LIMNOLOGY and OCEANOGRAPHY: METHODS

ASLO

Limnol. Oceanogr.: Methods 21, 2023, 734–760

doi: 10.1002/lom3.10578

Studying bioluminescence flashes with the ANTARES deep-sea neutrino telescope

Nico Reeb [1,2]* Sebastian Hutschenreuter,[1,3] Philipp Zehetner,[1,4,5] Torsten Ensslin,[1,4] A. Albert,[6,7] S. Alves,[8] M. André,[9] M. Anghinolfi,[10] G. Anton,[11] M. Ardid,[12] J.-J. Aubert,[13] J. Aublin,[14] B. Baret,[14] S. Basa,[15] B. Belhorma,[16] M. Bendahman,[14,17] V. Bertin,[13] S. Biagi,[18] M. Bissinger,[11] J. Boumaaza,[17] M. Bouta,[19] M. C. Bouwhuis,[20] H. Brânzaş,[21] R. Bruijn,[20,22] J. Brunner,[13] J. Busto,[13] B. Caiffi,[10] A. Capone,[23,24] L. Caramete,[21] J. Carr,[13] V. Carretero,[8] S. Celli,[23,24] M. Chabab,[25] T. N. Chau,[14] R. Cherkaoui El Moursli,[17] T. Chiarusi,[26] M. Circella,[27] A. Coleiro,[14] M. Colomer-Molla,[8,14] R. Coniglione,[18] P. Coyle,[13] A. Creusot,[14] A. F. Díaz,[28] G. de Wasseige,[14] A. Deschamps,[29] C. Distefano,[18] I. Di Palma,[23,24] A. Domi,[10,30] C. Donzaud,[14,31] D. Dornic,[13] D. Drouhin,[6,7] T. Eberl,[11] T. van Eeden,[20] N. El Khayati,[17] A. Enzenhöfer,[13] P. Fermani,[23,24] G. Ferrara,[18] F. Filippini,[26,32] L. Fusco,[13] Y. Gatelet,[14] P. Gay,[14,33] H. Glotin,[34] R. Gozzini,[11] R. Gracia Ruiz,[20] K. Graf,[11] C. Guidi,[10,30] S. Hallmann,[11] H. van Haren,[35] A. J. Heijboer,[20] Y. Hello,[29] J. J. Hernández-Rey,[8] J. Hößl,[11] J. Hofestädt,[11] F. Huang,[6] G. Illuminati,[14,26,32] C. W. James,[36] B. Jisse-Jung,[20] M. de Jong,[20,37] P. de Jong,[20] M. Jongen,[20] M. Kadler,[38] O. Kalekin,[11] U. Katz,[11] N. R. Khan-Chowdhury,[8] A. Kouchner,[14] I. Kreykenbohm,[39] V. Kulikovskiy,[10] R. Lahmann,[11] R. Le Breton,[14] D. Lefèvre,[40,41] E. Leonora,[42] G. Levi,[26,32] M. Lincetto,[13] D. Lopez-Coto,[43] S. Loucatos,[14,44] L. Maderer,[14] J. Manczak,[8] M. Marcelin,[15] A. Margiotta,[26,32] A. Marinelli,[45] J. A. Martínez-Mora,[12] K. Melis,[20,22] P. Migliozzi,[45] A. Moussa,[19] R. Muller,[20] L. Nauta,[20] S. Navas,[43] E. Nezri,[15] B. Ó Fearraigh,[20] M. Organokov,[6] G. E. Păvălaş,[21] C. Pellegrino,[26,46,47] M. Perrin-Terrin,[13] P. Piattelli,[18] C. Pieterse,[8] C. Poirè,[12] V. Popa,[21] T. Pradier,[6] N. Randazzo,[42] S. Reck,[11] G. Riccobene,[18] A. Romanov,[10,30] A. Sánchez-Losa,[8,27] F. Salesa Greus,[8] D. F. E. Samtleben,[20,37] M. Sanguineti,[10,30] P. Sapienza,[18] J. Schnabel,[11] J. Schumann,[11] F. Schüssler,[44] M. Spurio,[26,32] Th. Stolarczyk,[44] M. Taiuti,[10,30] Y. Tayalati,[17] S.J. Tingay,[36] B. Vallage,[14,44] V. Van Elewyck,[14,48] F. Versari,[14,26,32] S. Viola,[18] D. Vivolo,[45,49] J. Wilms,[39] S. Zavatarelli,[10] A. Zegarelli,[23,24] J. D. Zornoza,[8] and J. Zúñiga,[8] (ANTARES Collaboration)

[1]Information Field Theory Group, Max Planck Institute for Astrophysics, Garching, Germany
[2]TUM School of Computation, Information and Technology, Technical University of Munich, Garching, Germany
[3]Department of Astrophysics/IMAPP, Radboud University Nijmegen, Nijmegen, The Netherlands
[4]Department of Physics, Ludwig-Maximilians University, Munich, Germany
[5]CERN, Geneva, Switzerland
[6]Université de Strasbourg, CNRS, IPHC UMR 7178, Strasbourg, France
[7]Université de Haute Alsace, Mulhouse, France
[8]IFIC—Instituto de Física Corpuscular (CSIC—Universitat de València), Paterna, Valencia, Spain
[9]Technical University of Catalonia, Laboratory of Applied Bioacoustics, Barcelona, Spain
[10]INFN—Sezione di Genova, Genova, Italy

*Correspondence: nreeb@mpa-garching.mpg.de

Author Contribution Statement: T.E. and S.H. devised the main conceptual ideas. N.R. developed the method and performed the computations. T.E., P.Z., and S.H. verified the analytical methods. N.R. wrote the manuscript with support of all authors. T.E. supervised the findings of this work. T.E. advised about the ANTARES experiment and data. The ANTARES Collaboration provided the data.
S. Hutschenreuter, P. Zehetner, T. Ensslin, and T. Eberl equally contributed to this study.

Additional Supporting Information may be found in the online version of this article.

S. Hutschenreuter, P. Zehetner, T. Ensslin, and T. Eberl equally contributed to this study.

[11]Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen Centre for Astroparticle Physics, Erlangen, Germany
[12]Institut d'Investigació per a la Gestió Integrada de les Zones Costaneres (IGIC)—Universitat Politècnica de València, Gandia, Spain
[13]Aix Marseille Univ., CNRS/IN2P3, CPPM, Marseille, France
[14]Université de Paris, CNRS, Astroparticule et Cosmologie, Paris, France
[15]Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France
[16]National Center for Energy Sciences and Nuclear Techniques, Rabat, Morocco
[17]Faculty of Sciences, University Mohammed V in Rabat, Rabat, Morocco
[18]INFN—Laboratori Nazionali del Sud (LNS), Catania, Italy
[19]Laboratory of Physics of Matter and Radiations, University Mohammed I, Oujda, Morocco
[20]Nikhef, Science Park, Amsterdam, The Netherlands
[21]Institute of Space Science, Măgurele, Romania
[22]Universiteit van Amsterdam, Instituut voor Hoge-Energie Fysica, Amsterdam, The Netherlands
[23]INFN—Sezione di Roma, Roma, Italy
[24]Dipartimento di Fisica, Università La Sapienza, Roma, Italy
[25]LPHEA, Faculty of Science—Semlali, Cadi Ayyad University, Marrakech, Morocco
[26]INFN—Sezione di Bologna, Bologna, Italy
[27]INFN—Sezione di Bari, Bari, Italy
[28]Department of Computer Architecture and Technology/CITIC, University of Granada, Granada, Spain
[29]Géoazur, UCA, CNRS, IRD, Observatoire de la Côte d'Azur, Sophia Antipolis, France
[30]Dipartimento di Fisica dell'Università, Genova, Italy
[31]Université Paris-Sud, Orsay Cedex, France
[32]Dipartimento di Fisica e Astronomia dell'Università, Bologna, Italy
[33]Laboratoire de Physique Corpusculaire, Clermont Université, Université Blaise Pascal, CNRS/IN2P3, Clermont-Ferrand, France
[34]LIS, UMR Université de Toulon, Aix Marseille Université, CNRS, Toulon, France
[35]Royal Netherlands Institute for Sea Research (NIOZ), Texel, The Netherlands
[36]International Centre for Radio Astronomy Research, Curtin University, Bentley, Western Australia, Australia
[37]Huygens-Kamerlingh Onnes Laboratorium, Universiteit Leiden, Leiden, The Netherlands
[38]Institut für Theoretische Physik und Astrophysik, Universität Würzburg, Würzburg, Germany
[39]Dr. Remeis-Sternwarte and ECAP, Friedrich-Alexander-Universität Erlangen-Nürnberg, Bamberg, Germany
[40]Mediterranean Institute of Oceanography (MIO), Aix-Marseille University, Marseille Cedex 9, France
[41]Université du Sud Toulon-Var, CNRS-INSU/IRD UM 110, La Garde Cedex, France
[42]INFN—Sezione di Catania, Catania, Italy
[43]Departamento de Física Teórica y del Cosmos and C.A.F.P.E, University of Granada, Granada, Spain
[44]IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette, France
[45]INFN—Sezione di Napoli, Napoli, Italy
[46]Museo Storico della Fisica e Centro Studi e Ricerche Enrico Fermi, Rome, Italy
[47]INFN, CNAF, Bologna, Italy
[48]Institut Universitaire de France, Paris, France
[49]Dipartimento di Fisica dell'Università Federico II di Napoli, Napoli, Italy

Abstract

We develop a novel technique to exploit the extensive data sets provided by underwater neutrino telescopes to gain information on bioluminescence in the deep sea. The passive nature of the telescopes gives us the unique opportunity to infer information on bioluminescent organisms without actively interfering with them. We propose a statistical method that allows us to reconstruct the light emission of individual organisms, as well as their location and movement. A mathematical model is built to describe the measurement process of underwater neutrino telescopes and the signal generation of the biological organisms. The Metric Gaussian Variational Inference algorithm is used to reconstruct the model parameters using photon counts recorded by photomultiplier tubes. We apply this method to synthetic data sets and data collected by the ANTARES neutrino telescope. The telescope is located 40 km off the French coast and fixed to the sea floor at a depth of 2475 m. The runs with synthetic data reveal that we can model the emitted bioluminescent flashes of the organisms. Furthermore, we find that the spatial resolution of the localization of light sources highly depends on the configuration of the telescope. Precise measurements of the efficiencies of the detectors and the attenuation length of the water are crucial to reconstruct the light emission. Finally, the application to ANTARES data reveals the first localizations of bioluminescent organisms using neutrino telescope data.

The deep sea is one of the remotest habitats on Earth and its biological diversity is largely unexplored. Seventy-six percent of the inhabitants in the pelagic ecosystem emit light to communicate, attract prey or to protect themselves (Martini and Haddock 2017). The trait of bioluminescence is distributed over a diverse range of marine species, from bacteria to fish (Widder 2010). Over the last years, the distribution and quantification of bioluminescence in the deep sea and individual luminescent organisms have been studied using a variety of observational techniques (Martini and Haddock 2017; Herren et al. 2005; Martini et al. 2014; Frank et al. 2012; Martini et al. 2019). Most in situ observation techniques rely on actively triggering the light production by disturbing the environment and stimulating the organisms, since spontaneous emission does not occur at statistically sufficient rates for observation times in the order of hours (Priede et al. 2006; Widder et al. 1989; Craig et al. 2015). In this context, spontaneous emission refers to light emission which is not intentionally stimulated by observers, similar as in (Widder et al. 1989). The free-fall lander observations in the Atlantic Ocean off Cape Verde have detected five events per hour when fixing the sensor 250 m above the sea floor at 4700 m leading to no active stimulation due to movements of the sensor (Priede et al. 2006). With these studies, similarities between emitted bioluminescence flashes have been observed. Most eukaryotic organisms emit a single light flash or a series of flashes; a rapid increase of the luminosity indicates the start of a flash which—after reaching its peak value—is decaying exponentially with a time constant significantly longer than that of the initial rise. Studies have characterized such flash light curves from various species by the duration, the maximum photon flux, and the total photon emission of the flash (Esaias et al. 1973; Galt and Sykes 1983; Morin 1983; Lapota and Losee 1984; Galt et al. 1985; Herring 1988; Batchelder and Swift 1989; Lapota et al. 1989; Buskey and Swift 1990; Batchelder et al. 1992; Buskey 1992; Herring et al. 1993; Swift et al. 1995; Latz and Jeong 1996; Latz et al. 2004; Widder 2010; Lapota 2012; Valiadi and Iglesias-Rodriguez 2013; Johnsen et al. 2014; Craig et al. 2015; Cronin et al. 2016; Messié et al. 2019). An overview of these characteristics is provided in Table 1. Most species in the benthic and pelagic zone emit light flashes with their emission maxima within the range of $\lambda = 450 - 520$ nm (Widder 2010) which corresponds to the wavelength window of maximum light transmission in seawater (Aguilar et al. 2005).

Table 1. Characteristics of bioluminescent flashes taken from Widder (2010), Craig et al. (2015), and Messié et al. (2019).

Characteristics	Values
Mean duration of flash (s)	$0.1 - 59.0$
Maximum photon flux (photons s^{-1})	$4.9 \cdot 10^7 - 6.4 \cdot 10^{12}$
Total light emission (photons $flash^{-1}$)	$2.4 \cdot 10^8 - 2.3 \cdot 10^{13}$
Spectral wavelength (nm)	$450 - 520$

For years, deep-sea neutrino telescopes have been monitoring and recording the photon flux. These telescopes aim to detect Cherenkov radiation caused by charged secondary particles, which are induced by high-energy cosmic neutrino interactions with constituents of water molecules. Besides the use for research in fundamental physics, neutrino telescopes are also part of long-term monitoring strategies of the deep ocean including its biodiversity and indicators for climate change (Danovaro et al. 2017, 2020; Aguzzi et al. 2019). The records of these telescopes were used to analyze the dynamics of deep-sea bioluminescence (van Haren et al. 2011; Tamburini et al. 2013; Martini et al. 2014; Aguzzi et al. 2017). The majority of the recorded bioluminescence is assumed to be triggered by sea currents and turbulence around the detectors (Tamburini et al. 2013; Martini et al. 2014; Aguzzi et al. 2017; Chatzievangelou et al. 2021). In addition to analysis of long-term temporal changes, the experimental setup, and long-term observations of deep-sea neutrino telescopes offer also the possibility to analyze light emission of individual organisms. Although the observations of bioluminescence with these telescopes are conducted less invasively than the previously mentioned methods, the systematic impact of the telescope on the environment cannot be neglected when studying naturally occurring bioluminescence (Meighen-Berger et al. 2021).

In this paper, we present a method to reconstruct the movement and characteristic light curves of individual biological sources in the deep sea with data of a neutrino telescope. In particular, we use both synthetic as well as measured data of the ANTARES neutrino telescope located 40 km off the French coast on the bottom of the Mediterranean Sea (42°48′N, 6°10′E) and anchored to the sea floor at a depth of 2475 m. The method enabled us to do the localization of a luminescent organism using ANTARES data and the simultaneous reconstruction of the corresponding emitted bioluminescence light curve.

To do so, we developed a generative model of the measurement process of a neutrino detector and the signal generation of the biological sources to understand the origin of the measured data. The model parameters such as source movement and characteristic light curves are reconstructed by a Variational Bayesian Inference algorithm, called Metric Gaussian Variational Inference (MGVI) (Knollmüller and Enßlin 2020). The Numerical Information Field Theory (NIFTy) framework (Reinecke et al. 2013; Arras et al. 2019) provides an implementation of the MGVI algorithm and has been used to obtain results of this work. The code to our work is publicly available.[1]

Our method offers the opportunity to study the trait of bioluminescence of organisms over more than 10 yr.

[1] https://gitlab.mpcdf.mpg.de/ift/public/lumintracinggitlab.mpcdf.mpg.de/ift/public/lumintracing

The localization of individual organisms is needed to infer the absolute emitted luminosity, and hence necessary for a taxonomy of luminescent deep-sea organisms. Furthermore, the localization of organisms helps to understand the cause of the light emission better, for example, light emissions caused by physical contact with the telescope, by turbulences from the telescope or naturally occurring bioluminescence in undisturbed regions. This method allows us to gather information about the distribution of bioluminescence over time and space within the volume of the telescope.

The paper is structured as follows; in "Data" section, we summarize the data provided by the ANTARES experiment and used for the reconstruction. In "Methods" section, the proposed method is explained by describing the generative model in detail and highlighting our assumptions. In "Simulations" section, simulations and reconstructions on synthetic data sets are performed to discuss the opportunities and limitations of the model. The first localizations of deep-sea organisms and reconstructions of their emitted bioluminescence light curves using the ANTARES detector are presented and discussed in "Data analysis" section.

