Plymouth, 9/07/2015

I: Interviewer (Gregor Halfmann)

R: Respondent (David Johns)

I: Good, so …

R: So about twelve full-time equivalent analysts. That may be about fourteen or fifteen who actually spend some time doing it. Most of the guys down here, when they are doing analysis, will spend half a day at most and it works that maybe forty or fifty percent of their working week is spent doing this. And then we would kind of expect them all to do about 450 samples a year. That is what we say is a full load of samples. You could physically do more but we don't think it is very good to have somebody sit on the microscope for hours and hours. So four-hour sessions and that's it.

I: I was wondering about that, how tiring it is to do this work …

R: Yeah, I don't think it is … For health and safety guidelines, you could stay here all day. And we make sure that everybody, when they first start, has regular training, that they know how to set up their microscopes, They are all designed as ergonomic as we can make. Everyone is encouraged to take regular breaks. If you see, the people are [unclear] standing up a lot anyway to put their samples out, to put their samples away, there are guidebooks around, so people walk around all the time. So it is not too … you know, being crouched over in the same position all the time. Okay, so … We will find something to actually look at. So just briefly, this is one sample. And I will probably explain how the machine actually works with two pieces of silk, one piece going to [unclear] piece.

I: Yeah.

R: So this is the covering part, so where it comes in. And this section here represents about ten nautical miles of tow. So it would be moving through the band.

I: So one of these, you would consider as a sample.

R: That is one sample, yeah. So it is ten nautical miles of the CPR going through sea water. And it filters around three cubic metres of sea water. So it represents three cubic metres of sea water. So the initial stage of analysis is actually done before this. When it is unloaded, which Rob will have shown you, it is laid out and the green colour of it is assessed, the actual greenness. They have a chart, we call it our phytoplankton colour index, and the greenness is assessed against that. And the whole idea of that is, it gives you a very broad indicator of phytoplankton biomass, just by visual assessment. And I would think that sounds like really woolly science, you know, “it is kind of dark green, it is kind of pretty green.” But it actually matches up. We have got sixty or seventy years and it actually matches up with satellite data really well. So if we say “this is a really green sample in this area,” it represents high biomass and then we would check satellite records and they recorded high chlorophyll levels as well. So it is a good initial assessment.

I: And that is without any tools?

R: Just looking at it, yeah.

I: Because this one does not look green at all …

R: This does not look green at all. Yeah, this would be a zero, it is not green at all. So the sample is now laid out for … Claire is going to be looking at it. You can see it has got a number written on there. So this is number 33 and this is 612 LR. So the LR is the route; this is the LR route, it is going across the middle of the North Sea. It is 612, so it is the 612th time this route has been towed. So it has been going for quite a few years. And the 33 is the block number. So on that particular route, we would only look at the odd number blocks. We would look at block one, block three, block five … So this is 33, it's on the far end of that particular route. So once it is laid out like this, we would do three stages of analysis. In the first stage we would be looking for phytoplankton and for the phytoplankton, we look under high power. We look in twenty areas, so we look in ten areas going down this way and ten areas going down this way. And when we look at it, we are actually looking in one field of view which is actually one mesh in the silk. I'll show you that in a minute. So it is just a sub-sample, I think it would be one fifteen-thousandth sample. And we are only looking for presence and absence of phytoplankton, we don't count because it would just take too long. So once we have done the phytoplankton, we would then put on a lower power and we would scan across what we call a traverse sample. We scan across here and then here and here, five times, and five times on what would be the equivalent when it was sandwiched together. In that we are looking for zooplankton that is under two millimetres, so quite small stuff; and we would count all of that, everything is [unclear]. And if we saw any phytoplankton when we do that and which we have not seen in our phytoplankton count, we would just say it is present. There is no attempt to count it, it is just a presence/absence. Once that section is done, the final part is the large zooplankton, where we would look at this and we would take everything off; everything that is over two millimetres, pretty much everything is scraped off. And it is put into …

I: You are doing it with these tools over there?

R: Yeah, it is these little … sometimes you can use a brush, a paint brush. Normally we would put that in … because there is lots of it, it would go into one of these which is called a Bogorov tray. So you put it in with some fluids, some PGP. And then you would try to lay out … You would just start looking down one end with the microscope again and count everything. So that is what we call the eye-count stage. The phytoplankton, traverse zooplankton, and then the eye-count zooplankton stage, that is what the main stages are. And then each sample would take … [unclear] it is not going to take very long. If it is from the North Sea, where everybody here has been looking at for years and they know what they are looking for and there is not much on it, it could take half an hour, three quarters of an hour. Sometimes we get samples from more exotic locations, like the middle of the Pacific or the South Atlantic and they could be plastered. It could take a day, a long time. We try to make people sub-sample it then. We don't really want people spending hours and hours and hours by saying “you have got to do so many samples.” We don't really want somebody spending hours on one thing.

I: So you just said it is counting, or actually looking just for presence in some cases …

R: Yeah, for the phytoplankton.

I: But you are noting the different species or taxa, right?

R: Yes, yes, so if you are talking about taxonomic entities, so not species, not genus, we probably count over 800. A lot of these will be to species, but for certain taxa, the calanoids, the copepods that people are very interested in, we would say “Not only is it genus this and species this, but it is also a male or a female or it is in a stage where it is a juvenile or it is an adult.” So that is the information that is captured as well. But only for things that people are really interested in, otherwise it just adds to the workload. … Some of the things we are looking at, we cannot take them any further. We can only say that it is a decapod larva which is the larva of a crab or a lobster. It would just take forever if you wanted to identify exactly to the species. So some things that we look at are quite coarse, we just stop and say “that is good enough, that is what we need to know.” But Claire is a decapod expert, so she could identify it. … That is one of the good things, we keep the samples that we have collected. If somebody was to come back in ten years’ time or twenty years’ time saying “you know, I am really interested in looking at decapod larvae in the North Sea”, you go to the database and you find out which samples have those decapods on them and they can just crack on and identify themselves if they want to. We have done that with fish, with baby fish, because baby fish are really, really difficult to look at, because they go into the CPR machine at 20 knots and they just brrrrr. So somebody went back and looked at forty, fifty years’ worth of [unclear]. He would pick them all off and identify them himself, at least to genus.

I: Okay. So this is how they are packed?

R: Yes, so they all come back and they will get cut up, for example that LR route that Claire has is a big long strip of silk that gets cut into these samples. Then they are assigned to people, we have a computer programme that assigns the samples. So people will get a sample but they will never get the adjacent sample, because we don't want a bias there. So Claire would have 33LR, which is the one she has done, and then she goes to get the even sample next to it as well which she just wraps up. They get wrapped up and they are put into a box in there. So all these samples here are all the ones that are waiting for Claire to do. … People would do anything with that work, it does not affect our data because we have not looked at it. So there is a lot of interest in sort of genetic studies. They quite often end up destroying the samples and we'd rather they have that one we are not really that interested in. They can destroy it. And once the samples have been done like this, they are stored over here for a short period of time because the data Claire is writing down in her notebook right now … So this is what she has seen in the phytoplankton, for example. And this is the small zooplankton and this is the large zooplankton. So we have a bespoke piece of software that we enter this into. It runs some checks on the data anyway but it actually requires somebody with experience to look at it and say “well, that doesn't look right or that doesn't look right”. So that is part of our check block procedure. Then people would get a slip of paper like this which has a sample there for Claire and basically says “are there any calanus on it?” because from the data it looks like somebody has missed it. Maybe people [unclear] had lots and this person did not find any. So that is part of our QA procedure.

