2010) Global patterns and predictors of marine biodiversity across taxa. Nature Vol 466.
- 24. Thurber AR, et al. (2014). Ecosystem function and services provided by the deep sea. Biogeosciences 11:3941–3963
- 25. S. Våge, G. Bratbak, J. Egge, M. Heldal, A. Larsen, S. Norland, M. L. Paulsen, B. Pree, R. A. Sandaa, E. F. Skjoldal, T. M. Tsagaraki, L. Øvreås, T. F. Thingstad (2018). Simple models combining competition, defence and resource availability have broad implications in pelagic microbial food webs. Ecology Letters 21.
- 26. Wei C.-L., Rowe G.T., Escobar-Briones E., Boetius A., Soltwedel T., et al. (2010). Global Patterns and Predictions of Seafloor Biomass Using Random Forests. PLos ONE 5(12).
- 27. Weitz J. S., Stock C. A., Wilhelm S. W., Bourouiba L., Coleman M.L., Buchan A., Follows M.J. et al. (2015). A multitrophic model to quantify the effects of marine viruses on microbial food webs and ecosystem processes. The ISME Journal 9, 1352-1364
- 28. Wigington C.H., Sonderegger D., Brussaard C.P.D., Buchan A., Finke J.F. et al. (2016). Reexamination of the relationship between marine virus an microbial cell abundances. Nature Microbiology.
- 29. Winter C., Bouvier T., Weinbauer M.G., Thingstad T.F. (2010). Trade-Offs between Competition and Defense Specialists among Unicellular Planktonic Organisms: the "Killing the Winner" Hypothesis Revisited. Microbiology and Molecular Biology Reviews, p 42-57
- 30. Yinon M. Bar-On, Phillips R., Milo R. (2018). The biomass distribution on Earth. PNAS, 19.

Part 2. PhD student information on the overall year activity (courses/seminars/schools, mobility periods, participation to conferences)

List of attended courses/seminars/schools

- 1. "Hands-on Python for Machine Learning" with Dr. Marco Zappatore Adjunct Professor of CAT Laboratory, University of Salento (seminar)
- 2. LABORATORI TEORICO-PRATICI SULLA STRUMENTAZIONE AVANZATA: Laboratory of oceanographic instruments, Prof Pierpaolo Falco, dr. Francesco Memmola, dr. Alessandro Colucelli, Lab of Oceanography (course)
- 3. LEZIONI FRONTALI: Rischio climatico Prof. Pierpaolo Falco, Lab of Oceanography (course)
- 4. LEZIONI DI INFORMATICA: Getting Started with R: Environmental Computing Dr. Giuseppe d'Errico, Lab Ecotoxicology and Environmental Chemistry. (course)

List of periods spent abroad

List of conferences/workshops attended and of contributions eventually presented

Part 3. PhD student information on publications

If not yet published, please indicate the publication status (submitted, accepted, in preparation...)

List of publications on international journals

List of publications on conference proceedings List of other publications (books, book chapters, patents)

[14/10/2022]

Student signature Alemir Giorgetti

Supervisor signature

٨