Data

The ANTARES telescope consists of 12 lines that are distributed over an area of 0.1 km^2. The lines have a length of 480 m and are placed at a distance of around 60 m to each other. Each line, excluding the 12th, contains 25 storeys with a vertical separation of 14.5 m between them. The 5 top storeys of line 12 are not equipped with optical modules, but with different acoustic instruments (Adrián-Martínez et al. 2012*a*). The first storey of each line is located around 100 m above the seabed. A storey is defined as a collection of three optical modules (OMs) each containing a photomultiplier tube (PMT). The OMs are oriented downward looking under 45° (a zenith angle of 135°) and with an angle of 120° to each other in horizontal directions (Amram et al. 2002; Ageron et al. 2011). A schematic view of the ANTARES setup is given in Fig. 1.

Each optical module measures the light activity in terms of photon hits. The photon hits recorded in time frames of $\Delta t = 104.858$ ms are directly sent to shore and the rates of each time frame are calculated (Aguilar et al. 2007). According to a trigger system, the photon hits are stored for a specific period of time depending on the type of trigger. In the

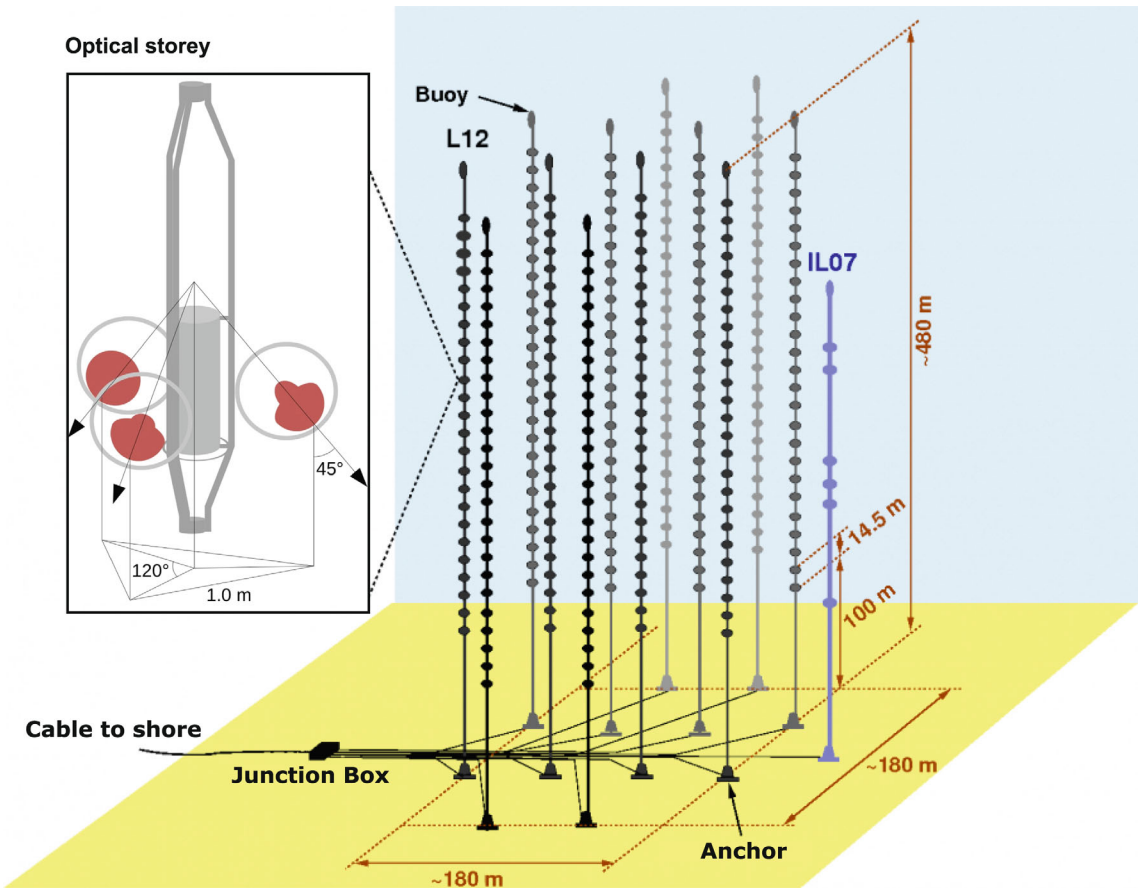

Fig. 1. Schematic of the ANTARES setup including an illustration of a storey with its three optical modules orientated downwards. The instrumentation line IL07 is colored blue (Ageron et al. 2011).

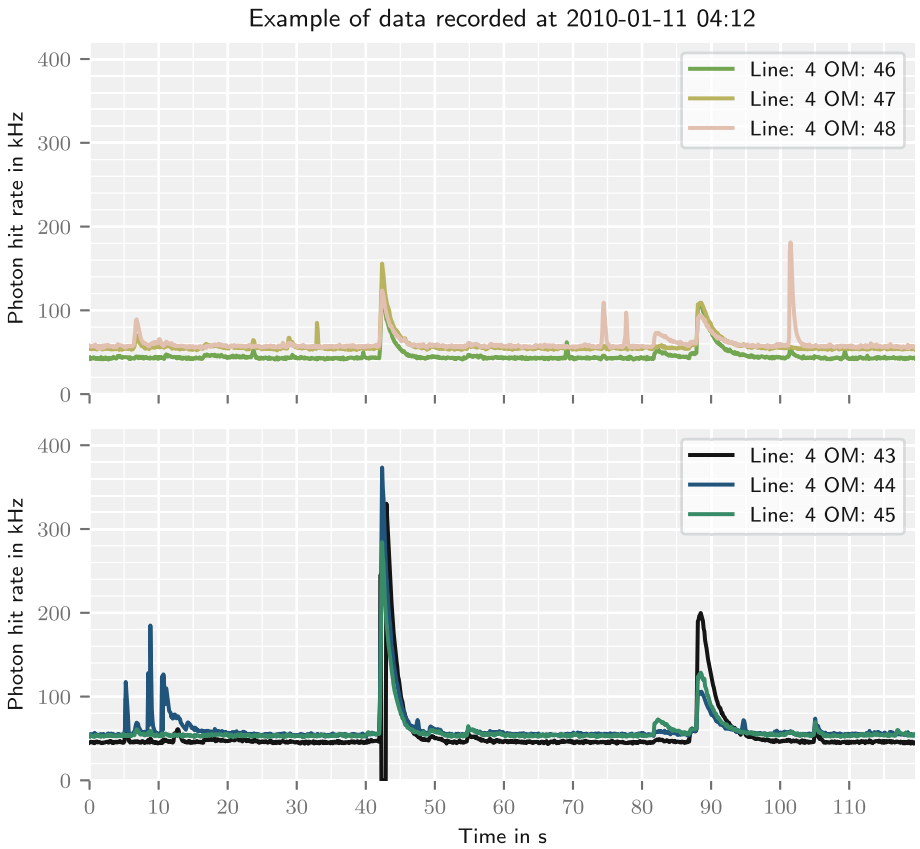

Fig. 2. Recorded photon hits of six OMs. The upper plot shows the photon hits of the optical modules in one storey. The lower plot shows the photon hits of the optical modules in the storey below.

following, only data samples with detection periods covering fully the bioluminescence light curve of interest are used to analyze the biological behavior. About 1–2 of such recordings are saved per day containing around 2 min of continuous raw data.

In addition to the photon count rates, the position and orientation, as well as the efficiency of each optical module is monitored precisely (Amram et al. 2002; Albert et al. 2018). The origin of the internal coordinate system used for the positioning of the OMs is located in the center of the detector volume. This internal coordinate system of ANTARES indicates the west–east direction as x-axis and the south–north direction as y-axis. The vertical direction is given by the z-axis (Amram et al. 2002, 2003). Within this work, we introduce the naming convention of each optical module as a tuple of line number and number of optical modules (l, n_{om}). The optical modules of one line are consecutively numbered starting at the bottom of the line.

Table 2. Measurements of the light attenuation length l_{att} in sea water at ANTARES site including only statistical errors (Aguilar et al. 2005; Adrián-Martínez et al. 2012*b*).

Time of measurement	l_{att} (m)
July 1998	60.6 ± 0.4
March 1999	51.9 ± 0.7
June 2000	46.4 ± 1.9
May 2008 to March 2010	$\sim 50 - 60$

An example data file is shown in Fig. 2. In order to represent the setup of the ANTARES detector, the photon counts of one storey are grouped together. In the figure, an almost constant background for each optical module with photon count rates around $40 - 60$ kHz can be identified. This background is assumed to be induced by ^{40}K nuclear decays, bioluminescence, photomultiplier intrinsic noise, and radioactive decays in the sea water and in the glass sphere (Amram et al. 2000). Furthermore, we assume to recognize two bioluminescence flashes that are recorded by six optical modules over two storeys. The occurrence times of the flashes are 40 and 84 s after a trigger started the recording. The flash recorded at $t = 40$ s surpassed the threshold limit of optical module $(4, 43)$, that is, the readout electronics is saturated which let the recorded photon rate drop to zero for this detector for a short period. These two flashes are analyzed in "Data analysis" section.

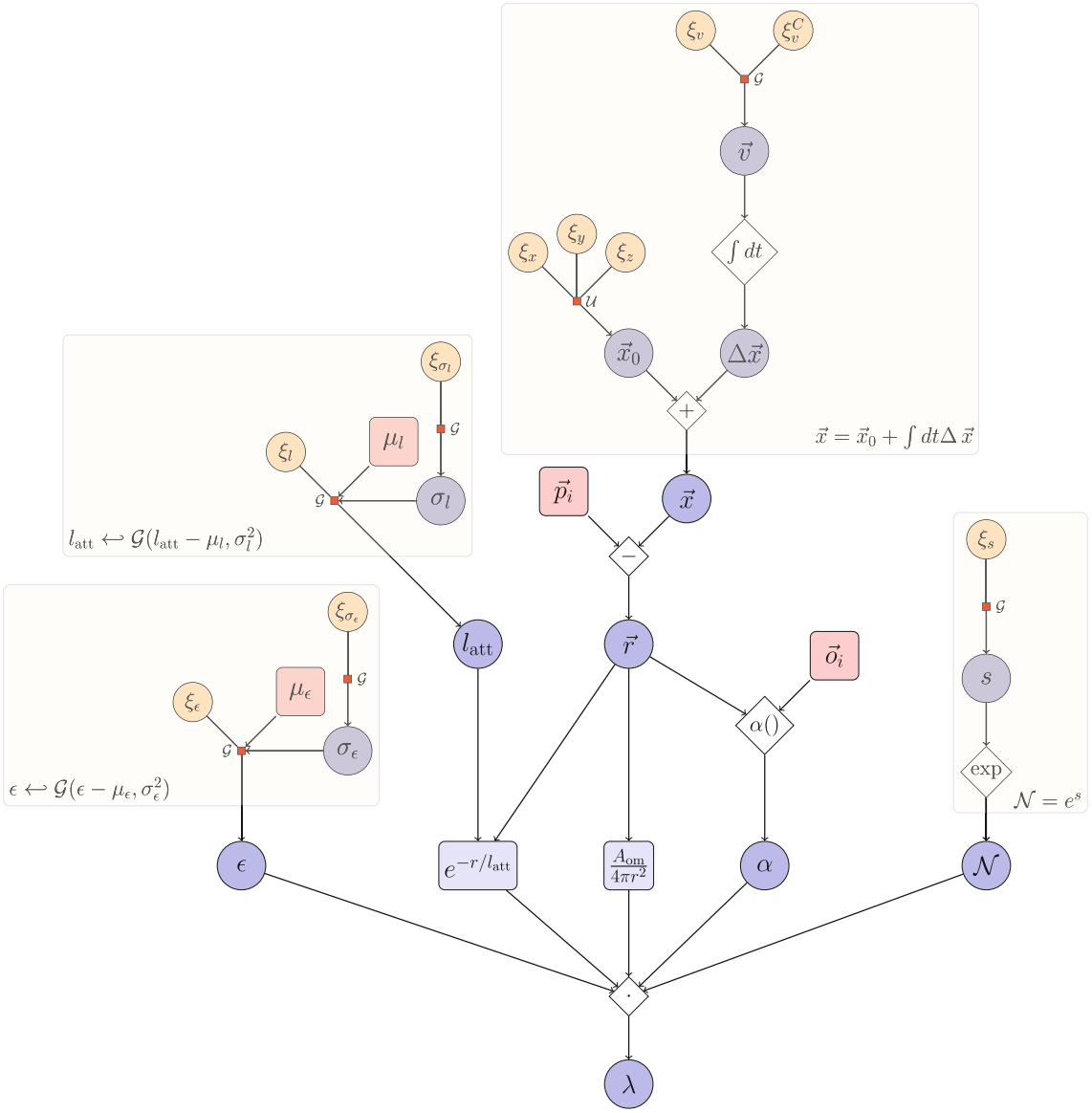

Fig. 3. Generative model of the response function. Blue shapes indicate in quantities that are in principle observable. The standardized variables are colored orange and operations have a diamond shape. The transformations from a standard Gaussian distribution into a target distribution are labeled as a small red square with the target distribution next to it. All red colored values include a priori assumptions, that is, distribution transformation, detector position and orientation.

The attenuation length of light in the sea water at the ANTARES site was measured by the ANTARES Collaboration. The light attenuation of water depends on its chemical and physical properties. Therefore, various measurements of the water light transmission properties have been made from 1997 to 2010 using an isotropic light source that emits blue light (Aguilar et al. 2005; Adrián-Martínez et al. 2012*b*). The light attenuation includes the effect of water absorption of light as well as the impact of scattered photons reaching the detector (Aguilar et al. 2005). A summary of the results of the study is given in Table 2. Variation in light attenuation depends on the amount of particles in the water that depends on oceanographic processes. A detailed model of the environment is crucial to perform reasonable reconstructions of biological light sources.

Methods

The data provided by the ANTARES experiment and knowledge about the organisms and their environment in the deep sea allow us to derive a mathematical model of the data generation process. This model depends on a set of

model parameters ξ and is able to describe a luminescent organism emitting a flash light curve, the propagation of the signal in the deep sea, and the detection of photon hits at a neutrino detector.

According to Bayes' theorem, measuring photon hits d updates the prior knowledge on ξ, expressed as probability density function $\mathcal{P}(\xi)$. The resulting posterior density $\mathcal{P}(\xi|d)$ can be calculated as follows by knowing the likelihood $\mathcal{P}(d|\xi)$ of the obtained data, given ξ, and the marginal probability of the data $\mathcal{P}(d)$,

$$\mathcal{P}(\xi|d) = \frac{\mathcal{P}(d|\xi)\mathcal{P}(\xi)}{\mathcal{P}(d)} \tag{1}$$

In case of high dimensionality and complexity of a model, the posterior distribution $\mathcal{P}(\xi|d)$ is often intractable. Our model, which will be explained in the following and is shown in Fig. 3, has up to 1000 dimensions depending on the final implementation and the size of the analyzed data set. For a static light source, the number of dimensions can be reduced to around 200. To overcome this issue of tractability we approximate this distribution with a simpler distribution $\widetilde{\mathcal{P}}_\eta(\xi)$ depending on variational parameter η. Within this work the posterior approximation is performed by the MGVI algorithm (Knollmüller and Enßlin 2018, 2020).

Table 3. Overview of the response function with its parameters and their explanations.

$$\lambda_i(\boldsymbol{t}) = \varepsilon_i \cdot \mathcal{N}(\boldsymbol{t}) \cdot \alpha\left(\vec{o}_i, \vec{r}_i(\boldsymbol{t})\right) \cdot e^{-(r_i(\boldsymbol{t})/l_{\text{att}})} \frac{A_{\text{om}}}{4\pi \cdot r_i^2(\boldsymbol{t})}$$

High-level parameters	Explanation
ε_i	Detector efficiency of optical module i, assumed constant over time
$\mathcal{N}(t)$	Emitted photons of an isotropic point source emitter, time dependent
$\vec{x}(t)$	Position of biological object, time dependent
$\vec{p}_i$	Position of optical module i, assumed constant over time
$\vec{r}_i(t) = \vec{p}_i - \vec{x}(t)$	Vector from source to optical module i
$r_i(t) = \lvert \vec{r}_i(t) \rvert$	Distance from source to optical module i
$\vec{o}_i$	Orientation of optical module i
$\alpha\left(\vec{o}_i, \vec{r}_i(t)\right)$	Angular acceptance
l_{att}	Attenuation length of light in sea water, assumed constant over time
$A_{\text{om}} = \pi r_{\text{om}}^2$	Effective area of the optical module assuming a circle
$A_{\text{light}} = 4\pi \cdot r_i^2(t)$	Area covered by spherical radiation at the location of optical module i

In order to apply Bayes' update rule and perform the posterior approximation a detailed understanding of the likelihood is crucial. First, we discuss the measurement process of a photon detector and derive an expression for the likelihood of multiple optical modules. Second, we build a mathematical model for the light emission of an organism and the light propagation through water reaching the photon detector. We explain each aspect of our model in detail, beginning with the generic formula of the expected photon number arriving at one optical module.