I: So is this your job, to look at it with the expertise and see if there are irregularities or who is doing it?

R: No, not at the moment, that is being done by Tanya who is in there and [unclear]. It is something that I should be doing but I am not at the moment. My job has kind of changed a little bit, maybe even since you came down here last time.

I: I will ask you about your tasks later …

R: Yeah.

I: But about these instruments here … I know that the Continuous Plankton Recorder has not really changed in any way, but the instruments for analysis, I am sure there have been some upgrades.

R: As in the microscopes?

I: Yeah.

R: Yeah, they have. When I first came here this was what we were using and in fact people still use this. This was the workhorse that everybody used.

I: So what difference does it make for the work of the taxonomist between this and this?

R: Well, these are all custom made. And they are custom made so they kind of mimic these. So the field of view and the magnification is all the same. The whole idea is that you keep the methodology the same. You don't want to make any mistakes with methodology, it has got to be the same. We pride ourselves on our 70-year time series, that's what we want. But to actually use these compared to this, this is horrible. When you sit down, everything is really tight, whereas these have been designed much more ergonomically and you can rest your arm on the thing here, you can move this stage. When you are analysing, you have to physically move it around like this, whereas this one, you can see there is a little wheel there to move it along, it is minimum movement. In fact we have got a new microscope, I use this one very much, this is the next version. This one is even nicer. It is a lot more ergonomic. So we are being most careful all the time that there is no chance that we are going to mess up our methodology. So the fields of view and the magnification are all exactly the same.

I: So these are custom-made for SAHFOS?

R: Yeah, but this one, I got the feeling that it might be more of an off-the-shelf one. But it does kind of work. So let's see if we can find one that you can have a look at … This look is very topical, microplastics. One thing that we picked up … Over the last ten years or so there was more interest in microplastic pollution. So we started to ask the analysts if they see any; that is obviously a really obvious piece because it is really big, but sometimes you see very, very small pieces. So we have been counting microplastics as a possible indicator of pollution for about ten years. We have not really got our heads around exactly how much of that could be contamination. Because when these silks are laid out and cut, they are exposed to the air and all sorts of stuff could fall onto the silk. So it is quite difficult to say when it comes back and you look under the microscope, if it is actual pollution from the sea or if it is just a form of cloth. So we have got some ongoing work trying to get that. … So quite a lot of phytoplankton, if there are adverse conditions, they can exist. So they can make this capsule that is very resistant to environmental change [unclear]. And they can remain viable for years and years. And then a lot of stuff that comes in is something where we had ballast water contamination come in, because they are really difficult to [unclear]. They don't care if they stay there for years. So prior to that they were probably in a big bloom of that particular species and then maybe food has run out or light level has changed and then it is time to protect their cells and sink.

I: Okay, thank you. ... So how often do you have to take a book and … ?

R: For most of the phytoplankton in that North Sea area or anywhere near the UK, most analysts here would not need to look that up. ... Yeah, some rare things and there are a couple of things that look fairly similar, for some of the zooplankton, people would look at that more often, even after years.

I: But you keep references here …

R: Yeah, there are all those references there and I have all these notes in here as well. So I got the phytoplankton ones at the front. These are quite good, these are coccolithophores, so they are calcareous and they are tiny, really really tiny, ten microns. When you look down there, you can see a sample and there is a silk mesh. The gap in the mesh is about 300 microns across. So it is quite a big gap. And we had people saying that there is no way that we can see coccolithophores, they said “no, it is going to go straight through your mesh because they are only ten microns.” But they do stay there, so we took photos and we published some of it and say “actually, we can see these.”

I: Why do they not go through?

R: Probably because of the other phytoplankton that are there, so you get spiny bits sticking out. The mesh is made of silk, which is spinous as well, so they get stuck to that. A lot of plankton produces [unclear] as well, so they will be trapped in the [unclear]. So if you were to count them abundance-wise we would not see very many, but if you towed it through a bloom of coccolithophores, you can pick up on them by satellite, they are just massive. If we towed trough a bloom, we would still see quite a lot. We would not see billions like there are there, we would undercount massively, but we would still see enough to get us to say “that is a bloom.” So you kind of ignore the values and just look at the trend. So yeah, everybody has got their own stuff like this that helps them identify. This is probably like twenty-five years old and [unclear]. … Let's just stick this under the microscope. Yeah, that's quite nice, that's a decapod, it's quite easy to see.

I: Oh, wow.

R: It is quite easy to see that that is a crab. So that is what we call a [unclear] state. So that is the final state before it settles down to the bottom and carries on. And then the one to the side of it is a slightly younger [unclear]. That is a stage it would go through before it becomes this one.

I: Okay.

R: You can kind of make it out, there is like a spine leading off and you can see a little tail.

I: It looks quite different though.

R: Quite different, yeah. And these are big. We got all the guide books here for people to look at. We have a reference collection over here as well, so when somebody finds a particular nice example of a plankton, of a zooplankton, they can chuck them in here and as you can see, actually that is a cephalopod, so that would be a baby squid. You can just see its eyes and tentacles.

I: So this is just zooplankton in the collection?

R: Just zooplankton, yeah, we don't keep the phytoplankton, it is just too small and it goes … it's a bit too difficult. If you put it in here, we would never have enough. We would not be able to get it off … The MBA did have a culture collection where they had lots of phytoplankton, just type cells that they just kept growing but that's closed now. So you could go and get a tube of those and have a look at them. But they don't do that any more. Yeah, so I don't think there is anything else to say about the analysis.

I: I am not sure if I have any more questions that we have to address here. I am wondering, because this is really manual work, basically all of it …

R: Yeah.

I: Have there ever been considerations to automate it? Or is that even possible?

R: No, it's … There has been a lot of work done on that sort of automated identification of the plankton. So there are projects going on at the moment to do that with using various software routines, Zooscan, I think, is the main. But that relies on a sample … not this type of sample. You would have to remove everything off. You have to take everything off it and then you would run it underneath the camera and then it takes images. The images are really, really good, but that system needs somebody to identify from the images. So it can't identify from images. The ones they are using that have automated identification are incredibly coarse. They are really, really coarse. They would probably be able to pick out that decapod. You can see how big it is and they might be able to tell you it was a decapod. They probably could. But a lot of the stuff that we look at is maybe three millimetres long. And to tell them apart you are looking at one of its swimming legs which is tens of microns across. You would need to look at it and manipulate it and say “actually, I can see these spines and therefore it is species A.” So at the moment the software just can't cope with it. It can do very broad … maybe it can give you an estimate of biovolume, but it won't do this.

I: And it hasn't been tried here at SAHFOS?