Likelihood and measurement process

The optical modules of neutrino telescopes detect single photons. Individual photon hits can be treated as independent events and therefore the photon detection is a classical Poisson process. Due to the assumption of Poisson statistics, the photon rate $r_{i,t}$ over the detection window $(t-\Delta t, t)$ needs to be converted to the total number of photon hits $d_{i,t} = \Delta t \cdot r_{i,t}$ detected by optical module i.

The likelihood of measuring $d_{i,t}$ photon hits at the optical module i at time t for a given expected number of photon hits $\lambda_{i,t}$ can be written as

$$\mathcal{P}(d_{i,t}|\lambda_{i,t}) = \lambda_{i,t}^{d_{i,t}} \frac{e^{-\lambda_{i,t}}}{d_{i,t}!} \tag{2}$$

The expected number of photon hits $\lambda_{i,t}$ is defined as the photon counts over the fixed detection time Δt. The measurement process is independent at different times $t, t+\Delta t, \ldots$ and at different optical modules $i, i+1, \cdots$. Hence, the likelihood of the count data vector d over a time frame $\Delta T = N \cdot \Delta t$ with N discrete time steps and elements $d_{i,t}$ is the direct product of the single measurement likelihoods,

$$\mathcal{P}(d|\lambda) = \prod_i \prod_{t=0}^{N-1} \lambda_{i,t}^{d_{i,t}} \frac{e^{-\lambda_{i,t}}}{d_{i,t}!} = \prod_i \mathcal{P}(d_i, \lambda_i) \tag{3}$$

Signal generation

Expected photon number

The number of expected photons depends highly on the light source itself, but also on specific attributes of the detector and its surroundings. The luminescent organisms are modeled as moving point sources with a position $\vec{x}(t)$ that generate specific time-varying light patterns spreading isotropically. The assumption of isotropy is an approximation to keep the model complexity feasible for the inference, especially regarding the computational cost and degeneracies. Recordings of such light patterns have been made by Mazzei et al. (2014) and show traces of anisotropic emission patterns, that is, the emitted photon number depends on the direction of emission. Despite this simplification, we believe that the reconstructions will be reasonable. Most of the bioluminescence flashes are

only recorded by a single line, hence only a limited part of the real anisotropy is detectable. As we will demonstrate in the following, the simple isotropic model performs well without any traces of residual anisotropy visible in, for example, the efficiency or the noise offset. In line with Occam's razor, we therefore opt for the simplest model that can explain the data.

The total amount of photons emitted by a point source over the detection period $(t-\Delta t, t)$ is described by the function $\mathcal{N}(t)$. Typically, $\mathcal{N}$ is a vector containing the photon numbers of a time frame ΔT at N discrete times with time steps of Δt. The luminosity $\mathscr{L}$ can be calculated as the rate of emitted photons per second, $\mathscr{L} = \mathcal{N}/\Delta t$. Therefore, the number of emitted photons $\mathcal{N}$ is given by multiplying the luminosity with the time step Δt.

Due to various factors such as detector quality, detector geometry and water absorption, the amount of photons reaching the optical modules is reduced. The efficiency ε_i of the optical module i is sampled ($\hookleftarrow$) from a Gaussian model

$$\varepsilon \hookleftarrow \mathcal{G}\left(\varepsilon - \varepsilon_{\text{ANTARES}}, \sigma_\varepsilon^2\right) \tag{4}$$

with efficiencies $\varepsilon_{\text{ANTARES}}$ provided by measurements of the ANTARES Collaboration and inferable standard deviation σ_ε (*see* Supporting Information). Deviations need to be allowed because missing optical modules lead to wrong assumptions about the efficiencies in one storey (*see* Albert et al. 2018).

In addition, the photon sensitive area of the optical modules is modeled as a circular surface with an area of $A_{\text{om}} = \pi r_{\text{om}}^2$. Only a fraction

$$\frac{A_{\text{om}}}{A_{\text{light}}(r_i(t))} \tag{5}$$

of the emitted photons can hit the detector i since the photons spread uniformly on the surface of the sphere $A_{\text{light}}\left(\vec{r}\right) = 4\pi \cdot r_i^2(t)$ at a distance $r_i(t) = |\,\vec{r}_i(t)\,|$ from the source. The vector between the light source position $\vec{x}$ and the detector position $\vec{p}_i$ is calculated by

$$\vec{r}_i(t) = \vec{p}_i - \vec{x}(t).$$

Table 4. Overview of the emitted photon number function with its parameters, their explanations and origin.

$$\mathcal{N}(\boldsymbol{t}) = \boldsymbol{e}^{\boldsymbol{s}(\boldsymbol{t})}$$

Parameters	Explanation	Origin
s	Correlated field	$s \hookleftarrow \mathcal{G}(s-\mu_s, C)$
μ_s	Inferable mean	See Supporting Information
C	Correlation matrix	Fixed prior value

Furthermore, the angular acceptance and accordingly the orientation of the detector play an important role. The angular acceptance $\alpha(\theta_{\text{optical}})$ as a function of the optical angle θ_{optical} is provided by the ANTARES experiment. The cosine of the optical angle depends on the orientation $\vec{o}_i$ of the optical module and can be obtained by

$$\cos\theta_{\text{optical}} = \frac{-1}{|\,\vec{r}_i(t)\|\vec{o}_i\,|}\,\vec{r}_i(t)\cdot\vec{o}_i$$

The polynomial fit of the angular acceptance stated as

$$\alpha(\theta_{\text{optical}}) = \alpha\left(\vec{r}_i(t), \vec{o}_i\right) \tag{6}$$

is used to calculate the percentage of photons coming from direction $\vec{r}_i(t)$ that actually hit the optical module oriented in $\vec{o}_i$ direction.

Finally, the impact of electromagnetic absorption by water and the photon scattering can be determined by the attenuation length l_{att} and the Beer–Lambert law. The fraction of photons reaching the detector after absorption and scattering can be calculated by

$$\frac{\widetilde{\mathcal{N}}_i(t)}{\mathcal{N}(t)} = e^{-r_i(t)/l_{\text{att}}} \tag{7}$$

with $\mathcal{N}(t)$ as defined before being the number of emitted photons and $\widetilde{\mathcal{N}}_i(t)$ the number of photons reaching the distance of the detector i. Despite small local changes of the water properties in the deep sea, the attenuation length l_{att} is assumed to be independent of the position of the optical modules. A loss of photons can still be explained either by a small attenuation length or by a darker light source. We aim to break this degeneracy between the attenuation length and the emitted number of photons via an informative prior on the former quantity. However, the measurements of the attenuation length mentioned in "Data" section (Table 2) show a high variability over the years. Therefore, we assume a Gaussian distribution as prior for the attenuation length with a mean $\mu_{l_{\text{att}}} = 50$ m and an inferable (*see* Supporting Information) wide standard deviation $\sigma_{l_{\text{att}}} = 20$ m. The implications of the degeneracy are discussed in more detail in the simulations discussed in "Static reconstruction of a synthetic static source" section.

Combining these effects on the emitted photons, the response function for the photon counts $\lambda_{i,t}$ detected by the optical module i can be expressed as

$$\lambda_i(t) = \varepsilon_i \cdot \mathcal{N}(t) \cdot \alpha\left(\vec{o}_i, \vec{r}_i(t)\right) \cdot e^{(-(r_i(t)/l_{\text{att}}))} \frac{A_{\text{om}}}{4\pi \cdot r_i^2(t)} \tag{8}$$

It is important to highlight that $\mathcal{N}$ describes the number of photons isotropically emitted by a hypothetical luminescent

Table 5. Overview of the position and movement model with its parameters, their explanations, and origin.

$$\vec{x}(t) = \vec{x}_0 + \Delta t \cdot \sum_{t_k=t_0}^{t} \vec{v}(t_k)$$

Parameters	Explanation	Origin
j_0	Starting position	$j_0 \curvearrowleft \mathcal{G}\left(j_0 - \mu_{j_0}, \sigma_{j_0}^2\right)$
	$j \in x, y, z$ independent	
μ_{j_0}	Mean position	Fixed prior values
σ_{j_0}	Standard deviation (inferable)	See Supporting Information
$\vec{v}$	Velocity vector	$v_j \curvearrowleft \mathcal{G}\left(v_j - \mu_{v_j}, K\right)$
	v_x, v_y, v_z independent	
μ_{v_j}	Inferable mean	See Arras et al. (2020)
K	Inferable Correlation matrix	A priori assumptions

organism modeled as point source. The complex structure of real biological organisms that may lead to anisotropic emission is not covered by our model. This mismatch between the assumed isotropic model and the real light emission of organisms in the deep sea can lead to unrepresentative uncertainty estimates of the MGVI algorithm, as discussed in "Data" section. A compact summary of the response function with its high-level parameters is given in Table 3. A visualization of the described generative model is given in Fig. 3 at the end of the section.

Luminosity

The bioluminescence light curves are assumed to be the dominant feature of the photon counts over time. The luminosity model $\mathcal{L}(t) = \mathcal{N}(t)/\Delta t$ has to be able to capture all features of a bioluminescence flash and hence be able to provide sensible prior samples.

Since the number of emitted photons is always positive, $\mathcal{N}(t)$ can be described to sufficient accuracy by a log-normal model. The light curve structure does not depend on absolute times t, but on the relative timing $\Delta t = t_l - t_k$. Therefore, a correlated signal $s(t)$ with a given correlation matrix $C_{t_l t_k} = C(t_l - t_k)$ is used to model the burst kinetics under the assumption of statistical homogeneity. Combining all prior assumptions, the equation for the luminosity model yields

$$\mathcal{N}(t) = e^{s(t)} \tag{9}$$

with $s(t)$ sampled from a Gaussian with inferable mean μ_s and fixed covariance C,

$$s(t) \curvearrowleft \mathcal{G}(s - \mu_s, C) \tag{10}$$

Although we do not know the underlying correlation structure exactly, we can use the recorded light curve structure to estimate the correlation. The main motivation for a fixed covariance is to reduce the computation time. Alternatively, the correlation could be inferred as well, as it is done for the source movement in "Source movement" section. The formulation and discussion of a reasonable correlation function as well as the distribution transformations used for the luminosity model can be found in the Supporting Information. The parameters of the luminosity model were chosen such that the attributes of recorded bursts given in Table 1 were fulfilled and similar bursts as in Fig. 2 could be constructed. A compact summary of the function modeling the photon rate with its parameters is given in Table 4 and illustrated in Fig. 3.

Source movement

The x, y, and z directions of the movement of the source within the coordinate system of the detector are handled independently from each other. For each direction $j \in x, y, z$ a velocity vector $v_j(t)$ can be reconstructed to describe the movement starting at point $\vec{x}_0 = (x_0, y_0, z_0)$.

Similarly to the luminosity model, statistical stationarity is assumed for the velocity as a function of time. But instead of using a fixed correlation matrix, this is inferred as well. In contrast to the luminosity model, we do not have access to previous recordings of the movement to estimate the correlation. Consequently, the velocity vectors

$$v_j(t) \curvearrowleft \mathcal{G}\left(v_j - \mu_{v_j}, K\right) \tag{11}$$

are sampled from a Gaussian with an inferable correlation matrix $K_{t_l t_k}$ as covariance. The covariance can be set such that sampled velocities meet criteria of biological plausible movements. A detailed discussion about reconstructing correlation functions of a signal can be found in the NIFTy documentation (Reinecke et al. 2013) and the corresponding papers (Arras et al. 2019, 2020).

The starting position of the movement $\vec{x}_0$ is assumed to be drawn from Gaussian prior distributions with their respective means (μ_{x_0}, μ_{y_0}, μ_{z_0}), which are set to be directly at the line below the storey with the highest photon count during a bioluminescence flash and inferable standard deviations $\sigma_{\{x,y,z\}}$ (*see* Supporting Information).

$$j_0 \curvearrowleft \mathcal{G}\left(j_0 - \mu_{j_0}, \sigma_{j_0}^2\right) \tag{12}$$

The absolute position $\vec{x}(t)$ at time t can be obtained by integrating the velocity from start time t_0 up to t and adding the start position $\vec{x}_0$. Hence, the expression of the position vector $\vec{x}(t)$ can be derived as

$$\vec{x}(t) = \vec{x}_0 + \Delta t \cdot \sum_{t_k=t_0}^{t} \vec{v}(t_k) \tag{13}$$

Table 6. GT and reconstructed parameters for a static inference source. The synthetic photon count data were generated from the simulated static source. The reconstructed attenuation length is also included.

Run	x (m)	y (m)	z (m)	$\mathscr{L}_{\max}$ (GHz)	l_{att} (m)
GT	50.00	1.00	23.00	45.59	45.00
0	50.03 ± 0.04	1.16 ± 0.03	22.94 ± 0.04	44.62	45.97 ± 1.32
1	50.02 ± 0.05	1.10 ± 0.02	22.97 ± 0.04	45.72	43.95 ± 0.36
2	49.98 ± 0.06	1.13 ± 0.03	22.95 ± 0.04	45.45	44.41 ± 1.02
4	50.05 ± 0.04	1.18 ± 0.07	22.99 ± 0.04	44.34	46.68 ± 1.65
5	50.03 ± 0.05	1.12 ± 0.04	22.99 ± 0.03	45.41	44.50 ± 0.59
6	50.00 ± 0.07	1.13 ± 0.03	22.99 ± 0.02	45.45	44.37 ± 0.62
7	50.08 ± 0.25	1.12 ± 0.03	22.96 ± 0.04	45.62	44.05 ± 0.56
8	49.99 ± 0.03	1.15 ± 0.04	22.97 ± 0.03	44.98	45.15 ± 0.60
9	50.02 ± 0.07	1.08 ± 0.09	22.87 ± 0.21	45.07	45.10 ± 0.47

Table 7. Model parameters of a static inference source. The model is used to reconstruct the photon count data generated from a simulated static source.

Observable	Model	Model parameters
$\vec{x}_0$ position	Gaussian	$\vec{\mu}_{\vec{x}_0} = (46\text{ m}, 10\text{ m}, 15\text{ m})$ $\mu_{\sigma_{l_0}} = 10\text{ m} \quad \forall j \in x, y, z$ $\sigma_{\sigma_{l_0}} = 1\text{ m} \quad \forall j \in x, y, z$
Flash shape	Correlated signal	See Supporting Information

Distribution transformations between Gaussian distributions and uniform distributions $x_0(\xi_{x_0})$ are used and discussed in the Supporting Information. A compact summary of the position and movement model with its parameters is given in Fig. 3 and in Table 5.

Optical background

In addition to the light curves emitted by an individual luminescent organism, photons from other sources are detected as well. As mentioned in "Data" section, an almost constant background light is assumed to be induced by nuclear decays and bacterial bioluminescence. We model this optical background as constant offset n_i for each individual optical module i. Each optical module monitors a different water volume containing a different amount of nuclear decays and of small luminescent organisms. Allowing such local variations of the water composition, each background value n_i for optical module i is sampled from a Gaussian distribution with inferable standard deviation,

$$n_i \hookleftarrow \mathcal{G}(n_i - \mu_n, \sigma_n^2) \tag{14}$$

This leads to the extended response function of an optical module i

$$\lambda_{i,\text{ext}} = \lambda_i + \varepsilon_i \cdot n_i \tag{15}$$

Table 8. Model parameters used to reconstruct the attributes of a simulated telescope and its surroundings.

Observable	Model	Model parameter
ε	$\mathcal{G}(\varepsilon - \mu_\varepsilon, \sigma_\varepsilon^2)$ with $\mathcal{G}\left(\sigma_\varepsilon - \mu_{\sigma_\varepsilon}, \sigma_{\sigma_\varepsilon}^2\right)$	$\mu_\varepsilon = \varepsilon_{\text{ANTARES}}$ $\mu_{\sigma_\varepsilon} = 0.001$ $\sigma_{\sigma_\varepsilon} = 0.001$
l_{att}	$\mathcal{G}\left(l_{\text{att}} - \mu_{l_{\text{att}}}, \sigma_{l_{\text{att}}}^2\right)$ with $\mathcal{G}\left(\sigma_{l_{\text{att}}} - \mu_{\sigma_{l_{\text{att}}}}, \sigma_{\sigma_{l_{\text{att}}}}^2\right)$	$\mu_{l_{\text{att}}} = 50\text{ m}$ $\mu_{\sigma_{l_{\text{att}}}} = 20\text{ m}$ $\sigma_{\sigma_{l_{\text{att}}}} = 5\text{ m}$
n_i	$\mathcal{G}(n_i - \mu_n, \sigma_n^2)$ with $\mathcal{G}\left(\sigma_n - \mu_{\sigma_n}, \sigma_{\sigma_n}^2\right)$	$\mu_n = 60\text{ kHz}$ $\mu_{\sigma_n} = 5\text{ kHz}$ $\sigma_{\sigma_n} = 1\text{ kHz}$

Reconstruction routine

For the inference of such a complex model we follow a fixed reconstruction routine for each data set. Most of the described parameters have the same prior values for all data sets, for example, the same assumptions about the luminosity and hence need no adjustments. However, the prior values of the starting position change for different data sets and have to be set manually (*see* "Source movement" section). Furthermore, a mask is defined to exclude uninformative data points, that is, dead times of the PMTs (*see* "Data" section), low-efficiency PMTs ($d_i < 5$ kHz) and additional flashes. A mask is represented by a vector consisting of binary values 0 and 1 indicating informative and uninformative data points, respectively.