R: No, at SAHFOS we have connections to people who were doing it. We keep our eye on it. We had a guy who was looking at it years ago and we have done some work with Plymouth University because they are doing it. I think there is a guy at the university running it, but at the moment it is not a viable alternative to this method. You still need the expert input here. What happened recently is that funding is very tight. Somebody will come up with that system and really push it to say “this is the system and we can do everything with it.” They will say that ours is old-fashioned, it is time-consuming, and it is expensive, and their system is the way forward. But actually, when you compare the data, their work is so coarse, you can't actually get anything from it. If you wanted to know just pure, super coarse biomass, in fact not even biomass, but biovolume, that's fine. It's perfect and it will give it to you really fast. But if you wanted to know anything else about the ecosystem, it is pretty useless. I am sure at some stage there will be better automated routines. They can produce really nice images now and they can produce really, really fast images. They can run this sampler, they can run the stuff through the sample and they can come back from the ship and they got this massive database of photos. That is the first part that is done. It is just then, how do you get the software to identify everything?

I: So this is the main lab of SAHFOS, right?

R: Yeah.

I: Do you have any more rooms for further analysis or any other …

R: Ah …

I: Is this the main workplace?

R: It is for this because this is kind of our main work at SAHFOS. This is what we would call our routine CPR data. This is it. There is another lab which a couple of our researchers use for more like physiology stuff. I have no idea about that. I know where the lab is because I am responsible for health and safety, so I have checked it, but that's it. I have no idea what they are doing in there. I don't really think that fits with what we are doing, so I [unclear]. But we also have got another researcher, Rowena, who works on the genetics side of stuff. But she works in a lab of the MBA and again, I don't know anything about it.

I: So as a lab manager, this is your …

R: Yeah, this is my bit, yeah. We did have another taxonomy lab, but they have moved that and they don't have that any more.

I: Okay, so I guess all the other questions we can discuss upstairs. I think I have seen enough …

R: Yeah, you can come back down, but there is not much else to see in here … But we got this fancy microscope, that's very nice. I think we almost kind of found that and forgot why they have it. I am not even certain what that was used for. And this microscope over here is for doing phytoplankton work, but not sort of routine samples. You would need a separate sample and settle the sample. So you get a phytoplankton sample and they have this settling chamber and then the phytoplankton would settle on the bottom and then it looks upwards. A couple of people have done some specific training to become not exactly accredited, but proved for that type of phytoplankton work. So if we happen to get any external consultancy work on phytoplankton, they have done the training on this. It has run out of [unclear]. So that is useful. We have the facilities, but on a general basis it doesn't get used.

I: This is also an older one right?

R: Yeah, this is not a very good place. [unclear]. It is a lot bigger now than it used to be when I first started. The lab was probably around here and this was all office space. This was done a few years ago, it was modernised.

I: Actually I was wondering, because there were lots of papers on the CPR Survey in 2003 about the history and all that. I wondered what happened since then. So has SAHFOS still expanded since then, or?

R: Yeah, it has expanded, the number of routes has gone up and the area we cover. What we would call our core business is here. Most of our funding comes from here, NERC and Defra, Canada and the US and Norway. They fund basically most of these routes. This is our bread and butter and gets done all the time. For the last maybe ten years, we have had stuff in the North Pacific, we had regular routes there. A lot of it came out of the funding of the oil tanker that went down. I forgot the name, but it was that fund.

I: I think I read about that.

R: Yeah it was the one that was really bad and lost masses of oil, so that is a part of their responsibility; they say “here is funding for marine research.” These are all done on a monthly basis. Every month, a route will go across here. This one is not quite so, it is maybe three or four times a year.

I: And these samples also come back here?

R: Always. These also come back here. And these ones, as you can imagine, is a completely different set of animals and some of the stuff over here, they are more interested in … So for example, there is a particular type of copepod called calanus. And there is calanus finmarchicus and calanus helgolandicus. They are the stuff that everybody is interested in. If you google them, or actually any work on zooplankton, it is all about those two. Because they are the food for larvae fish. They are so important for the food chain. And they are also difficult to tell apart. You could not use a machine to tell those two apart. One of them is a cold water and one of them is a warm water one. So it is a good indicator. If we notice that the cold water one has shifted this way and the warm water one has shifted, it is very interesting. But we really need to say whether it is calanus finmarchicus or calanus helgolandicus. That's all I need to [unclear]. Whereas over here, there is a neocalanus, which fills kind of the same niche. But they want to know if they are adult females, adult males, if their stage is … They are a bit like an insect, they go through stages like an [unclear]. They want to know what stage they are, if they are juveniles. So there is a lot more information on these samples and they take a lot longer to do, a lot longer …

I: So why do they want to know all this additional information?

R: Those organisms are just so much more important, I believe. That ecosystem is quite different. Here you have a spring bloom pretty much every year, maybe in March, and it is due to the water starting to warm up and the light levels increasing. So the phytoplankton just explode, millions of them, and then you get the zooplankton and they graze it down. Then occasionally in autumn, the water gets stirred up again and you get another bloom. So that happens every year. Over here, this doesn't happen. We don't have a spring bloom, it is just a constant low number of phytoplankton and you get peaky zooplankton in there. So that is what they are interested in: They want to know if they are juvenile or adults, so they want to know what stage of development everything is at. I don't do any research on that stuff so … And we also have got tows around here, around Falkland and the South Atlantic. They are very [unclear], they take even longer, because there is a lot of zooplankton and again, people there are interested in euphausiids, which look like shrimp. Euphausiids there are kind of the main thing higher trophic levels eat, the whales eat them and the penguins even. So everybody is interested in those to identify them, and if we can, to species, which is so difficult. You would be looking for a tiny spine, it is just a pain in the arse, it really is. So yeah … These bits here are how we keep track of what everybody is doing. So everybody has an analyst number. [unclear]. And then for each month it says how many have been allocated and how many have been analysed and what their backlog is. So we keep track of what everybody's backlog of work is. And we have a separate thing, which is a timetable that we work with. For example, the April samples should all have been finished two weeks ago. So there are a couple outstanding, but if we don't have that timetable to say when everything is done it is just …

I: So who is deciding who is doing what?

R: As in who gets what samples?

I: Yeah.

R: We know who is due to get samples and then it goes into the programme that we use and it just [unclear] them out. We keep on eye on it, so I might say “actually, they have quite a lot of samples here. Why did they get the samples? Is It because they have been too busy or is it [unclear] to the lab?” And then we would say that there is no point in giving them any this week. So yeah, it is a fun part [laughing]. What have you been doing the last two weeks? You haven't done any samples …

I: Okay, good. I think that's all down here, thank you.

R: So that is where I am moving in to. So I am going to lose my nice office upstairs …

I: So your position is changing?

R: Yeah, when did you come down?

I: In May.

R: May. So I have been lab manager for a while. Actually I still look at the samples and I manage the whole process, with the workload, the data, etc. I do some research as well but there was an aspect of the analysis management that the person who was downstairs in the office was kind of doing but also kind of not doing. So there was a bit of a grey area about who was in charge of it. But the person is sort of semi-retiring from October. So it was a good idea then to say “okay, we are not clear who is doing what. I am going down there,” although I didn't really want to go down there, and I take on the aspects that were unclear before.

I: Okay.

R: So the bit that I was kind of responsible for was making sure that all the analysis work was done, the workload would be healthy, an adequate workload, etc. But once the samples are being physically analysed, there is a whole quality assurance procedure and that was not managed as closely. So we were kind of falling down a bit. So [unclear] all the samples analysed to our deadlines and the other deadlines are set a bit … wishy-washy, you know what I mean?

I: Oh yeah, this office is much nicer.

R: Yeah, it is much nicer. It makes sense for me being downstairs because there has been this confusion as to who is in charge of what and people would see the person down by the lab and would go inside and say “Can I do this?” and the person would say “yeah, yeah.” Then I find out six weeks later and I am like “What? You shouldn't be doing that.” … Although it is a nice view, you can see from the window, it leaks and they can't change the window because it is listed. It's a listed building. And the wind comes in, give it an hour, two hours and it will be about forty degrees in here. I usually never have the blinds up. I always have the blinds down, I don't really look at it that much.

I: Alright. So I have quite a lot of questions, actually.

R: That's alright.

I: I mainly came to talk about SAHFOS, but in the end we may have time to talk about NMBAQC.

R: NMBAQC, okay.

I: Okay, you kind of addressed already the first question that I had. What was your role as the laboratory manager and your tasks? What has been taking the majority of your time, basically?

R: Well, at the moment, in that role, we are trying to get ISO accreditation. And we are going for ISO 17025. And that has taken a lot of time. We have always had set procedures to do everything, set standards and set procedures, but everything needs to be … I don't know if you have looked at any ISO stuff, it's boring, everything has to be done exactly the same way. Okay, we have always done this procedure the same way, but bits of it have been written down here, bits of it have been written down there. Everything needs to be put in a proper structure. This is the procedure and this procedure links to this procedure and this procedure links to a technical manual and yeah that at the moment is taking …

I: But that's a one-time …

R: Yeah, it will be a one-time, but after setting it, it will be ongoing because you have a lot more robust quality assurance and quality control programme for that internal audit. So that is going to take more time going on. I suppose from what takes time now, a lot of it are probably staff issues. It would be trying to make sure that everybody does their job, you know, I don't have any complaints about it, but there will always be somebody who will be moaning about one thing, or they may be making mistakes in one area of their work, or they are getting pulled away from doing their analysis. People should be down there for forty percent of their time and you find that they have gotten slightly more interested in a project and they just drift … So that takes quite a lot of …

I: You said that you were also doing a bit of research yourself …

R: Yeah, my research tends to be a lot of collaborative stuff. So another area of my job is to … I got a lot of those bits … If anybody wants CPR data, externally, they would contact me. Pretty much in the first case, it would be me. Then we would have a discussion about what they want, what they are going to use it for, what they can do with it, where it is from. Then nine times out of ten, I would then get my colleague to [unclear]. But he doesn’t have a biological background, he is an IT guy, to him it is just numbers. So there is no point in going to him originally because he just wouldn’t be able to answer any questions. So what tends to happen a lot is that I will get into discussions with people, I go meet them or they come down here, and then decide “actually, this makes more sense to do it as a collaborative piece of work.” So they want to use the CPR data, they will do some stuff with CPR data, and they ask me “Does this make sense?” or “What do you think about this?” And that is how my research kind of evolves. I have not had time to sit down and just lead my own research for ten years or so.

I: So the procedure is pretty much that you get contacted via email?

R: Yeah, or phoned up.

I: Then what are these discussions typically about? What expertise can you provide and help them with?

R: So quite often people would say … I am thinking of a general one, “I am interested in basking sharks” or “basking sharks around the UK”. This is one that came up a little while ago. “We have not seen many basking sharks in the last two years, there might not be any around.” Basking sharks preferably eat plankton, so they phoned me up via skype, two guys, one in London and one in Cornwall and they said “Right, what is the plankton telling us?” So I would look at the plankton and say “You know what, I looked at this.” First it was kind of anecdotal from what I would have known from this year, because everybody has done their analysing of samples now and it goes into this sort of database, but that won’t be available for [unclear] to look at for a while. But I have been down there, I have looked at the samples, so I could say “You know what, I have been doing this for twenty years. To me, it looks really odd this year.” So the first conversation was “I think it looks quite strange.” Then I would pull the data out and start going through what I would think would be the driving factor for maybe why there are no basking sharks. There is none of this species here, which is really unusual. I have looked at the last fifty years’ worth and it is not there. So I suppose that is …

I: Okay, so it is really about the scientific question?

R: Yeah, yeah. And sometimes people would ask and say “I want to do this type of work with this species in this area” and I might say “Actually, that is maybe not the appropriate species to use in that area” or I might say “Actually, not that I am a hundred percent confident, but that species might be a bit undercounted.” You know, some things are quite difficult to see and I might say “Actually, you should use something else instead.”

I: I read in the most recent annual report of SAHFOS and in one of the bits that you have written it says that the CPR data is used a lot in “blue-sky research”. So can you talk a little bit about this kind of research and perhaps explain why the CPR data might be particularly useful for blue-sky research? If that is the case …

R: Yeah, it is the case. That is not something that I do now, really. But there are a few guys here, who would do blue-sky research, or what I call blue-sky research, but maybe it is not the right term. A lot of the data goes into research that is kind of policy-driven, so there is already an identified output for it, you are feeding in that type of work. So “We need to know about the health of the ocean, let’s look at plankton.” But there are some other guys who are interested in species X’s physiology. So they have got it in a lab and they are looking at its physiology, its thermo-preferences, etc. So then they would say “You find this, you have got seventy years’ worth of data and the whole spatial information.” So that is quite often what it is very useful for. It puts a lot of those fine-scale bits of research into some sort of spatial context, because nobody else has plankton data over such a wide area. I think that is probably why it is most useful. It is the spatial aspect and the temporal; it has been running for so many years.

I: I have one more question about the ways to access the data. You said that people contact you, but I have seen on many other databases or projects that there is on the internet an interactive map or a browser, something like that.

R: Yeah.

I: Are there any plans to do something like that?

R: Yeah [laughs]. That is being worked on at the moment. That should have been done a little while ago. We have only got one guy left in IT, two have left and we are struggling to find a new person. So that is going to be on hold. One of the issues we have kind of had is: We are happy to give away some of our data for free, but … I sort of mentioned it downstairs when I was saying who funds the routes. So we have funding coming in from NERC, from Defra, from lots of different people. So we are not one hundred percent funded from one group. So if you know that you are going to get all your money every year from NERC regardless of what you do, which some organisations are like, they are very happy for you to have that data. But if you are more dependent on your funding coming from you writing research products that use your data or science that uses your data, as another funding stream, you don’t want to give it away, really. It becomes quite difficult. It has been an ongoing thing since I have been here: that we want people to use our data, but on the other hand we are not going to say “Here, have everything.” Because what tends to happen then has happened: We would say “Here you go, you want all our data. Have all our data.” Then next thing we know is that person is then using that data in a huge EU-funded project where they get all the money and we don’t get anything. The way we are funded, we can’t operate like that.

I: So if there was like an interactive access, you would not be able to get all the data?