Additionally, the reconstruction is split up into two parts. We first assume a simpler model by neglecting the movement of the source. This assumption reduces the complexity of the model and allows for a more stable inference. After analyzing the results of this static reconstruction, one can conclude whether a more complex model including the movement of

the source is required. This is done by inspecting the residuals between model and data, that is, by evaluating the posterior predictive power of the model. If the data can be reconstructed by the assumption of a static source, the real source does not move or its movement cannot be resolved. If this is not the case, the reconstructed position, luminosity, and efficiencies are used as initial samples for the dynamic reconstruction. The samples of remaining variables of the model are randomly initiated according to their model priors. This scheme reduces the risk of overfitting to which such a complex model is prone to. In the next section, the splitting of the reconstruction is performed on simulated data to analyze this routine, as well as to illustrate the limitations and capabilities of the model.

Simulations

In the last section, we developed the response function of the expected photon counts of an optical module. Due to the form of the response, being a product of the factors $\mathcal{N}$, ε, $e^{r/l_{\text{att}}}$, and $(A_{\text{om}}/4\pi r^2)\cdot\alpha\left(\vec{r},\vec{o}\right)$, increases of one factor can be compensated by decrease of another factor, and vice versa. This leads to a degeneracy between the emitted photon number $\mathcal{N}$, the source position $\vec{x}$, the efficiency ε of an optical module and the attenuation length l_{att}. This degeneracy can only be reduced by using the data of multiple OMs and assuming a constant background light. In the following sections, we analyze the degeneracy between those variables by performing the reconstruction on synthetic data sets of multiple OMs. First, we focus on a static reconstruction of a simulated static source in order to examine whether the degeneracy can be reduced to such a degree that a light source can be localized and the corresponding bioluminescence light curve reconstructed. Second, we apply the complete dynamic reconstruction to the simulated static source, as well as a simulated dynamic source. These results are used to discuss the spatial resolution of the reconstruction and the possibility of a movement reconstruction.

For the simulation of static and dynamic sources, the configuration of the ANTARES telescope from some past moment in time is used to create the detector setup. The efficiencies are randomly drawn from Gaussian distributions (Eq. 4) with fixed standard deviation for each optical module. To simulate a realistic environment, the attenuation length is set to $l_{\text{att}} = 45$ m and the constant light background of each optical module is drawn from a Gaussian distribution (Eq. 14) with fixed mean $\mu_n = 60$ kHz and fixed standard deviation $\sigma_n = 2$ kHz. Furthermore, the shape of a burst is extracted from real data and scaled to reach realistic luminosity values.

The initial position in both cases is manually set to ($x_0 = 50$ m, $y_0 = 1$ m, $z_0 = 23$ m). While the static source stays at its initial position, we simulate a linear movement with a velocity $v = 0.2$ m s^{-1} for the dynamic source. The light signal is detected by 27 optical modules distributed over 9 storeys of one line. The mean of the recorded photon hits of each optical module is calculated according to Eq. 15. Tables 6 and 11 provide the ground truth (GT) observables of the static and dynamic simulated source, respectively.

Static reconstruction of a synthetic static source

The parameters for the initial position of the static inference model (Eq. 12) used to reconstruct data generated from the static source are provided in Table 7. The parameters of the photon number model are given in the Supporting Information. Assumptions about the efficiencies of the optical modules, the attenuation length and the light background are given in Table 8. The position and orientation of each optical module used for the simulation are also used for the inference.

Several reconstruction runs were performed, each with a different random seed to assess numerical stability. In order to analyze the degeneracy between source position, luminosity, and attenuation length, we compare the reconstructed values with their GT summarized in Table 6. For most runs, we were able to recover the GT values with acceptable precision. The position of the light source $\vec{x} = (50\text{ m}, 1\text{ m}, 23\text{ m})$ has been reconstructed with a maximum deviation to the GT of $\Delta y_{\max} = 18$ cm in y direction for run 4, whereas a standard deviation of $\sigma_y = 7$ cm was inferred. For all runs, a similar shift for the y position can be identified. We argue that the detector setup, especially the orientation of the PMTs, leads to systematic errors and regions with variant spatial resolutions. Furthermore, the degeneracy between the attenuation length and luminosity becomes apparent in the reconstruction. A higher attenuation length will lead to a lower maximum luminosity (e.g., runs 0 and 4) and vice versa (e.g., run 1). We assume that our algorithm is able to infer the attenuation length for a perfect scenario, that is, an isotropic light source and only constant background sources. For more realistic scenarios, this might not be the case. Run 3 did not converge due to

Table 9. Model parameters of a dynamic inference source. The model is used to reconstruct the photon count data generated from a simulated static source.

Observable	Model	Model parameters
v_i velocity	Correlated signal with $i \in \{x,y,z\}$	$\mu_v = 0$ m s^{-1}
		$\sigma_v = 0.1$ m s^{-1}
		$\sigma_v = 0.1$ m s^{-1}
$\vec{x}_0$ position	Gaussian	$\vec{\mu}_{\vec{x}_0} = (46\text{ m}, 10\text{ m}, 15\text{ m})$
		$\mu_{\sigma_{l_0}} = 20$ m $\quad \forall j \in x,y,z$
		$\sigma_{\sigma_{l_0}} = 5$ m $\quad \forall j \in x,y,z$
Flash shape	Correlated signal	See Supporting Information

Table 10. GT and reconstructed parameters using a dynamic source model. The synthetic photon count data were generated from a simulated static source. The positions are given at during the time of highest photon emission t_h and at the end of the data set (t_{max}) to reflect the movement of the source.

Run	x_{t_h} (m)	$x_{t_{max}}$ (m)	y_{t_h} (m)	$y_{t_{max}}$ (m)	z_{t_h} (m)	$z_{t_{max}}$ (m)	$\mathscr{L}_{max}$ (GHz)	l_{att} (m)
GT	50.00	50.00	1.00	1.00	23.00	23.00	45.59	45.00
0	49.99 ± 0.06	49.95 ± 0.09	1.09 ± 0.09	0.83 ± 0.20	22.99 ± 0.06	22.80 ± 0.15	45.74	43.98 ± 0.65
1	50.07 ± 0.13	50.09 ± 0.19	1.12 ± 0.13	1.08 ± 0.18	22.99 ± 0.05	22.96 ± 0.09	45.39	44.58 ± 0.58
2	50.04 ± 0.13	50.16 ± 0.21	1.12 ± 0.08	1.08 ± 0.15	23.01 ± 0.09	22.99 ± 0.14	45.44	44.46 ± 0.85
4	50.07 ± 0.10	50.13 ± 0.18	1.16 ± 0.08	0.96 ± 0.16	22.93 ± 0.06	22.89 ± 0.11	45.63	43.80 ± 0.59
5	50.04 ± 0.07	49.80 ± 0.15	1.05 ± 0.09	0.93 ± 0.21	23.01 ± 0.06	22.93 ± 0.10	45.56	44.23 ± 0.61
6	50.04 ± 0.16	50.09 ± 0.20	1.19 ± 0.08	1.33 ± 0.17	23.00 ± 0.09	22.98 ± 0.11	44.79	45.76 ± 0.79
7	49.93 ± 0.48	49.79 ± 0.50	1.21 ± 0.09	1.03 ± 0.21	22.93 ± 0.06	22.90 ± 0.08	44.86	44.95 ± 0.64
8	49.98 ± 0.07	49.86 ± 0.14	1.09 ± 0.06	1.06 ± 0.11	23.00 ± 0.08	22.92 ± 0.14	45.36	44.32 ± 0.87
9	50.03 ± 0.08	49.84 ± 0.17	1.00 ± 0.08	1.02 ± 0.13	22.98 ± 0.15	22.83 ± 0.20	45.59	44.26 ± 0.88

Table 11. GT and reconstructed parameters using a static source model. The synthetic photon count data were generated from a simulated dynamic source. The GT positions are given during the time of highest photon emission t_h and at the end of the data set (t_{max}) to reflect the movement of the source. The summary includes the reconstructed attenuation length.

Run	x_{t_h} (m)	$x_{t_{max}}$ (m)	y_{t_h} (m)	$y_{t_{max}}$ (m)	z_{t_h} (m)	$z_{t_{max}}$ (m)	$\mathscr{L}_{max}$ (GHz)	l_{att} (m)	$\overline{e^2}$
GT	51.15	52.10	2.15	3.10	24.15	25.10	45.71	45.00	1.02
0	51.19 ± 0.03		2.32 ± 0.02		24.23 ± 0.04		44.24	46.75 ± 0.64	1.32
1	51.17 ± 0.04		2.27 ± 0.02		24.23 ± 0.05		44.91	45.08 ± 0.23	1.36
2	51.12 ± 0.05		2.33 ± 0.06		24.20 ± 0.06		44.10	46.81 ± 1.74	1.35
4	51.22 ± 0.02		2.32 ± 0.06		24.29 ± 0.02		44.72	45.84 ± 1.29	1.35
5	51.17 ± 0.05		2.27 ± 0.03		24.25 ± 0.04		44.81	45.51 ± 0.63	1.33
6	51.16 ± 0.05		2.28 ± 0.03		24.25 ± 0.02		44.85	45.33 ± 0.56	1.34
7	51.20 ± 0.16		2.29 ± 0.04		24.22 ± 0.05		44.85	45.08 ± 0.45	1.37
8	51.14 ± 0.03		2.30 ± 0.03		24.23 ± 0.03		44.47	46.02 ± 0.48	1.32

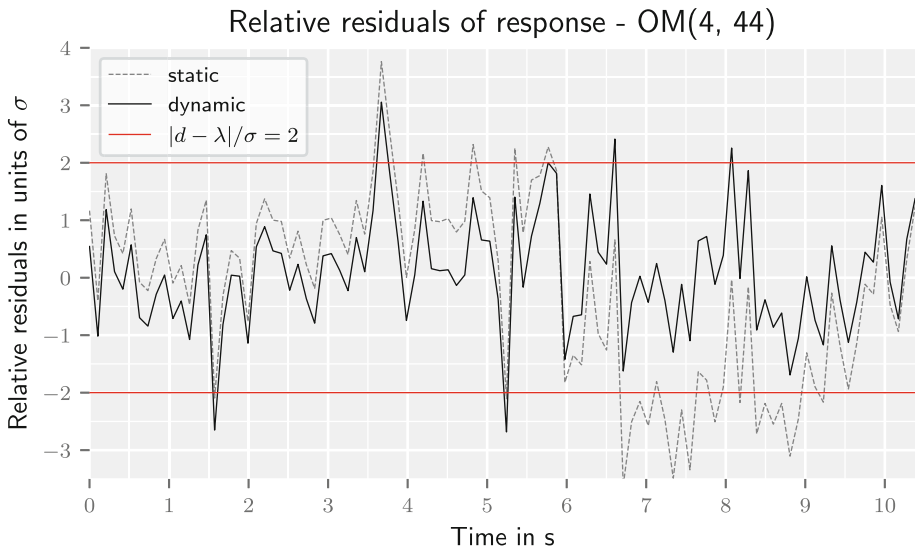

Fig. 4. Comparison of relative residuals of a static (dashed, gray) and dynamic (solid, black) inference source for run 2. The synthetic data were generated from a simulated dynamic source. The red line marks deviations as large as twice the shot noise σ of a Poisson process.

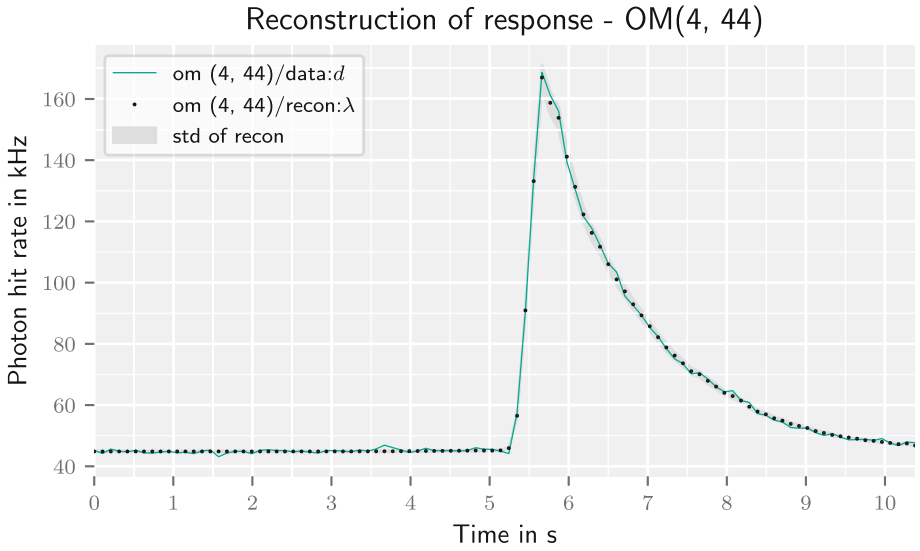

Fig. 5. Reconstruction (run 2) of the response of a dynamic inference source (black dots) with standard deviation (gray) of the posterior estimate is shown together with the synthetic data (green).

Table 12. GT and reconstructed parameters using a dynamic source model. The synthetic photon count data were generated from a simulated dynamic source. The positions are given during the time of highest photon emission t_h and at the end of the data set (t_{max}) to reflect the movement of the source.

Run	x_{t_h} (m)	$x_{t_{max}}$ (m)	y_{t_h} (m)	$y_{t_{max}}$ (m)	z_{t_h} (m)	$z_{t_{max}}$ (m)	$\mathscr{L}_{max}$ (GHz)	l_{att} (m)	$\overline{e^2}$
GT	51.15	52.10	2.15	3.10	24.15	25.10	45.71	45.00	1.02
0	51.05 ± 0.06	51.62 ± 0.08	2.14 ± 0.07	2.57 ± 0.22	24.17 ± 0.06	24.68 ± 0.17	45.42	45.01 ± 0.59	1.00
1	51.13 ± 0.12	51.78 ± 0.19	2.21 ± 0.13	2.82 ± 0.17	24.15 ± 0.06	24.78 ± 0.16	44.79	46.11 ± 0.66	1.05
2	51.10 ± 0.12	51.78 ± 0.22	2.20 ± 0.08	2.98 ± 0.19	24.18 ± 0.11	24.95 ± 0.13	44.87	45.80 ± 1.63	1.00
4	51.12 ± 0.10	51.86 ± 0.15	2.24 ± 0.08	2.88 ± 0.19	24.10 ± 0.06	24.81 ± 0.14	45.07	45.38 ± 0.51	1.07
5	51.07 ± 0.07	51.38 ± 0.16	2.11 ± 0.09	2.83 ± 0.21	24.18 ± 0.07	24.71 ± 0.15	45.18	45.05 ± 0.41	1.16
6	51.09 ± 0.12	51.96 ± 0.21	2.23 ± 0.08	3.00 ± 0.18	24.18 ± 0.09	24.84 ± 0.14	44.62	46.75 ± 0.76	1.01
7	51.02 ± 0.29	51.54 ± 0.33	2.26 ± 0.08	3.02 ± 0.23	24.11 ± 0.06	24.69 ± 0.11	44.66	46.07 ± 0.45	1.13
8	51.03 ± 0.06	51.56 ± 0.16	2.13 ± 0.07	2.83 ± 0.24	24.17 ± 0.08	24.82 ± 0.24	45.11	45.10 ± 0.86	1.02

numerical instabilities and is therefore removed. Despite all degeneracies, the position could be reconstructed within an error range of 20 cm and the maximum luminosity with an error less than 3%. The MGVI algorithm only gives lower bounds on the statistical uncertainties (Knollmüller and Enßlin 2020). Due to the strongly nonlinear nature of the model, we assume that the provided errors are most likely too small and hence have to be taken with caution.

Dynamic reconstruction of a synthetic static source

In order to demonstrate the consequences of using the wrong model on a static source, we also employed the dynamic model on the simulated static data set. This test will provide some intuition about movements that cannot be resolved by the detector setup and produce the same photon count data as a static source. As explained in "Reconstruction routine" section, the flash light curve, the position, and the detector efficiencies of the static reconstruction of the previous section are used as the initially assumed position for the dynamic reconstruction. In Table 9, the parameters of the dynamic model are summarized. The model parameters of the detector and its environment are the same as in the static reconstruction and provided in Table 8.

Since the reconstructions of the remaining observables were already discussed in the previous section and similar results could be observed for assuming a dynamic model, we focus on the discussion on the velocity. For the specific detector setup and source location, a source movement for a static source was reconstructed with a mean velocity $v \simeq 0.05\ \mathrm{m\ s^{-1}}$ in all runs. The results of the different reconstruction runs with varying seeds are summarized in

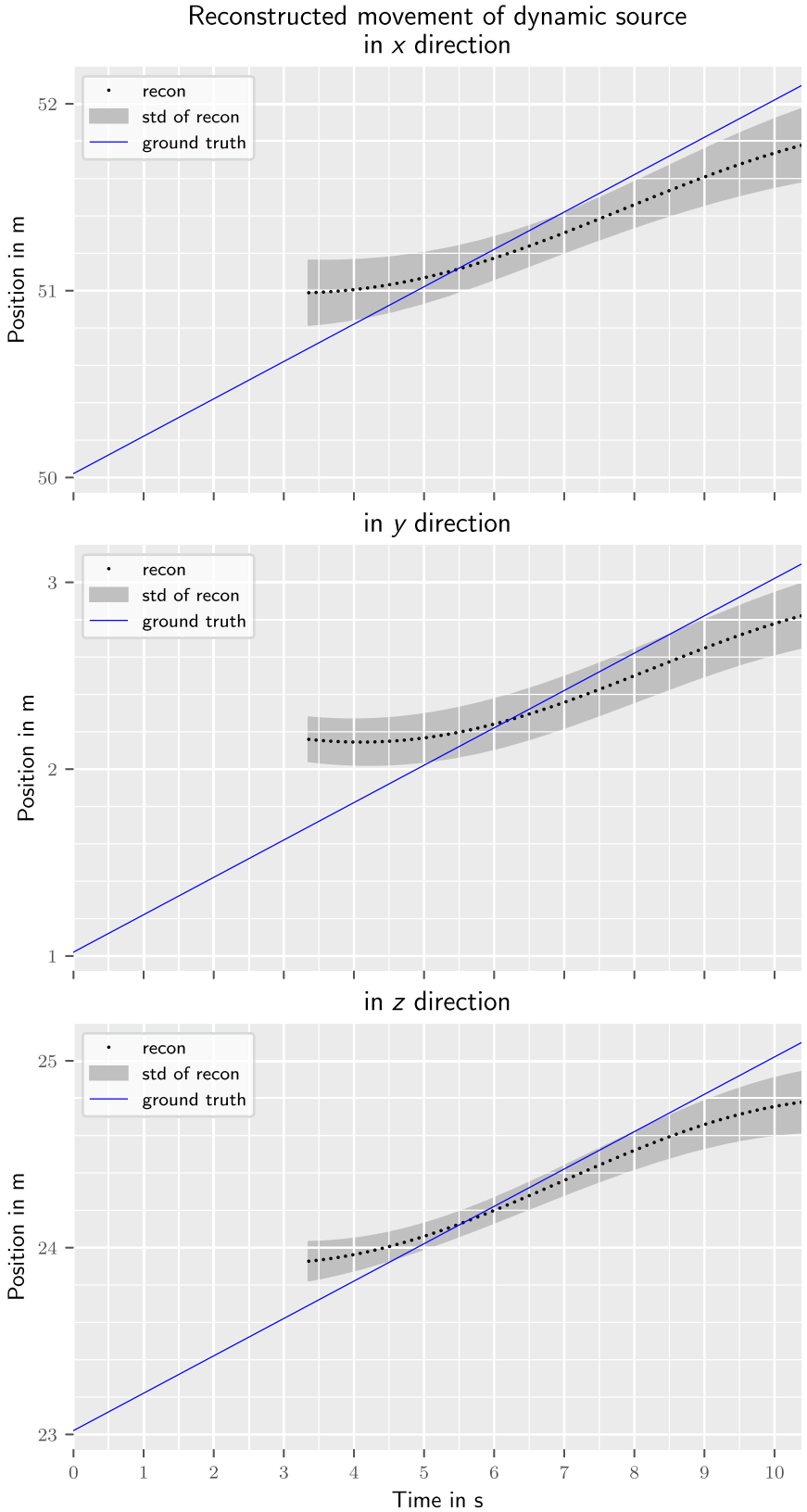

Fig. 6. Reconstruction (run 2) of the source position with the dynamic model. The simulated position (solid, blue) is shown in comparison to the reconstructed (circles, black) position. The reconstruction is only given for the time duration of the burst. Shown is also the uncertainty of the position reconstruction, shaded in gray.

Table 13. Data samples recorded by the ANTARES telescope reconstructed within this work.

Label	Time stamp (UTC)	OMs	Duration
Flash 1	11 Jan 2010, 04:12:35	$(4, 34-60)$	$\sim$ 10 s
Flash 2	11 Jan 2010, 04:13:20	$(4, 34-60)$	$\sim$ 9 s
Flash 3	19 Jan 2010, 22:28:10	$(4, 25-56)$	$\sim$ 11 s

Table 10. An illustration of the movement reconstructions in relation to the optical modules and the reconstructions of the nuisance parameters for the different runs in are given in the Supporting Information.

These results show that a source movement with a mean velocity $v = 0.05$ m s^{-1} cannot be distinguished from a static source for this detector setup. This detection limit depends highly on the setup of the detector, that is, number of optical modules and angular acceptance, and on the position of the source. Therefore, no global spatial resolution of this method can be defined. Our simulation result gives insights about the spatial resolution of a specific area and hence only about an excerpt of the global spatial resolution.

Future neutrino telescopes with different detector setups might increase the resolution. In the next section, we demonstrate that a movement reconstruction is theoretically possible with the ANTARES setup by applying the method to synthetic data drawn from a simulated dynamic source.

Complete reconstructions of a dynamic source

Static reconstruction

For the first part of the reconstruction routine, the same model parameters of the detector and source are used as in the static scenario given in Tables 7 and 8. The results of the reconstruction of a dynamic source using the static model are provided in Table 11.

Analyzing the results shows that a dynamic reconstruction might increase the accuracy of the estimates as the static model did not provide an optimal fit. Exemplarily, in Fig. 4, the residuals $e = ((d-\lambda)/\sigma)$ relative to the shot noise of the Poissonian measurement process $\sigma = \sqrt{\lambda}$ of optical module $(4, 44)$ are presented for run 0. During the light flash, an increased level of deviations $(2 < |e|)$ can be recognized. Residuals of optical module $(4, 44)$ close to zero can be found at $t_{\min} \simeq 6$ s. Detailed analysis shows that the relative residuals are positive $e > 0$ before the minimum $t < t_{\min}$ and negative $e < 0$ afterwards $t > t_{\min}$. Even though these variations are only slightly above the shot noise of the measurement process σ, they can be explained by a moving source. Hence, another reconstruction is performed to reduce these residuals by introducing a model for the source movement. The mean of the noise weighted squared residuals $\overline{e^2}$ over all optical modules are given in Table 11 for all reconstruction runs. In the following, this mean will be used as posterior check to quantify the accuracy of the reconstruction.

Table 14. Model parameters for the dynamic source model, the ANTARES telescope, and its surroundings on 11 January 2010 at 04:12 (UTC). This model is used to reconstruct biological sources detected by the ANTARES telescope on 11 January 2010 at 04:12 (UTC) and 04:13 (UTC) that emitted flashes 1 and 2.

Observable	Model	Model parameter
ε	Gaussian	$\mu_\varepsilon = \varepsilon_{\text{ANTARES}}$
		$\mu_{\sigma_\varepsilon} = 0.001$
		$\sigma_{\sigma_\varepsilon} = 0.001$
l_{att}	Gaussian	$\mu_{l_{\text{att}}} = 50\text{m}$
		$\mu_{\sigma_{l_{\text{att}}}} = 20\text{m}$
		$\sigma_{\sigma_{l_{\text{att}}}} = 5\text{m}$
n_i	Gaussian	$\mu_n = 60$ kHz
		$\mu_{\sigma_n} = 5$ kHz
		$\sigma_{\sigma_n} = 1$ kHz
$\vec{x}_0$	Gaussian	$\vec{\mu}_{\vec{x}_0} = (46\text{m}, 10\text{m}, 20\text{m})$
		$\mu_{\sigma_{j_0}} = 20\text{m} \quad \forall j \in x, y, z$
		$\sigma_{\sigma_{j_0}} = 5\text{m} \quad \forall j \in x, y, z$
v_i velocity	Correlated signal with $i \in \{x, y, z\}$	$\mu_v = 0$
		$\sigma_v = 0.1$
		$\sigma_{\sigma_v} = 0.05$
Flash shape	Correlated signal	See Supporting Information

Dynamic reconstruction

We use the same dynamic model as in the previous "Static reconstruction" section. The parameters used to define the source model are given in Table 9 and the parameters for the detector setup and its environment are presented in Table 8.

The increased level of variations between simulated and reconstructed response of the static model can be almost reduced to $0 < |e| < 2$ by the dynamic reconstruction as shown in Fig. 5, an improvement over the static model. The mean of the squared residuals $\overline{e^2}$ over all optical modules are given in Table 12 for all reconstruction runs. The results show a slight improvement compared to the residuals in Table 11 for the static reconstruction. The reconstructed movement visualized in Fig. 6 highlights the limits of the ANTARES detector for tracking individual organisms. A reasonable estimate of the movement could be reconstructed for a simulated light source moving linearly with a velocity $v = 0.2$ m s^{-1}. Although the reconstructed locations display deviations from the GT, the data residuals are close to the shot noise of the measurement process, that is, the deviations in the position cannot be resolved by the detector. These results show that a linear movement with a velocity $v = 0.2$ m s^{-1} is in principle

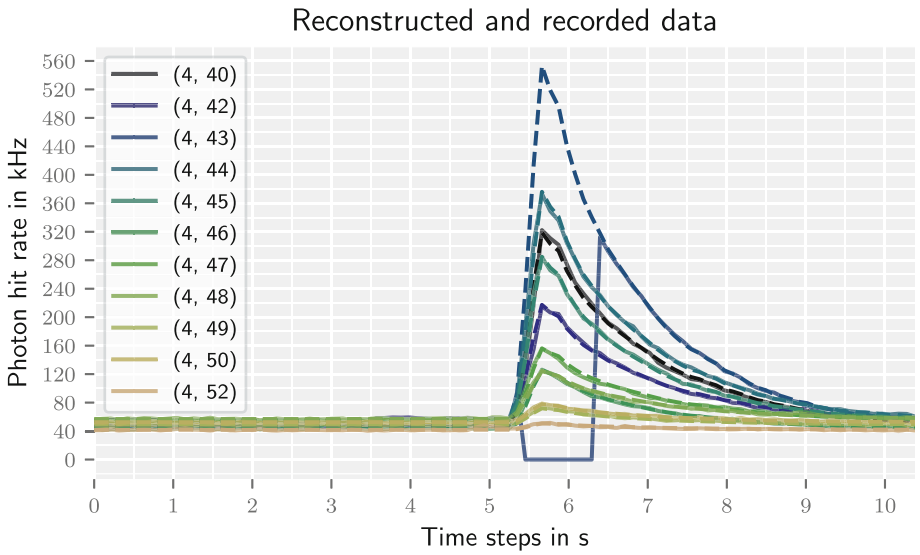

Fig. 7. Reconstruction (run 5) of the photon hit rate (dashed) during flash 1 detected by the ANTARES detector on 11 January 2010 at 04:12 (UTC). The recorded data are given in solid lines. Only an excerpt of optical modules is presented. For optical module (4,43), the recording was stopped when readout electronics were saturated.

Table 15. Reconstructed parameters assuming a dynamic light source model. The flash 1 light curve used for the reconstruction was recorded on 11 January 2010 at 04:12 (UTC) by the ANTARES telescope. The positions are given at different times as in Table 12.

Run	x_{t_h} (m)	$x_{t_{max}}$ (m)	y_{t_h} (m)	$y_{t_{max}}$ (m)	z_{t_h} (m)	$z_{t_{max}}$ (m)	$\mathscr{L}_{max}$ (GHz)	l_{att} (m)	$\overline{e^2}$
0	42.68 ± 0.07	42.34 ± 0.13	0.76 ± 0.18	0.28 ± 0.28	21.24 ± 0.10	20.84 ± 0.12	48.27	43.57 ± 1.41	2.95
1	42.86 ± 0.10	42.28 ± 0.21	1.57 ± 0.11	1.41 ± 0.21	21.04 ± 0.09	20.82 ± 0.15	40.82	51.83 ± 3.16	3.20
2	42.91 ± 0.12	42.06 ± 0.23	1.48 ± 0.12	1.01 ± 0.27	20.95 ± 0.17	20.35 ± 0.18	42.30	50.56 ± 1.73	2.78
3	45.57 ± 0.02	45.48 ± 0.06	15.79 ± 0.34	15.95 ± 0.74	16.09 ± 0.18	16.04 ± 0.19	83.42	32.09 ± 0.84	8.08
5	42.84 ± 0.11	42.07 ± 0.13	1.29 ± 0.09	1.26 ± 0.17	21.15 ± 0.10	20.75 ± 0.14	44.69	48.51 ± 2.00	2.73
6	45.62 ± 0.05	45.54 ± 0.05	15.03 ± 0.27	15.11 ± 0.22	16.07 ± 0.05	15.93 ± 0.05	79.86	34.97 ± 1.37	6.66
7	42.82 ± 0.07	42.04 ± 0.11	1.13 ± 0.17	1.03 ± 0.16	21.21 ± 0.07	20.99 ± 0.13	46.19	47.25 ± 2.55	2.87
8	42.93 ± 0.12	42.26 ± 0.21	1.51 ± 0.22	1.52 ± 0.26	21.02 ± 0.02	20.27 ± 0.06	41.57	49.90 ± 0.56	3.07
9	42.82 ± 0.12	42.07 ± 0.24	1.27 ± 0.25	1.27 ± 0.24	21.09 ± 0.07	20.85 ± 0.22	43.10	51.62 ± 1.61	2.96

resolvable, but is not free of systematics. Besides the degeneracies, which are already discussed in "Static reconstruction of a synthetic static source" section, the angular acceptance has great impact on the positioning of a light source. The optical modules of the ANTARES detector have a wide angular acceptance (Amram et al. 2002). Therefore, changes of the source position in nonradial directions do not lead to significant changes in the photon count numbers, which reduces the possibility to recognize position changes. Multiple optical modules with smaller angular acceptance might increase the spatial resolution. The reconstructed flash characteristics differ only slightly between the reconstruction with a static and a dynamic model as shown in Table 12. Therefore, with regard to the bioluminescence flash light curves, both models are sufficient to reconstruct a reasonable estimate for velocities below $v = 0.2$ m s^{-1}. An illustration of the movement reconstructions and the nuisance parameters for the different runs is also given in the Supporting Information.

Data analysis

After presenting the capabilities and limitations of the method, the reconstruction routine is applied on data sets of the ANTARES Collaboration using different random seeds. The complete routine consists of a reconstruction using a static model and a dynamic model as explained in "Reconstruction routine" section. The outcomes of the static and dynamic reconstruction provide similar results regarding the position during the time of highest photon emission and the flash light curve. Therefore, it is sufficient to present only the results of the final step using the dynamic model.

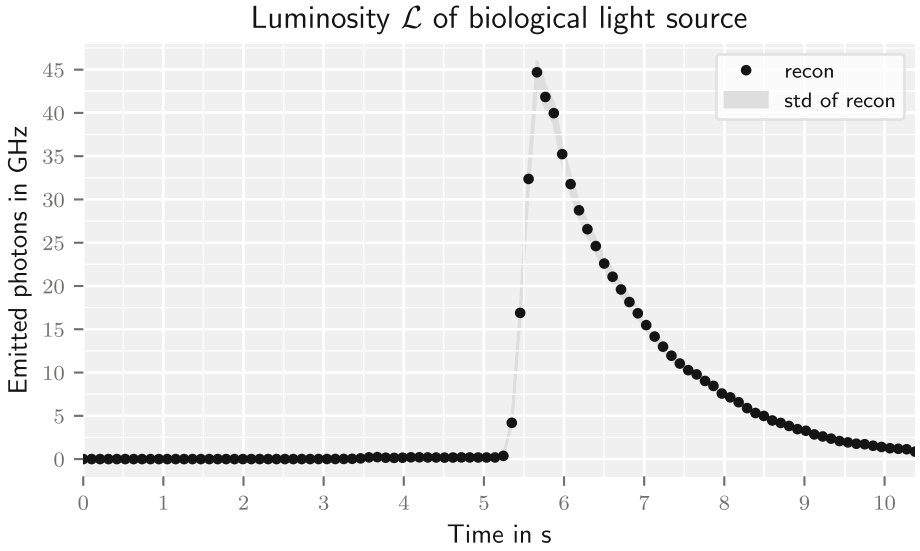

Fig. 8. Reconstruction (run 5) of the bioluminescence flash 1 light curve detected by the ANTARES detector on 11 January 2010 at 04:12 (UTC).

For the reconstruction, we rely on flash light curves that were detected over several storeys to be able to reduce the degeneracy. Since the process of finding such light emissions has not been automated, only a small excerpt of flash patterns found in the ANTARES data is analyzed here. We identified three suitable bioluminescence events in the data of early 2010, which we analyze in the following. The data samples cover observation times of $9-11$ s. In Table 13, we label the samples used for the following reconstructions and state the array of optical modules that detected the flash neglecting malfunctioning modules.