R: No, no. You can group up the data into maybe functional types. So, you have got a nice dataset to look at and it would tell you quite a lot, but it wouldn’t tell you the super detail that you could get out of the data if you had everything. For copepods which are one of the main things the fish and basking sharks eat, everyone is interested in them, we could say “Here is an index for large ones which are over two millimetres and here is an index for small ones.” Then people would say “Ah, that is really interesting because if I want to do modelling and look at biomass that fits with my modelling really nicely.” But if you wanted to talk about long-term changes in the community, it wouldn’t give you the information you needed. You would need it at a species level. So that would be kind of like our hook to say “Okay, you can have this coarse data for free. Do what you can, but if you want anything more involved, you need to come here and speak with us and either do collaborative research or you involve us in your project and give us some funding.”

I: We just talked about your expertise that goes into the projects when somebody asks for data. Do you think about that going away if the process was purely automated?

R: Yeah, and it would. Because I think it is the same with any dataset. You can write so much down in the metadata to cover everything, can’t you? You could list every species we look at and write something about every single one, but it still won’t give you that kind of expert knowledge, I think. That’s why we would always say to people “Come here and speak to us.” There have been occasions where people have had data and they got hold of it through one person or another and then we have seen stuff luckily not quite get to publication stage and they said “Look, I have noticed all these changes” and then I said “Well, if you had spoken to me I could have told you that there was an issue.”

I: Okay, so we have talked about access to the data now. I have one more question about your background. You said you have been here for quite a long time.

R: Yeah.

I: Did you also start as a plankton taxonomist?

R: Yeah.

I: And you are still also doing it sometimes?

R: Yeah. So it started off as a part-time job just doing analysis, I think it was 18 hours a week. So I started doing it then, yeah. Now, I don’t spend as much time. I don’t get given as many samples. We call it a load. So a full load is maybe 450 samples a year and I think I am on about three quarters which will go down to half soon, because I just don’t have the time to do it.

I: You already told me about how many taxonomists there are. But I met Rob on Friday and he is doing the taxonomist work, he is also a technician, and he is curating the archive. Is that usual, that the staff have different tasks or jobs?

R: Yes, well … If you speak to the guys in the operations, they only do that. There are a lot of aspects to it, but they only do the operations. Then when you go to the analysts, for the people involved in analysis, a lot of the time the analysis takes forty or fifty percent of their time. So there are only a couple of people who spend all their time on analysis-associated extra tasks. So I don’t know if you saw the girl who came in she [unclear] spends all her time only on analysis-related tasks. She spends a lot of time physically looking at samples and then the rest of the time is doing the sort of quality check. So she looks what everybody else has done and says “That’s not right and that’s not right.” Plus, she is involved in cutting the samples and distributing the samples. But only a couple of them are like that. Most of them would do forty or fifty per cent of their week in analysis and then the rest of it is filled up with something else. It will either be the research side, or in Rob’s case the instrumentation side. There is one person down there, the rest of her time is kind of public engagement work and starting to go down the policy route. So yeah it is quite normal.

I: I read about the training of the taxonomists in the annual report, too. And if I understand correctly it takes about two or three months to get a basic training, and there is a test in the end …

R: Yeah, kind of.

I: Okay, can you explain that a little bit? Because I am also wondering if you have learning material or if it is mostly learning by practising? How does it work?

R: So, it probably takes a few, maybe three months is a generous time, for somebody to go through and say they have looked at the three stages of analysis, the phytoplankton, the small zooplankton, and the large zooplankton. During that time, they will have one to one training for most of the time with one person who leads the training. But they could ask any of the analysts because most people have been here for years. They will be given a set of notes. There is sort of standard set of notes that people have got but most people make their own notes as well, as you can imagine. And then there are all the resources downstairs. Then after a period of time, they will be given a sample that somebody else has already looked at, so we have got the results for that. They would then do that sample again, they would analyse that sample and then the two results can be compared. Then after they have done maybe ten or so, if our senior analyst down there is happy with it, they will be given samples to start doing them on their own. But that takes a while, it would certainly be four to six months. I know it says that it takes two or three months for training, but I would say it is probably six months before they get their first sample to do on their own, unless they come here and they have got a lot of experience, which does not tend to happen that much. When we advertise for an analysis role, we advertise for some marine science background or whatever, but plankton is quite a small field of people and we very rarely get anybody applying who has got any knowledge of plankton. So we would kind of look … because I am sort of responsible for when we take new people on. I would be looking for people who I think would fit in with the team, who would be happy to do that job. You really want them to do their job for three or four years at least. So you kind of think “Are they going to fit in? Are they going to stay? Is that going to be their line of work?” And to have some sort of knowledge of natural history, to be interested in natural history, is something I think is very important. Before I was doing that, we had somebody else who was taking people on. They just wanted people who would stay and treat it like a factory and do their work. So that is great from a work perspective but it means that they have got no interest in the biology or the ecology of what they are looking at. So they don’t tell you if they see anything different. They don’t think “Actually, you know what, that is really weird at this time of the year.” So you kind of lose that. So I am more like … One of the last ones we took on, she was interested in bats, I think it was. [unclear] So she actually has got a license to look at bats, and I was thinking “That’s what we want.” You want somebody who is interested in nature, to say “Actually, this is quite unusual.” Otherwise it becomes a bit of a conveyor belt of churning out samples.

I: So how many taxonomists do you have in training right now?

R: At the moment just one. Yeah, just one, she is the last person we took on. She has not taken that much time to train because she has got a background in euphausiids. So she has done a lot of plankton. In fact, we had only people applying recently, who had a background in plankton anyway. In her case, it was a bit strange because she applied for a junior position that is not very well paid, as a plankton analyst, part-time. And she has got a PhD, a big publication record, all sorts of pieces, and normally we would have stayed clear, because if you take on somebody who is very research-focused, experience has shown us, they come in and do the job for six months and then get bored stiff and they are off to do something else, they would not stay at SAHFOS. But then you lost them, you put all that training in and you lost them. But she wanted to, I think she has got a family, she wants to stay put and that’s what she wants. She doesn’t want to do the research side any more, so she is in training at the moment. In fact, I think she is probably doing her first samples now, so.

I: Can you tell me perhaps how much of the skill of a taxonomist can be learned by text books and how much is experience or informal training?

R: [Laughs] Probably most of it is informal and on-the-job stuff, yeah. Textbooks are obviously really useful, but it is not the same as looking down and actually seeing a physical specimen there. One of the things, that probably was not as obvious from looking at the CPR, is that the stuff is squashed. It is squashed between the plankton [sic] and it is squashed because it is being hit at twenty knots. So it is very, very flat. So if you look at a textbook and you look at that … They do look quite different, so you need to manipulate the organism. What we are trying to do as well is we try to get samples from around here. Guys who work in this building have got a regular sampling point right out there, the Eddystone. And they use a net to catch their zooplankton. So their zooplankton looks lovely. They are not squashed, so we also give people that to look at as well. So we can say “Here is the book, this is what it looks like in the book. This is what a real one looks like, it is nice. And this is what ours look like.” So you get a nice broad understanding. You can start to identify things at weird squished angles which you would not pick out from a book.