ANTARES recordings of flash 1

Starting with 11 January 2010 at 04:12:35 (Coordinated Universal Time [UTC]), a flash pattern was recorded by 25 optical modules over 9 storeys. The recordings of six optical modules have already been shown in Fig. 2 in "Data" section. Flash 1 peaked at around 40 s after the start of the recording. Two optical modules $(4,41)$ and $(4,51)$, within the optical module array $(4,34-60)$, did not record any photon counts. Furthermore, optical module $(4,43)$ did not record any photon for the period of highest luminosity, since the readout electronics were saturated. The model parameters used for the reconstruction are presented in Table 14.

The reconstruction was performed with 10 different random seeds referred to as runs in the following. For seven runs (0, 1, 2, 5, 7, 8, and 9), a match between recorded and reconstructed photon data could be observed. Exemplarily, the posterior mean of the reconstructed data of run 5 is shown together with the recorded data in Fig. 7. Only a limited number of optical modules are presented. However, for runs 3 and 6, a clear mismatch between reconstructed and recorded data could be identified (*see* residuals in Table 15). As a consequence, we regard these results as less likely. Run 4 had to be stopped due to numerical issues. The reconstruction of the position and movement as well as the characteristics of the bioluminescence light curve for the nine different random seeds are summarized in Table 15. For the seven runs, similar results could be observed. The maximum of the absolute deviations of each dimension $(\Delta x_{\max} = 0.25\ \text{m},\ \Delta y_{\max} = 0.81\ \text{m},\ \Delta z_{\max} = 0.29\ \text{m})$ for different runs are calculated at the positions with highest luminosity of these runs. For the attenuation length, absolute deviations of up to 8 m and for the maximum luminosity deviations of up to 18% can be recognized. We assume that the variability of the luminosity for different runs can be ascribed to the degeneracy between luminosity and attenuation length. Therefore, precise measurements of the attenuation length and hence a more informative prior will reduce this variability and increase the consistency of the reconstruction. In addition, the reconstruction shows that the source is located above one storey containing one optical module that did not record any data. These missing data increase also the uncertainty of the localization. Nevertheless, the position as well as the movement estimates for the seven runs are consistently within a range of 1 m, which we regard as acceptable precision. Hence, we conclude that we localized a light source a few meters away from the optical modules and could even identify the direction of the movement away from the detector. In Fig. 8, the reconstructed flash light curve of run 5 is presented as this run provides the smallest data residuals (*see* Table 15). We also illustrate the estimated movement in Fig. 9. The reconstructed nuisance parameters are provided in the Supporting Information.

ANTARES recordings of flash 2

A second flash was observed on the same day at 04:13:20 (UTC) recorded by optical modules within the array

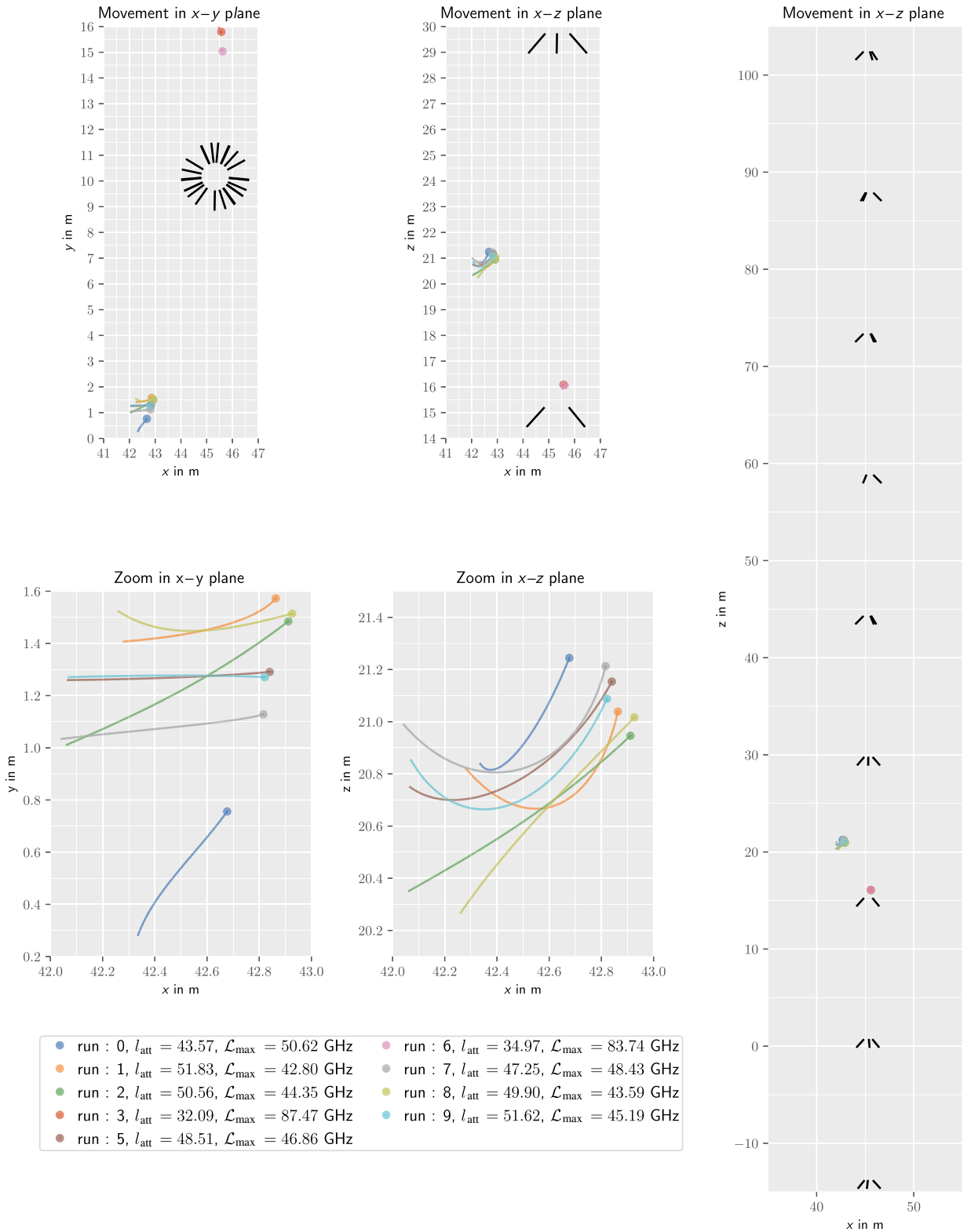

Fig. 9. Movement reconstruction of a light source emitting flash 1 detected by the ANTARES detector on 11 January at 04:12 (UTC). The light sources are represented by colored dots and lines. The dot indicates the position with the highest luminosity and the line represents the movement afterwards. The position and direction (downwards in *z*-direction) of optical modules are shown as dashes. (Left, middle column) Source movement in relation to the detector and zoomed excerpts. (Right column) Overview of all optical modules used for the reconstruction.

Table 16. Reconstructed parameters assuming a dynamic light source. The flash 2 light curve used for the reconstruction was recorded on 11 January 2010 at 04:13 (UTC) by the ANTARES telescope. The positions are given at different times as in Table 12.

Run	x_{t_h} (m)	$x_{t_{max}}$ (m)	y_{t_h} (m)	$y_{t_{max}}$ (m)	z_{t_h} (m)	$z_{t_{max}}$ (m)	$\mathscr{L}_{max}$ (GHz)	l_{att} (m)	$\overline{e^2}$
0	48.42 ± 0.95	45.17 ± 1.75	−9.18 ± 0.81	−7.40 ± 1.10	19.70 ± 0.28	19.97 ± 0.50	42.77	58.11 ± 6.44	2.92
1	48.05 ± 0.89	45.97 ± 0.99	−8.84 ± 0.61	−9.46 ± 0.88	19.49 ± 0.69	19.86 ± 0.84	39.23	70.25 ± 13.93	2.83
2	48.59 ± 0.75	46.28 ± 1.52	−9.83 ± 0.88	−8.49 ± 1.28	18.94 ± 0.83	20.00 ± 1.06	50.88	48.58 ± 1.91	4.47
3	48.83 ± 1.06	46.03 ± 2.07	−10.22 ± 0.75	−10.51 ± 1.05	19.50 ± 0.51	20.14 ± 0.80	50.95	49.83 ± 2.59	2.61
4	47.72 ± 0.53	45.97 ± 0.55	−9.28 ± 0.40	−9.36 ± 0.72	19.53 ± 0.38	19.65 ± 0.45	42.82	60.08 ± 6.07	2.92
5	48.06 ± 1.13	45.96 ± 1.08	−9.10 ± 0.51	−9.30 ± 0.80	20.02 ± 1.14	20.60 ± 1.10	43.03	59.48 ± 5.62	3.07
6	48.62 ± 0.44	46.29 ± 0.76	−9.94 ± 0.42	−9.06 ± 0.68	20.22 ± 0.48	20.13 ± 0.62	47.52	54.43 ± 4.12	3.10
8	48.29 ± 1.27	45.13 ± 1.58	−9.38 ± 0.46	−10.55 ± 0.64	19.72 ± 0.38	19.95 ± 0.66	42.97	57.59 ± 5.97	2.63
9	48.16 ± 0.36	46.36 ± 0.96	−9.27 ± 1.03	−9.14 ± 1.41	19.33 ± 0.60	19.50 ± 0.64	43.04	60.30 ± 5.91	3.29

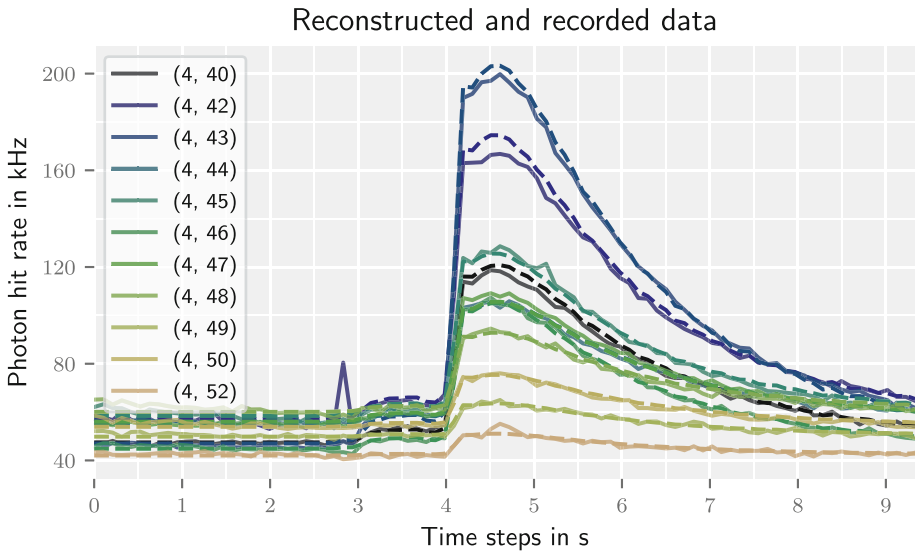

Fig. 10. Reconstruction (run 3) of the photon hit rate (dashed) during flash 2 detected by the ANTARES detector on 11 January 2010 at 04:13 (UTC). The recorded data are given in solid lines. Only an excerpt of optical modules is presented.

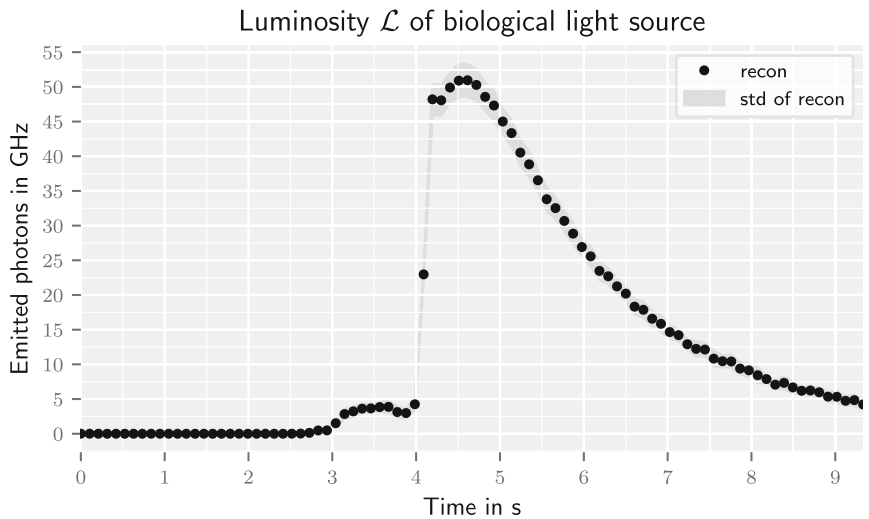

Fig. 11. Reconstruction (run 3) of the emitted bioluminescence flash 2 light curve detected by the ANTARES detector on 11 January 2010 at 04:13 (UTC).

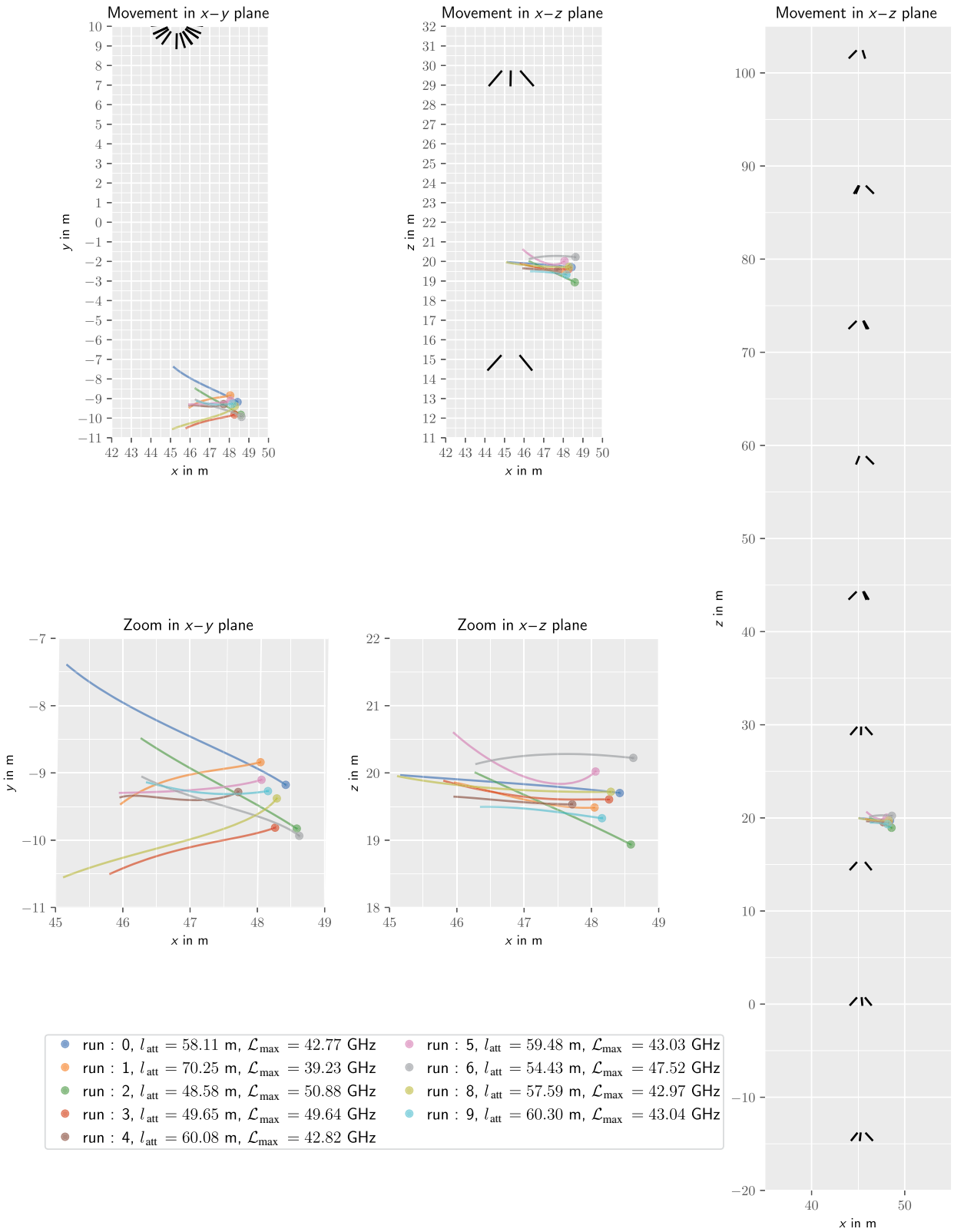

Fig. 12. Movement reconstruction of a light source emitting flash 2 detected by the ANTARES detector on 11 January at 04:13 (UTC). The light sources are represented by colored dots and lines. The dot indicates the position with the highest luminosity and the line represents the movement afterwards. The position and direction (downwards in *z*-direction) of optical modules are shown as dashes. (Left, middle column) Source movement in relation to the detector and zoomed excerpts. (Right column) Overview of all optical modules used for the reconstruction.

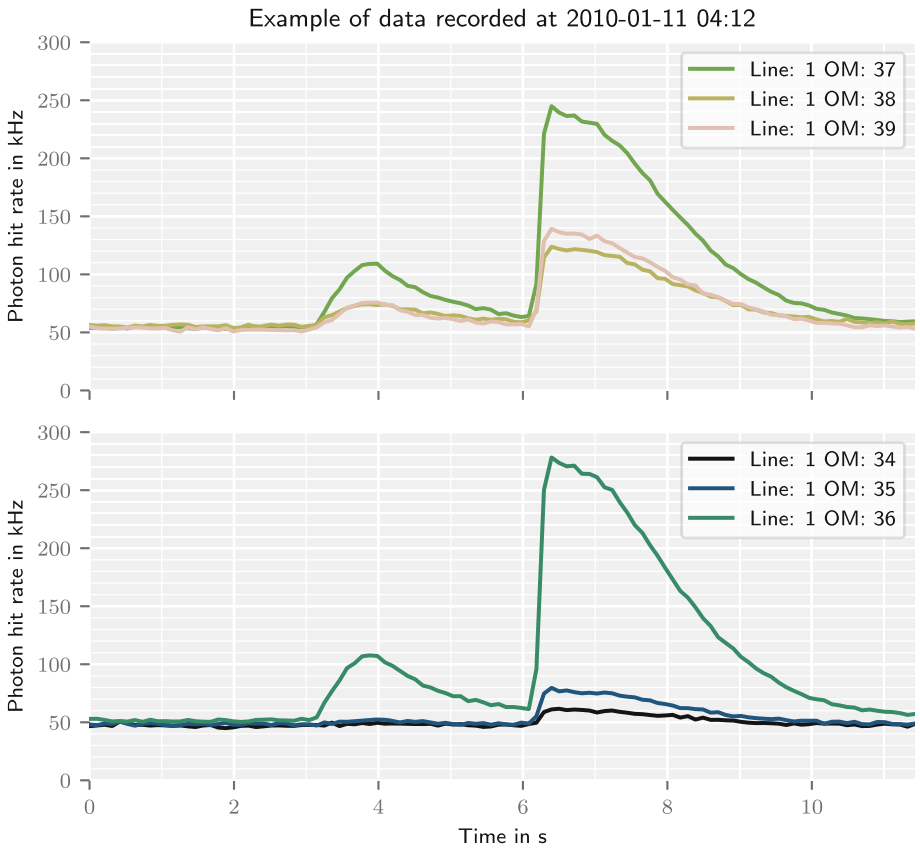

Fig. 13. Recorded photon hits of flash 3 taken from six optical modules (OM). The two plots show each the photon hits of the optical modules in one storey. Two consecutive flashes can be observed.

$(4,34-60)$. Flash 2 peaked at around 85 s after the start of the recording shown in Fig. 2 in "Data" section. In Fig. 2, the two consecutive flashes can be seen. The optical modules $(4,41)$ and $(4,51)$ did not detect the second flash. All model parameters are the same as for flash 1 given in Table 14. The ANTARES detector is still in a similar configuration, because the second flash occurred around 1 min after the first. Furthermore, due to similar photon count data of flashes 1 and 2, the same a priori assumption of the position can be taken.

The results of the 10 runs with different seeds are summarized in Table 16. The reconstructed nuisance parameters are provided in the Supporting Information. The photon hit rate could be reconstructed for most runs sufficiently (*see* residuals in Table 16). Run 7 had to be stopped due to numerical issues and run 2 has higher residuals than the remaining runs. As example, the recorded and reconstructed photon data of run 3 are presented in Fig. 10. In contrast to flash 1, the maximum of the absolute deviations of the position (excluding run 2) $(\Delta x_{\max} = 1.11\ \text{m},\ \Delta y_{\max} = 1.12\ \text{m},\ \Delta z_{\max} = 0.89\ \text{m})$ between the several runs are higher. We argue that this increase in variability can be explained by the varying spatial resolution throughout the detector volume. The maximum deviation of the attenuation length is around 20 m and the maximum deviation of the maximum luminosity around 30%. Precise knowledge of the attenuation length should reduce this variability as already explained in "ANTARES recordings of flash 1" section. In addition, three recording optical modules close to the source will also increase the accuracy of the reconstruction. Despite the variabilities between the different runs, similar movements can be observed. It is noticeable that each run reconstructed a velocity that is higher than the velocity of flash 1 and reaches values of up to $v \simeq 0.5\ \text{m s}^{-1}$. The main movement is along the x dimension which is similar to the movement of flash 1. The impact of the detector setup on the reconstruction of the movement has to be studied in detail since a spatial resolution depending on direction might introduce a movement during the reconstruction. Due to reduced residuals for the dynamic reconstruction compared to the static reconstruction, we argue that a movement of the source is likely. In addition, analysis of the water current could give insights whether the organism actively moved or a passive movement was recorded. In Fig. 11, we present the flash light curve of run 3, since the smallest data residuals could be calculated for this run (*see* Table 16). The reconstructed movement of run 3 is illustrated in Fig. 12. The second flash occurred only 44 s after and around 13 m away from the first flash. A causal connection between both light emission is likely. Despite the movement reconstruction for flash 1 and the

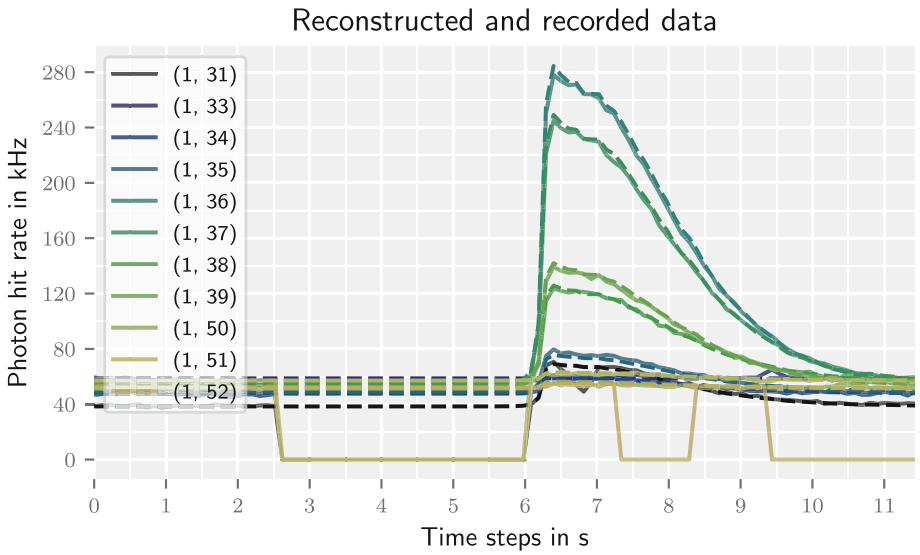

Fig. 14. Reconstruction (run 1) of the photon hit rate (dashed) during flash 3 detected by the ANTARES detector on 19 January 2010 at 22:28 (UTC). The recorded data are given in solid lines. Only an excerpt of optical modules is presented.

Table 17. Model parameters for the initial position of a dynamic source. The model is used to reconstruct a biological source detected by the ANTARES telescope on 19 January 2010 at 22:28 (UTC) that emitted flash 3. The other model parameters are taken from Table 14.

Observable	Model	Model parameter
$\vec{x}_0$	Gaussian	$\vec{\mu}_{\vec{x}_0} = (4.5\ \text{m},\ 97\ \text{m},\ -20\text{m})$
		$\mu_{\sigma_{j_0}} = 20\ \text{m}\ \ \forall j \in x, y, z$
		$\sigma_{\sigma_{j_0}} = 5\text{m}\ \ \forall j \in x, y, z$
v_i velocity	Correlated signal with $i \in \{x, y, z\}$	$\mu_v = 0$
		$\sigma_v = 0.1$
		$\sigma_{\sigma_v} = 0.05$

deviation between the light curve structures of the two flashes, it cannot be excluded that both flashes were emitted by the same organism.

ANTARES recordings of flash 3

The last bioluminescence flash that is analyzed within this work occurred on 19 January 2010 at 22:28 (UTC) (Fig. 13). In comparison to the previous models, we adjust the position model parameter since the flash was detected by different optical modules (given in Table 13). For this recording, we noticed two flashes within 7 s. Since our model describes a single light source, we mask the first flash in order to localize only the second. In Fig. 14, the reconstructed photon data are presented and the masks used for the reconstruction can be recognized. We argue that the first flash affects the photon number during the second flash (*see* Fig. 13) and hence it might have impact on the localization of the second flash if the two flashes were emitted from different locations. We also observed that six optical modules, $(1,28)$, $(1,32)$, $(1,40)$, $(1,44)$, $(1,47)$, and $(1,48)$, did not record data or had a low efficiency. Nevertheless, we apply our routine onto this data set and discuss the results. The position model parameters are given in Table 17. The remaining model parameters are the same as in Table 14.

Table 18 summarizes the reconstruction runs and shows that the residuals are higher compared to the previous flashes. Run 3 shows the highest residual and will be neglected in the following discussion. Run 2 had to be stopped due to numerical issues. We recognize small maximum deviations of the position $(x_{\max} = 0.82\ \text{m},\ y_{\max} = 0.30\ \text{m},\ z_{\max} = 0.22\ \text{m})$ (excluding run 3). For the attenuation length, the algorithm reconstructed unphysical values up to 107 m. We argue that the occurrence of two flashes as well as six missing optical modules had impact on the inference. As mentioned in "Static reconstruction of a synthetic static source" section, the reconstruction of the attenuation length might only be reasonable in a close to perfect scenario, that is, constant background noise and hence only a single isotropic light source. Future work could improve this issue by introducing a reconstruction routine that includes multiple light sources. The reconstructed nuisance parameters are provided in the Supporting Information.

Discussions

Strengths, limitations, and improvements

The application of the method on synthetic and real data reveals both the strengths and the limitations of our approach. Despite the degeneracy between most of the variables (position, efficiency, attenuation length, luminosity, and background noise) and missing optical modules, we were able to obtain reasonable estimates for the positions of flashes 1 and 2. Multiple

Table 18. Reconstructed parameters assuming a dynamic light source. The flash 3 light curve used for the reconstruction was recorded on 19 January 2010 at 22:28 (UTC) by the ANTARES telescope. The positions are given at different times as in Table 12.

Run	x_{t_h} (m)	$x_{t_{max}}$ (m)	y_{t_h} (m)	$y_{t_{max}}$ (m)	z_{t_h} (m)	$z_{t_{max}}$ (m)	$\mathscr{L}_{max}$ (GHz)	l_{att} (m)	$\overline{e^2}$
0	−8.67 ± 0.23	−9.51 ± 0.29	103.60 ± 0.18	104.46 ± 0.36	−20.48 ± 0.06	−21.28 ± 0.15	28.53	106.18 ± 4.47	3.95
1	−8.87 ± 0.25	−9.22 ± 0.54	103.67 ± 0.18	104.69 ± 0.44	−20.44 ± 0.09	−21.05 ± 0.18	29.61	93.98 ± 5.11	3.67
3	−8.36 ± 0.28	−9.73 ± 0.40	102.87 ± 0.13	104.66 ± 0.49	−20.79 ± 0.09	−21.39 ± 0.16	23.85	147.96 ± 14.79	6.44
4	−8.85 ± 0.16	−9.14 ± 0.29	103.64 ± 0.12	104.57 ± 0.36	−20.47 ± 0.06	−21.12 ± 0.17	29.51	107.59 ± 8.18	3.91
5	−9.13 ± 0.17	−9.67 ± 0.28	103.71 ± 0.13	104.81 ± 0.43	−20.32 ± 0.06	−21.09 ± 0.20	29.52	98.75 ± 5.17	4.21
6	−9.01 ± 0.26	−9.41 ± 0.37	103.81 ± 0.15	104.68 ± 0.39	−20.38 ± 0.07	−21.01 ± 0.17	29.06	107.65 ± 5.68	4.15
7	−9.49 ± 0.78	−10.03 ± 0.90	103.90 ± 0.16	105.02 ± 0.35	−20.26 ± 0.04	−21.18 ± 0.13	30.08	103.63 ± 6.49	4.41
8	−9.13 ± 0.26	−9.51 ± 0.43	103.79 ± 0.10	104.94 ± 0.37	−20.27 ± 0.02	−21.03 ± 0.13	30.08	97.30 ± 5.54	4.21
9	−9.28 ± 0.12	−9.59 ± 0.15	103.78 ± 0.12	104.53 ± 0.44	−20.32 ± 0.04	−20.94 ± 0.13	30.21	91.68 ± 6.71	3.67

runs of the reconstruction routine with different random seeds of the used MGVI algorithm showed mostly consistent results. For a synthetic source, we were also able to infer the attenuation length and luminosity correctly. The reconstruction of those parameters with the ANTARES data sets shows a larger uncertainty but physical reasonable results. Therefore, we note that precise prior knowledge of the attenuation length has to be available to ensure accurate reconstructions of the luminosity. Nevertheless, even with deviations of the attenuation length, we could reconstruct the source position consistently within a range of 1 m. The reconstruction runs of flashes 1 and 2 could also give evidence on the varying spatial resolution, that is, the position estimates of flash 2 showed a much larger diversity of the different reconstruction runs than the estimates of flash 1. The reconstructed source of flash 2 has a larger distance to the optical modules than the source of flash 1 and hence the spatial resolution might be reduced. Synthetic tests also verified that the reconstruction of movements is possible. However, due to the varying spatial resolution the estimates of the movement, especially the total velocity, have to be taken with caution. A poor spatial resolution is less constraining for movements of the reconstructed source. Here, a measure of spatial resolution might be helpful to identify such regions in the future. The errors given by the MGVI algorithm are only lower bounds to the true statistical errors (Knollmüller and Enßlin 2020). Since the magnitude of the mismatch depends on the nonlinearity of the problem (i.e., the deviation of the posterior from Gaussianity) and our model is very nonlinear, the inferred uncertainties are most likely far too small and hence should be taken with strong caution.

The accuracy of the position reconstructions and the spatial resolution can be increased by a different detector setup. Reducing the angular acceptance and increasing the number of OMs will increase the spatial resolution. An OM with sharp angular acceptance is able to resolve small position changes more precisely, since changes of the angle between OMs and light source lead to significant changes in the arriving photon numbers. However, a small angular acceptance leads also to a smaller monitored volume, which can be compensated by using more OMs. Furthermore, precise measurements of the attenuation length will help to break the degeneracy with the luminosity and hence increase the total accuracy of the reconstructions.

In this work, our verifications have been limited to tests with synthetic data sets. During measurements of the attenuation length or in similar experiments, light has been emitted from a fixed position and recorded by the optical modules (Aguilar et al. 2005; Adrián-Martínez et al. 2012*b*). These data sets could be used to verify our reconstructions, but were not available to us. Furthermore, our model is limited to a single light source which leads to unrepresentative reconstructions if multiple sources are present as discussed in "ANTARES recordings of flash 3" section. We argue that our algorithm could be extended to multiple sources.

Interpretations

Although only a small number of photon data recorded by the ANTARES telescope has been analyzed in this study, the results already provide some insights on the biological sources. Our localizations of bioluminescence events showed that the light emitting organisms were sufficiently far from optical modules, such that no direct collision with the detector structure caused the bioluminescence response. The majority of the bioluminescence activity, which has been recorded by neutrino telescopes, show inertial rhythms that match inertial current periodicity and hence are assumed to be provoked by collisions with the detector infrastructure (Aguzzi et al. 2017). Further measurements of the environment, for example, sea current measurements, need to be taken into account to verify the natural occurrence of bioluminescence events, that is, to exclude that the detector caused the light emission. In a study by Meighen-Berger et al. (2021), it has been simulated that OMs induce turbulence even at a distance 20 m, which can cause bioluminescence responses. Assuming that the analyzed bioluminescence flashes were not triggered by the flow of water masses, the PMTs recorded organisms in their natural environment utilizing their bioluminescence trait for communication, predation, or mating.

According to previous studies non-stimulated light emissions are rare (Widder et al. 1989; Priede et al. 2006; Tamburini et al. 2013; Martini et al. 2014) and hardly distinguishable from stimulated ones using moving detectors (Priede et al. 2006). In addition, the monitoring of a given volume to quantify the occurrence of light emissions comes with drawbacks. The monitored volume depends highly on the luminosity of the individual organisms, that is, bright organisms can be monitored within the whole ANTARES detector infrastructure, whereas darker organisms can only be monitored in regions near the PMTs. Therefore, the monitoring volume will always depend on the emitted luminosity. Our method opens the possibility to study in situ bioluminescence within the large data sets of deep-sea neutrino telescopes. The localization helps to distinguish flashes triggered by the turbulence around the detector and rare events of naturally occurring bioluminescence. Furthermore, the inferred location and absolute luminosity allows extrapolate densities even into "dark" areas, thereby tackling the aforementioned problem of the luminosity-dependent detector volume.