I: You already talked a bit about the quality assurance and quality control. Perhaps you can tell me a bit more about that. So, do you keep any statistics, I mean there are probably several different mechanisms to …

R: Yeah, so I showed you that slip of paper which is what we call a check block. Everybody has a record of the samples they have seen in the year. And [unclear], I don’t think she is in at the moment, she keeps a record of … They go through the samples, they look at the samples and they look for things that they think are a mistake. So they would look through it and say “This is a mistake, this is no mistake.” And they would issue a check block to a person. But you don’t do it yourself, it gets given to somebody else. So they would issue the check block against my work and that person checks it. Then when it comes back, it is either that I was right and they were wrong to give the check block or I missed something, or I have miscounted it, or it is an acceptable error. So there is a whole load of criteria and then they keep statistics of that throughout the year; we are meant to get it every six months, but … So at the end of the year you can say “Actually, you were given so many check blocks against you and you were wrong on these and these are the things that you are making mistakes with.” And then that is how you guide into a training. We had a guy here who was simply not taking enough stuff off. So he was just looking at the sample and he would pick off what he thought he needed to take off. But there was a whole load of other stuff that he was missing. So after a couple of months you would say “Look, we have looked in your statistics and you are missing this stuff all the time.” But on the other hand, what it can do is you can say “Okay, we have given this person loads of check blocks,” because we basically don’t believe what they are saying. And we get the results back and all the results are fine. So that is then how you start to think about the person who is giving the check blocks. They have got a bias [unclear] as well. It’s their bias against the person for whatever reason …

I: And you mentioned that you are not giving consecutive samples to one person, so that is also …

R: Yeah, that is another sort of check.

I: Are there any more like simple things you can do to make sure the quality …

R: Those are probably the main. Those are the main checks, they actually go through the samples. What we are going to be doing now, it is part of the ISO, is that each person will have one of their samples totally re-analysed every month, which we were not doing before. That is another step we are required to put in. It is kind of a requirement for ISO. So it means that they will just take a sample that I have done and they will just completely redo it. So that person who does it will be one of our senior people. And then it means that you actually are going to get less of the routine work out because they will have to go back over another sample. What we are in the process of doing now is really tying down … Okay, I have looked at a sample, you can imagine that for parts of it I have sub-sampled. So I sub-sample it, they look at it, and they are not going to look at the same sub-sample that I looked at because they almost certainly are going to look slightly different. So what do you then say is correct? So, you are just trying to thrash that out. If it is vastly different, it is wrong, if it is not that different, you will ignore it because it is sub-sampling …

I: You could not put a marker or something …

R: You probably could do, yeah you might be able to do it. But even in the case of you looking at the sample, and you wrapping it up, and you put formalin on it to preserve it, so it has got fluid on it, and then it is wrapped, and then it is folded, and it is put in a box … For the stuff where they are supposed to take everything off, for the eye count, the large stuff, you should expect that to be almost the same, shouldn’t you? There are going to be a few that might be different, but the part that is sub-sampled is just proving to be difficult to tie down.

I: I don’t know much about what the ISO is and what it means. What would change when you have that?

R: The main thing is that when you want to get external, contractual work, it is kind of a requirement. If we wanted to … I don’t know … For example, the environment agency takes water samples around the whole UK and they look at phytoplankton. And the main reason for that is for being an indicator of water quality and if they have certain harmful plankton in there, they would shut the oyster farms there, or whatever. Now if you wanted that work, you would have to have an ISO accreditation. They just would not give it to you otherwise. So because funding is so tight at the moment, you know, we are trying to identify other funding streams. And one of them could be doing that sort of statutory monitoring. So that is one side. The other side is something called the Marine Strategy Framework Directive which is a new big marine policy that covers the whole EU. And that has to have indicators that you produce. A country has to say that it is monitoring their seas to say they are healthy. And some of those indicators were probably better served if you can say that your dataset has come from an ISO accredited. It might not be necessary, but if in a few years’ time, the UK decides to use an indicator to measure how healthy our sea is, there is a whole [unclear] of indicators. And you say, for example “From that indicator it shows, actually, the sea is not very healthy.” So you are in breach of that EU law. You go into infringement procedures, so then you have to go back and say “Okay, this is our data. We have got confidence in our data. It is not the data that is wrong and telling us this story. This is really what is happening in the sea.” So if you have got ISO accreditation on your dataset, then you are kind of building a case. It is not fun doing it [laughs].

I: How long does it actually take until data are sort of finished? Like until a sample is really finished and it goes into the final database?

R: Okay, it could be done quicker than it is done, put it that way. At the moment we have just … Pretty much all the samples from this April are finished. So they have been looked at, they are all analysed, they are all waiting there. But what takes a while is somebody doing the quality check, going through. And then it has to go finalised into the database. Ideally, I reckon we can get to five months. So from the sample being towed to it being in the database, maybe five months. We could possibly get a little bit quicker than that. It is probably around seven or eight maybe. Because there is always something, somebody is away, or there is some link in the chain that does not quite work. Part of the way of managing it now, the way I have sort of managed it, is that we have got to a certain point … So yeah, we are working on eight months at the moment. So from when a sample comes in to when it is in the database, it is eight months, although a lot of time it works out being seven months. Nobody is asking us really for the data any faster than that, so I kind of got it to that stage and we just stick to that level. It would take quite a lot of more work to get it under six months. We could do it but unless our main customers or funders like NERC or Defra say “We want it quicker,” I am not going to push it.

I: I am wondering, mainly the limiting factor to work more or faster is the funding, right?

R: Yeah …

I: I mean you could do more probably?

R: Yeah, we could do more samples. We set a limit of 450, but that, to be honest, people could do more. That is a comfortable number for most people to be able to do fairly easily without me cracking a whip down there.

I: So can you tell me a little bit about the outlook in terms of funding for SAHFOS? Is it like secured? You said there are a lot of different sources.

R: Ahm …

I: I read about when SAHFOS was shut down in the 80s and then it turned into a charity. So is it like safe or what would you say?

R: It’s … There is a shortfall in funding, which is looking fairly bleak. I think we are okay for a couple of years, but if we don’t start getting some big projects in then it is going to look pretty bad. But saying that, we have really pushed ourselves with Defra into the sort of policy needs of the UK and the EU. So we are saying “Okay, that MSFD, the Marine Strategy Framework Directive, is being looked at on a national level and on EU level and one of my colleagues, probably a lot more than me, I have had a little bit to do with it, but she has actually left now and has gone to Plymouth University, we have really pushed to say “Actually, our dataset is perfect for this, it is ideal.” So it is kind of written into it. The UK will have to fund it because it is being used to monitor the health of the sea. So funding from Defra should be reasonably secure. And I think we have gone into … Now they have changed because there are different ways of funding, I think we are in marine evidence now, or something, which was a good move for us in Defra. So it means they kind of have to do it. And in NERC we are under what they call national capabilities, national capabilities that they think are of prime importance for UK marine science. So that side of it is quite good. They will probably ramp our funding down, I should mention. Canada is always a bit … We get money from Canada and they get some data from us but they are really odd. We give them the data, but … That’s from DFO, the Department of Fisheries and Oceans, their scientists are not really encouraged to do anything with that data. They sort of publish lots of internal reports, they don’t really publish to the outside, so you don’t know. That data just disappears, it will go into internal reports. And we get funding from the States and that is kind of on a rolling maybe four or five year pace. So we have got a guy on our council who sort of sourced that out and he is very pro CPR. So we have done quite well there. So funding … yeah, it is okay-ish. We need some big grants coming in.