Besides the localization of the light source, our method also reconstructs the characteristic light curves of individual organisms. To our knowledge, currently, no data set exists that allows classifications of deep-sea bioluminescence flashes. The similarities of the light curves between different taxes will make classifications difficult. Further analysis of data recorded by deep-sea neutrino telescopes could give insights about the frequency and distribution of such spontaneous events. This offers the possibility to estimate the density of luminescent organisms in different regions of the detector infrastructure. Here, interdisciplinary collaborations become essential to study the behavior of luminescent organisms with data of neutrino telescopes, for example, to extract information about active migrations as already proposed by Chatzievangelou et al. (2021). Our method could be a useful tool for this task.

Conclusion

This work shows the potential of bioluminescence studies with a neutrino telescope in the deep sea and highlights the biological activity information that can be extracted. But it also clearly points out the limitations of the bioluminescence studies in regard to the spatial resolution due to the architecture of the detector.

The proposed method is generic and can be applied on data sets of different underwater neutrino telescopes. The development and work on the new neutrino telescope in the Mediterranean Sea, KM3NeT (Adrián-Martínez et al. 2016), offer a detector architecture which is even more suitable for the study of luminescent organisms. Each optical module of KM3NeT will be equipped with 31 PMTs having a narrower angular acceptance compared to the ANTARES setup (Adrián-Martínez et al. 2016). This increases the spatial resolution of the positioning.

Alternatively, the method can be adjusted to be applied on data of the Pacific Ocean Neutrino Experiment (P-ONE) (Agostini et al. 2020) and its precursor STRAW (STRings for ABsorption length in Water) (Bailly et al. 2021). This would allow a global analysis of bioluminescence events.

For future systematic surveys of bioluminescence, the method needs to be automatized and optimized. An optimized framework can be used to build a catalogue of various types of bioluminescence light curves including the position of the source. In this work, we also showed that the tracking of light sources is possible. Therefore, this method can also be used to analyze the movement behavior of deep-sea organisms, which still little is known about.

Data availability statement

Release and preservation of data used by the ANTARES Collaboration as the basis for publications is guided by the ANTARES policy as written in its Memorandum of Understanding among the collaborating Institutions. Parts of the data are available on the website https://antares.in2p3.fr/publicdata.html and periodically updated.

References

Adrián-Martínez, S., and others. 2012*a*. The positioning system of the ANTARES neutrino telescope. J. Instrum. **7**: 08002. doi:10.1088/1748-0221/7/08/t08002

Adrián-Martínez, S., and others. 2012*b*. Measurement of the group velocity of light in sea water at the ANTARES site. Astropart. Phys. **35**: 552–557. doi:10.1016/j.astropartphys.2011.12.003

Adrián-Martínez, S., and others. 2016. Letter of intent for KM3NeT 2.0. J. Phys. G Nucl. Part. Phys. **43**: 084001. doi:10.1088/0954-3899/43/8/084001

Ageron, M., and others. 2011. ANTARES: The first undersea neutrino telescope. Nucl. Instrum. Methods Phys. Res. Sect. A **656**: 11–38. doi:10.1016/j.nima.2011.06.103

Agostini, M., and others. 2020. The pacific ocean neutrino experiment. Nat. Astron. **4**: 913–915. doi:10.1038/s41550-020-1182-4

Aguilar, J. A., and others. 2005. Transmission of light in deep sea water at the site of the Antares neutrino telescope. Astropart. Phys. **23**: 131–155. doi:10.1016/j.astropartphys.2004.11.006

Aguilar, J. A., and others. 2007. The data acquisition system for the ANTARES neutrino telescope. Nucl. Instrum. Methods Phys. Res. Sect. A **570**: 107–116. doi:10.1016/j.nima.2006.09.098

Aguzzi, J., and others. 2017. Inertial bioluminescence rhythms at the Capo Passero (KM3NeT-Italia) site, Central Mediterranean Sea. Sci. Rep. **7**: 44938. doi:10.1038/srep44938

Aguzzi, J., and others. 2019. New high-tech flexible networks for the monitoring of deep-sea ecosystems. Environ. Sci. Tech. **53**: 6616–6631. doi:10.1021/acs.est.9b00409

Albert, A., and others. 2018. Long-term monitoring of the ANTARES optical module efficiencies using ^{40}K decays in sea water. Eur. Phys. J. C **78**: 1–8. doi:10.1140/epjc/s10052-018-6132-2

Amram, P., and others. 2000. Background light in potential sites for the ANTARES undersea neutrino telescope. Astropart. Phys. **13**: 127–136. doi:10.1016/s0927-6505(99)00118-8

Amram, P., and others. 2002. The ANTARES optical module. Nucl. Instrum. Methods Phys. Res. Sect. A **484**: 369–383. doi:10.1016/s0168-9002(01)02026-5

Amram, P., and others. 2003. Sedimentation and fouling of optical surfaces at the ANTARES site. Astropart. Phys. **19**: 253–267. doi:10.1016/s0927-6505(02)00202-5

Arras, P., and others. 2019. Nifty5: Numerical information field theory v5. Astrophysics Source Code Library.

Arras, P., P. Frank, P. Haim, J. Knollmüller, R. Leike, M. Reinecke, and T. Enßlin. 2022. Variable shadow of M87* from space, time and frequency resolved interferometry. Nat Astron **6**: 259–269. doi:10.1038/s41550-021-01548-0

Bailly, N., and others. 2021. Two-year optical site characterization for the pacific ocean neutrino experiment (p-one) in the Cascadia basin. The. Eur. Phys. J. C **81**: 1071. doi:10.1140/epjc/s10052-021-09872-5

Batchelder, H., E. Swift, and J. R. Van Keuren. 1992. Diel patterns of planktonic bioluminescence in the northern Sargasso Sea. Mar. Biol. **113**: 329–339. doi:10.1007/BF00347288

Batchelder, H. P., and E. Swift. 1989. Estimated near-surface mesoplanktonic bioluminescence in the western North Atlantic during July 1986. Limnol. Oceanogr. **34**: 113–128. doi:10.4319/lo.1989.34.1.0113

Buskey, E., and E. Swift. 1990. An encounter model to predict natural planktonic bioluminescence. Limnol. Oceanogr. **35**: 1469–1485. doi:10.4319/lo.1990.35.7.1469

Buskey, E. J. 1992. Epipelagic planktonic bioluminescence in the marginal ice zone of the Greenland Sea. Mar. Biol. **113**: 689–698. doi:10.1007/BF00349712

Chatzievangelou, D., and others. 2021. Integrating diel vertical migrations of bioluminescent deep scattering layers into monitoring programs. Front. Mar. Sci. **8**. doi:10.3389/fmars.2021.661809

Craig, J., I. Priede, J. Aguzzi, and A. Jamieson. 2015. Abundant bioluminescent sources of low-light intensity in the deep Mediterranean Sea and North Atlantic Ocean. Mar. Biol. **162**: 1637–1649. doi:10.1007/s00227-015-2700-2

Cronin, H., J. Cohen, J. Berge, G. Johnsen, and M. Moline. 2016. Bioluminescence as an ecological factor during high Arctic polar night. Sci. Rep. **6**: 36374. doi:10.1038/srep36374

Danovaro, R., and others. 2017. An ecosystem-based deep-ocean strategy. Science **355**: 452–454. doi:10.1126/science.aah7178

Danovaro, R., and others. 2020. Ecological variables for developing a global deep-ocean monitoring and conservation strategy. Nat. Ecol. Evol. **4**: 181–192. doi:10.1038/s41559-019-1091-z

Esaias, W. E., H. C. Curl Jr., and H. H. Seliger. 1973. Action spectrum for a low intensity, rapid photoinhibition of mechanically stimulable bioluminescence in the marine dinoflagellates *Gonyaulax catenella*, *G. acatenella*, and *G. tamarensis*. J. Cell. Physiol. **82**: 363–372. doi:10.1002/jcp.1040820306

Frank, T., S. Johnsen, and T. Cronin. 2012. Light and vision in the deep-sea benthos: II. Vision in deep-sea crustaceans. J. Exp. Biol. **215**: 3344–3353. doi:10.1242/jeb.072033

Galt, C., M. Grober, and P. Sykes. 1985. Taxonomic correlates of bioluminescence among Appendicularians (Urochordata: Larvacea). Biol. Bull. **168**: 125–134. doi:10.2307/1541178

Galt, C. P., and P. F. Sykes. 1983. Sites of bioluminescence in the appendicularians *Oikopleuradioica* and *O. labradoriensis* (Urochordata: Larvacea). Mar. Biol. **77**: 155–159. doi:10.1007/BF00396313

Herren, C. M., S. H. D. Haddock, C. Johnson, C. M. Orrico, M. A. Moline, and J. F. Case. 2005. A multi-platform bathyphotometer for fine-scale, coastal bioluminescence research. Limnol. Oceanogr. Methods **3**: 247–262. doi:10.4319/lom.2005.3.247

Herring, P. J. 1988. Copepod luminescence. Hydrobiologia **167**: 183–195. doi:10.1007/BF00026304

Herring, P. J., M. Latz, N. J. Bannister, and E. A. Widder. 1993. Bioluminescence of the poecilostomatoid copepod *Oncaea conifera*. Mar. Ecol. Prog. Ser. **94**: 297–309. doi:10.3354/meps094297

Johnsen, G., M. Candeloro, J. Berge, and M. Moline. 2014. Glowing in the dark: Discriminating patterns of bioluminescence from different taxa during the Arctic polar night. Polar Biol. **37**: 707–713. doi:10.1007/s00300-014-1471-4

Knollmüller, J., and T. A. Enßlin. 2018. Encoding prior knowledge in the structure of the likelihood. 1812.04403. doi:10.48550/arXiv.1812.04403

Knollmüller, J., and T. A. Enßlin. 2020. Metric Gaussian variational inference. 1901.11033. doi:10.48550/arXiv.1901.11033

Lapota, D. 2012. Seasonal changes of bioluminescence in photosynthetic and heterotrophic dinoflagellates at San Clemente Island. IntechOpen. 2. doi:10.5772/35341

Lapota, D., and J. R. Losee. 1984. Observations of bioluminescence in marine plankton from the Sea of Cortez. J. Exp. Mar. Biol. Ecol. **77**: 209–239. doi:10.1016/0022-0981(84)90121-7

Lapota, D., M. Geiger, A. Stiffey, D. Rosenberger, and D. Young. 1989. Correlation of planktonic bioluminescence with other oceanographic parameters from a Norwegian fjord. Mar. Ecol. Prog. Ser. **55**: 217–227. doi:10.3354/meps055217

Latz, M. I., and H. J. Jeong. 1996. Effect of red tide dinoflagellate diet on the bioluminescence of *Protoperidinium* spp. Mar. Ecol. Prog. Ser. **132**: 275–285. doi:10.3354/meps132275

Latz, M. I., J. C. Nauen, and J. Rohr. 2004. Bioluminescence response of four species of dinoflagellates to fully developed pipe flow. J. Plankton Res. **26**: 1529–1546. doi:10.1093/plankt/fbh141

Martini, S., D. Nerini, and C. Tamburini. 2014. Relation between deep bioluminescence and oceanographic variables: A statistical analysis using time-frequency decompositions. Prog. Oceanogr. **127**: 117–128. doi:10.1016/j.pocean.2014.07.003

Martini, S., and S. Haddock. 2017. Quantification of bioluminescence from the surface to the deep sea demonstrates its predominance as an ecological trait. Sci. Rep. **7**: 45750. doi:10.1038/srep45750

Martini, S., L. Kuhnz, J. Mallefet, and S. Haddock. 2019. Distribution and quantification of bioluminescence as an ecological trait in the deep sea benthos. Sci. Rep. **9**: 14654. doi:10.1038/s41598-019-50961-z

Mazzei, L., S. Marini, J. Craig, J. Aguzzi, E. Fanelli, and I. Priede. 2014. Automated video imaging system for counting deep-sea bioluminescence organisms events. 2014 ICPR Workshop on Computer Vision for Analysis of Underwater Imagery, Stockholm, Sweden. 57–64. doi:10.1109/CVAUI.2014.15

Meighen-Berger, S., L. Ruohan, and G. Wimmer. 2021. Bioluminescence modeling for deep sea experiments. arXiv. doi:10.48550/ARXIV.2103.03816

Messié, M., I. Shulman, S. Martini, and S. H. D. Haddock. 2019. Using fluorescence and bioluminescence sensors to characterize auto- and heterotrophic plankton communities. Prog. Oceanogr. **171**: 76–92. doi:10.1016/j.pocean.2018.12.010

Morin, J. G. 1983. Coastal bioluminescence: Patterns and functions. Bull. Mar. Sci. **33**: 787–817.

Priede, I., P. Bagley, S. Way, P. J. Herring, and J. Partridge. 2006. Bioluminescence in the deep sea: Free-fall lander observations in the Atlantic Ocean off Cape Verde. Deep-Sea Res. I Oceanogr. Res. Pap. **53**: 1272–1283. doi:10.1016/j.dsr.2006.05.004

M. Reinecke, T. Steininger, M. Selig 2013 NIFTy—Numerical Information Field TheorY. Available from https://gitlab.mpcdf.mpg.de/ift/NIFTy

Swift, E., J. M. Sullivan, H. P. Batchelder, J. Van Keuren, R. D. Vaillancourt, and R. R. Bidigare. 1995. Bioluminescent organisms and bioluminescence measurements in the North Atlantic Ocean near latitude 59.5°N, longitude 21°W. J. Geophys. Res. **100**: 6527–6548. doi:10.1029/94JC01870

Tamburini, C., and others. 2013. Deep-sea bioluminescence blooms after dense water formation at the ocean surface. PLoS One **8**: 1–10. doi:10.1371/journal.pone.0067523

Valiadi, M., and M. D. Iglesias-Rodriguez. 2013. Understanding bioluminescence in dinoflagellates—How far have we come? Microorganisms **1**: 3–25. doi:10.3390/microorganisms1010003

van Haren, H., and others. 2011. Acoustic and optical variations during rapid downward motion episodes in the deep north-western Mediterranean Sea. Deep Sea Res. I Oceanogr. Res. Pap. **58**: 875–884. doi:10.1016/j.dsr.2011.06.006

Widder, E. A. 2010. Bioluminescence in the ocean: Origins of biological, chemical, and ecological diversity. Science **328**: 704–708. doi:10.1126/science.1174269

Widder, E. A., S. A. Bernstein, D. F. Bracher, J. F. Case, K. R. Reisenbichler, J. J. Torres, and B. H. Robison. 1989. Bioluminescence in the Monterey Submarine Canyon: Image analysis of video recordings from a midwater submersible. Mar. Biol. **100**: 541–551. doi:10.1007/BF00394831

Acknowledgments

We thank our colleagues of the Information Field Theory Group who provided insight and expertise on the NIFTy framework and Statistical Inference. We would also like to show our gratitude to Thomas Eberl for sharing his expertise on the ANTARES experiment and the data set. SH acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 772663). The authors acknowledge the financial support of the funding agencies: Centre National de la Recherche Scientifique (CNRS), Commissariat à l'énergie atomique et aux énergies alternatives (CEA), Commission Européenne (FEDER fund and Marie Curie Program), Institut Universitaire de France (IUF), LabEx UnivEarthS (ANR-10-LABX-0023 and ANR-18-IDEX-0001), Région Île-de-France (DIM-ACAV), Région Alsace (contrat CPER), Région Provence-Alpes-Côte d'Azur, Département du Var and Ville de La Seyne-sur-Mer, France; Bundesministerium für Bildung und Forschung (BMBF), Germany; Istituto Nazionale di Fisica Nucleare (INFN), Italy; Nederlandse organisatie voor Wetenschappelijk Onderzoek (NWO), the Netherlands; Council of the President of the Russian Federation for young scientists and leading scientific schools supporting grants, Russia; Executive Unit for Financing Higher Education, Research, Development and Innovation (UEFISCDI), Romania; Ministerio de Ciencia, Innovación, Investigación y Universidades (MCIU): Programa Estatal de Generación de Conocimiento (refs. PGC2018-096663-B-C41, PGC2018-096663-A-C42, PGC2018-096663-B-C43, PGC2018-096663-B-C44) (MCIU/FEDER), Generalitat Valenciana: Prometeo (PROMETEO/2020/019), Grisolía (ref. GRISOLIA/2018/119), and GenT (refs. CIDEGENT/2018/034, CIDEGENT//2019/043, CIDEGENT//2020/049) programs, Junta de Andalucía (ref. A-FQM-053-UGR18), La Caixa Foundation (ref. LCF/BQ/IN17/11620019), EU: MSC program (ref. 101025085), Spain; Ministry of

Higher Education, Scientific Research and Professional Training, Morocco. We also acknowledge the technical support of Ifremer, AIM, and Foselev Marine for the sea operation and the CC-IN2P3 for the computing facilities. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

None declared.

Submitted 07 October 2021
Revised 13 October 2022
Accepted 11 September 2023

Associate editor: Ivona Cetinic