I: Would you say that it is difficult to get under these umbrellas with a technology that basically cannot change?

R: [Laughs] Yes, but … the actual CPR, you know the towed CPR, and we have our plankton data which is the stuff that we [unclear]. But there are other things, Rob probably mentioned it. There is other stuff that you can stick on there. So there is … Did he talk with you about the water sampler?

I: Yeah, a little bit, yeah.

R: Yeah, so the water sampler which would take discrete samples, which they can do all sorts of DNA tests on. There is a whole load of new instrumentation they are putting on there. So as long as we can get at least around the table in these meetings, I would never say to the people “Look, you either want CPR data or nothing else.” I would say “Look, we have got huge logistical knowledge here. We have got guys who are towing stuff all around the world, they know how to get those ships and things on those ships. There is that logistics side. We have got people who know how to use large datasets, we have got people who know how to do this.” So you are kind of selling the expertise of SAHFOS, not just the method, that old method. But then that puts a lot of stuff into context. People will say “Okay, we are going to do a new project, we are going to do a cruise, and we are going to do such and such and such in an area.” And then [unclear] will say “Oh, can we just tow a CPR and we can see what you guys see and then we can compare it to what we have got.” So we do that quite a lot. So it is an old method, but it’s …

I: It is expandable …

R: Yeah, yeah, and it is everything that goes with it.

I: And what are the considerations when you think about putting on new instruments or testing new things. Who is deciding on that and how do you even become aware of the new technologies?

R: So we have got a guy [unclear] who is in charge of that area. He is in charge of the instrumentation. So they would think about what they want; part of it is him going and talking to people about what we want and also being in the right places where people say what they want. But there is a whole aspect of, you know, if you stick something on a CPR, what is it going to do to the CPR? So at the moment, there is actually a boat out today that is towing a CPR with a pitch and roll sensor on it. So they bolted some new piece of instrumentation on it and they have a pitch and roll sensor on it to make sure that it is not weighing it down, so it is flying like this or going to the side. So that sort of aspect is taken into consideration. So if they have a new piece of kit, how is it going to affect the flight of the CPR? But a lot of times it is just people going into the meetings and saying “Actually we are towing through that area, do you want to bolt something on it?” And as long as they have got some money, then …

I: Is any of these technologies or instruments developed here or does it all come from external?

R: Yeah, all the stuff comes from external. I think the guy would like to do that, maybe try and get ideally a PhD student for something, so we can say “This is actually kind of what we want.” But you would need an electronics lab for that, wouldn’t you? A lot of that comes from the NOC in Southampton, although we just got some … I am not certain how much input he has had, but we have got seal tags now that are going to be fitted to the CPR. Seal tags, they are like, as soon as the CPR breaks the water it can beam back temperature data. So it will take temperature data as soon as it hits the water. As soon as it goes into the water it starts recording temperature and then it gets towed and as soon as it comes out of the water, I don’t know what it does, it connects with a mobile phone signal and beams the data straight back. So that is not a normal seal tag. I think that is being chained and I think he has had some input into how that is done. Apparently he has got a 3D printer now.

I: Okay, I think that covers about everything of my CPR questions. So maybe we can talk a little bit about the NMBAQC. So you are the chair of it … How much work is it?

R: It’s in fits and starts, that is.

I: What exactly is your role as the chair? What are the things you have to care about?

R: Well, how long have I been there? I’ve been there for a couple of years. The very first thing that dragged on for ages was a dispute between contractors. So I took on the role, the guy who used to do it has been doing it for years, ten, fifteen years, and he was too busy, but he still retained, he is still on the sort of board. So the very first thing is there were two competing contractors and they hate each other. I think one was supposed to have stolen staff from the other one. So that was the very first thing and that just dragged on for years. So it would be quiet for three months and then I would get an email from one of them saying “Well, the guy over there is not doing this properly.” And then it was just a case of how am I going to diplomatically say some of these things and I would normally have to speak to the guy who sort of … So I am the chair of that group. My group answers to HBDSEG, the Healthy and Biologically Diverse Seas Evidence Group which is responsible for the Marine Strategy Framework Directive in the UK. So the chair of that group is kind of who I would answer to. He works for the Environment Agency, so I would always contact him and say “I have got these guys and …” And then it is composed of these components which go out to tender every couple of years. So I suppose a couple of months last year were taken up looking up the tenders. So the tenders came in, they actually came in from the two contractors, they were the only ones who applied. They hate each other, so that was really difficult to manage. There is going to be some more. And then we are trying to develop a new component on zooplankton, which we are leading here at SAHFOS. So that takes up some bits of time.

I: Yeah, I was reading about the zooplankton standards. Why is it that there has never been a standard for zooplankton?

R: It is because it has never gone into any sort of statutory monitoring. So if you wanted to look at the health of the sea there, for public health, it would pretty much all be phytoplankton, because a lot of the phytoplankton [unclear] to produce nasty stuff. What they found recently, although I don’t know if they ever do it like that, I don’t know if you read it in the research, they found cholerae, vibrio cholerae kind of sticks to the kiting on zooplankton. So they think now that the zooplankton tends to be the source or that there is a population of cholerae on it. So maybe in the future you could say “actually, we should monitor zooplankton levels because that quite often is a natural source of cholerae.”

I: How long does it take to establish standards like that?

R: It can take ages. Well, it depends on how many people are doing that work. One has been going on for years, it is what they call equibiota which is the stuff that lives on the sea floor. And that is … JNCC has just produced a best practice guide on it. That has taken … I have been doing it for three years and probably about five years’ work beforehand because there are lots of people involved in that epibenthic work. They go out and do a certain … You can imagine, if you go out and put a wind farm in somewhere, you need to know what is on the bottom. So there are lots of people who are doing that type of survey work and they all got their own methods. So they will all say “This is adequate. This is fine for what I want to do.” And then somebody else would say “Actually, that is really crap. You can’t tell anything from that. You should do it this way.” So … The zooplankton would not take that long. There is all kinds of best practice stuff written already.

I: I thought about that. The NMBAQC is probably not the only scheme or set of standards.

R: No.

I: What is its specific role? Or does it have a particular role?

R: It is the body that, in the UK, advises on best practice for marine biology data. If you want to do any statutory monitoring for anything, they call them competent monitoring authorities, or anything for the EA, they will insist that you are part of the NMBAQC group. You have to take part in it because it offers external quality assurance. If you go down the ISO route, you have to have some of your samples looked at externally. So NMBAQC offers that for the people who are doing competent monitoring. For example, if you are doing phytoplankton monitoring around here for shellfish beds, the EA would want to know if you are actually taking part in the NMBAQC phytoplankton part. It is just a way of saying “Actually, what you are doing, is kind of up to standard.”

I: So it has become mandatory at some point, right?

R: Yeah. Well, actually, what it says is “Take part in NBMAQC or such scheme.” And there is only one in the UK.

I: Okay. Do you have any feedback from scientists who take part and do this? What is their reaction to do this? Is it like “Oh, this again …”, or …

R: Yeah, it is a funny one, because most of it goes into statutory monitoring which is kind of removed from site research scientists. So most of the people, who are doing it, they do it, they follow that system and they pass their data off. And I found that very difficult, because … Well, no, the phytoplankton is a bit complicated. The phytoplankton component is actually called BEQUALM and it is almost like a separate thing. So BEQUALM, everybody who takes part in that is involved in phytoplankton monitoring. So I went along to their meeting in Denmark last year. And all the guys there are fantastic in identifying phytoplankton. “How much do you know about it?” “Nothing.” “Who looks at your data?” “No idea.” They were just so far removed, a lot of them are like that. They do it, it is statutory monitoring. “Here you go, here is my dataset. I send it to [unclear] over there and then I don’t know what happens to it.” So it is quite odd. When we did our zooplankton workshop here a couple of months ago there were similar sorts of people. They were looking at their zooplankton, “oh yeah, it is really interesting” and they identified all these things. “Oh yeah, that is really useful.” “So who is going to look at this data and what can you tell me?” “Nothing.” [laughs]

I: That’s interesting.

R: Yeah.

I: Maybe coming towards the end, you can be a bit more general. NMBAQC is about the quality of marine biological data.

R: Yeah.

I: I don’t know if it is possible, but can you perhaps generalise a little bit what actually determines the quality of data or what can be most threatening to marine biological data?

R: It is inconsistency. Inconsistencies that creep in. I suppose the people I have talked about that go and do their statutory monitoring and then pass the data on is one part of it. They all tend to be following a very specific set of guidelines, so you are kind of confident in the quality of their data. They always have their little instructions and they do exactly the same thing every time. Where it got a bit difficult is where we mentioned this type of work to the research community, to the research scientists. For example, our zooplankton work was brought up by one of my colleagues at ICES, the International Council for the Exploration of the Sea. She is in a working group on zooplankton ecology. She brought it up about the standards of people’s work and the quality of what they were producing and they did not like that at all. They took it as being offensive. They have done this all their life and they know exactly what they are doing: “You should be confident in my work. I am an expert in this work. Who are you to say …” But what we found out is that they have not been following the same procedure. They have not been following that strict procedure all the time. So then you automatically start to think “Well, the quality of your data is a bit suspect because what if you changed your analyst? What if you changed your microscope? What if you changed a slight method? What about the chemical preservation? Have you got any records of any of that sort of stuff?” “No.” That instantly makes me think it is a bit suspect. Obviously, you have heard of Helgoland. The Helgoland Roads dataset has been collecting plankton for years and years and years. It is a great dataset, but they had a case where they had not been keeping some records. I think it is Karen Wiltshire who is the head over there by the way, she showed us some very interesting stuff, great step changes in the data. It just turns out it was a new analyst. So unless you have kept all those records and you have your standard procedures, it is …

I: Is perhaps one effect of doing these quality assurance things to make researchers aware of such …

R: Yeah, I suppose so. It seems to be a real novel concept to some people. In the phytoplankton community, a lot of them know that anyway. But some of the other guys are like “I never thought about this.” And as far as I am concerned, if you are then publishing a paper and you have some sort of background, you have more sort of confidence in your data. When your reviewer might say “Well, your data is a bit suspect”, you can say “Well no, we have got a standard methodology, etc.”

I: Can you say whether the quality of marine biological data or the usability of data has improved since NMBAQC or similar …

R: I would say so. I think it is a lot more robust now. People are a lot more careful with keeping decent metadata records. Some other sort of stuff that has been improved is the epibiota. They used to send a camera down just to record the seabed. I saw some of the early video footage and it is appalling. It is really, really bad. Whereas now they would send down a video camera with a sledge with lasers pointing forward, so the laser dots are exactly ten centimetres apart. So when you look at the video footage you have got a size reference. When you look at it now, you think “That’s really good.” You get a lot of information from this and when you look at the initial stuff from a few years ago and it is just … People thought it was okay but it does not give you that level of information. And it is in people’s benefit, I think, because if you are happy with the quality of the data, you can start answering more questions.

I: My final question is also a bit broader. Do you see the status or relevance or the attention to data having changed over the past years or decades, since you have been doing marine biology? I mean within the scientific community, is there more attention being paid to these issues?

R: Yes, I would say so, certainly from the policy side of work which is kind of the area that I know about. If you are talking about pure research, there probably is as well. People are a lot more rigorous, so you would probably find papers that were published in the 60s, for example, on lots of different types of marine data that probably wouldn’t get anywhere now, because they wouldn’t have considered various things. Imagine the processing power that you have got now, you can actually [unclear] … From the policy side point of view, yes. I have talked about the MFSD, there is another one that is called the water framework directive, which is more coastal. We don’t sample in close coastal water. All those sort of things have driven people to say “We really need to know that this is proper, decent-quality data”, because if the EU or somebody comes to us and says “Your water quality is really low”, they need to know that the data is reliable. But from the research side, yeah I think it is so hard to publish with marine data, I presume, probably I should say with any data, that people are a lot more rigorous, I would say. We have got guys here with a lot more stats background. I did a degree in fisheries science, so quite [unclear], I would say. So I have kind of learned on the job the stuff that I am doing. But the guys who have been trained over the past few years in a PhD are very much focused on using Matlab and R and the routines. I would be looking and saying “Oh, that is really interesting, look at this trend in the data!” And they go “yeah, but look at the sampling effort, you cannot say anything about this.” So it is a lot more rigorous.

I: So that also comes with development and new technology …

R: Yeah, definitely. You can just crunch the data so much more. If you look at some of the original CPR papers from the 60s, they are pretty much hand-drawn distribution maps.

I: So do people go back and re-analyse these papers?

R: Yeah, definitely, they looked at stuff and said “okay, what does this mean? Can we do this?” There was a paper published, maybe twenty years ago about a type of zooplankton and how it correlated with a big atmospheric driver in the North Atlantic. So if you plotted it like a regression, it fitted it exactly. It was so closely correlated. Then at a certain point the relationship broke down and they were like “Oh this is really interesting, the relationship of the last couple of years has broken down and I think it is because rising sea surface temperature. This particular thing is a cold-water organism and it is moved away. So there is no link any more.” Great. So a guy contacted us a few years ago and said “I want to redo that to see what is happening.” And he could not get it to match. And that is because the guy who did it originally, I actually contacted him, had no idea how he did it. [laughs] So he produced this paper that has been published and it is not reproducible. And I am sure that is the case with so many things.

I: But by now there are probably also higher standards for publishing those things, right?

R: Yeah, so we get picked up … And again I am [unclear]. Part of why we are doing all the quality and standard procedures is to counteract that argument. I would say at SAHFOS, certainly over the past ten years, maybe more, we have been very open about what you can use our data for and what you can’t use it for. We did have a previous boss, who said “it can solve everything for you. It will solve world hunger and everything, it is an amazing tool.” He was very zealous about the work, so I think we had a bit of a battle after that when people would say “It is no good for this, it is no good for that.” And we would say “yeah, we know”. So now we always say what you can do with it and what you can’t do with it.

(end of recording)