Acquisition Directorate

Research & Development Center

Report No: CG-D-15-13

Automated Protist Analysis of Complex Samples: Recent Investigations Using Motion and Thresholding

January 2012

Homeland Security

N O T I C E

This document is disseminated under the sponsorship of the Department of Homeland Security in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

The United States Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of this report.

This report does not constitute a standard, specification, or regulation.

Timothy R. Girton
Technical Director
United States Coast Guard
Research & Development Center
1 Chelsea Street
New London, CT 06320

Technical Report Documentation Page

1. Report No	2. Government Accession Number	3. Recipient's Catalog No.
CG-D-15-13		

4. Title and Subtitle	5. Report Date
Automated Protist Analysis of Complex Samples: Recent Investigations Using Motion and Thresholding	January 2012
	6. Performing Organization Code Project No. 4101
7. Author(s) Nelson, Bruce; First, Matthew; Riley, Scott; Robbins-Wamsley, Stephanie; Herring, Penny; and Drake, Lisa	8. Performing Report No. UDI # 1331
9. Performing Organization Name and Address Naval Research Laboratory Center for Corrosion Science Bldg F14, Trumbo Annex Key West, FL 33040 U.S. Coast Guard Research and Development Center 1 Chelsea Street New London, CT 06320	10. Work Unit No. (TRAIS)
	11. Contract or Grant No. MIPR HSCG32-08-X-R00015
12. Sponsoring Organization Name and Address U.S. Department of Homeland Security U.S. Coast Guard Office of Operating & Environmental Standards Washington, DC 20593-0001	13. Type of Report & Period Covered Final
	14. Sponsoring Agency Code Commandant (CG-5224) USCG Headquarters Washington, DC 20593-0001

15. Supplementary Notes

The R&D Center's technical point of contact is Penny Herring, 860-271-2868, email: penny.r.herring@uscg.mil

16. Abstract (MAXIMUM 200 WORDS)

Ballast water management systems must be tested relative to the U.S. Coast Guard's proposed ballast water discharge standard before they can be approved for routine shipboard use. Analysis for live organisms in treated samples is time-limited and labor intensive. Currently, skilled staff must use a microscope to observe, count, and verify the viability of live organisms ≥ 10 µm and < 50 µm in minimum dimension. This report provides technical details on the results obtained when applying automation algorithms developed by Naval Research Laboratory to complex samples using the Protocols for Automated Protist Analysis. This work has shown that the number of viable organisms that are detected in complex samples depends on the thresholds used to display and analyze data, that organism size cannot be uniquely determined in these samples without significant effort, and that motility is essential for determining organism viability in this size class when using automated methods.

This report's findings are the result of continued research into the potential to automate protist analyses and provide the most current information available at the conclusion of this initial research effort. Additional automation research efforts are ongoing and may affect details reported herein.

17. Key Words	18. Distribution Statement
Automated analysis, protist, motion detection, ballast water treatment, pattern recognition, vital stains, epifluorescence	Distribution Statement A: Approved for public release; distribution is unlimited.

19. Security Class (This Report)	20. Security Class (This Page)	21. No of Pages	22. Price
UNCLAS//Public	UNCLAS/Public	26	

(This page intentionally left blank.)

EXECUTIVE SUMMARY

Ballast water is a known pathway for the introduction of aquatic nuisance species (ANS). In an effort to reduce the number of introductions of ANS into United States waters, the U.S. Coast Guard (USCG) has proposed regulations requiring ships to meet stringent discharge standards. The proposed concentration for organisms ≥ 10 microns (μm) to < 50 μm in minimum dimension (nominally protists) is less than 10 living organisms per milliliter of ballast water. This concentration level may later be reduced (i.e., a more-stringent standard, Phase II, may be enacted) if USCG determines ballast water management systems (BWMSs) can meet the more stringent standard and doing so would be practicable. Before the USCG can approve BWMSs for routine use aboard ships, the BWMS vendor must demonstrate they are capable of meeting a given discharge standard.

A protocol for testing BWMSs at full scale has been developed by the Environmental Protection Agency's Environmental Technology Verification Program in cooperation with the USCG. Among other metrics, testing requires evaluating treated samples to determine the number of living organisms in the 10 to 50 micron (≥ 10 μm to < 50 μm) size class. Current manual methods require skilled personnel using microscopes to observe, enumerate, and determine the viability of organisms in concentrated samples before mortality is shown to occur from the artificial conditions of holding samples. This visual analysis is labor intensive, requires skilled personnel, is subject to operator fatigue, and provides no archive of results. The USCG Research and Development Center therefore sought a method to automate analyses of organisms in the ≥ 10 μm to < 50 μm size class.

Previously conducted Naval Research Laboratory (NRL) research efforts have shown that pattern recognition algorithms could be applied to sequential photomicrographs of organisms ≥ 50 μm to identify motion and, therefore, the viability of organisms. These motility algorithms have also proven to be applicable for the analysis of organisms ≥ 10 μm to < 50 μm samples.

This report provides descriptions of research and development that was conducted in 2010 and 2011 in support of refining the Protocols for Protist Automation. The emphasis of this year's work was in the evaluation of complex samples, such as those associated with BWMS evaluations. This research has resulted in an improved method for holding ≥ 10 μm to < 50 μm samples under the microscope and has also demonstrated the need for two microscope modalities and motility algorithms when working with complex samples in this size class. Work was also performed to explore the impact of sample complexity on the ability to accurately enumerate viable organisms and determine their size. Additional work with "heat killed" samples has shown that many objects observed in samples with significant epifluorescent signals cannot be attributed to viable "non-motile" organisms in complex samples.

Some of the more significant results of the current research effort are that the number of viable organisms that are detected in complex samples depends on the thresholds used to display and analyze the image data, that organism size cannot be uniquely determined in these samples without significant effort, and that motility is essential for determining organism viability in this size class when using automated methods.

The findings in this report are the result of continued research into the potential to automate analyses of organisms ≥ 10 μm to < 50 μm and provide the most current information available at the conclusion of this research effort. Additional automation research efforts are ongoing and may affect details reported herein.

(This page intentionally left blank.)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ACRONYMS, ABBREVIATIONS, AND SYMBOLS

ANS	Aquatic nuisance species
BWMS	Ballast water management system
CMFDA	5-Chloromethylfluorescein diacetate
FDA	Fluorescein diacetate
LUT	Look Up Table
mL	Milliliter
mm	Millimeter
NRL	Naval Research Laboratory
NRL KW	Naval Research Laboratory in Key West, Florida
s	second
SR	Sedgwick Rafter (counting chamber)
U.S.	United States
°C	degrees Centigrade
µm	Micrometer (micron)
<	Less than
>	Greater than
≤	Less than or equal to
≥	Greater than or equal to

(This page intentionally left blank.)

1 BACKGROUND AND INTRODUCTION

The protocols for automation of analyses for organisms ≥10 µm (microns) to < 50 µm are less mature than those for organisms ≥50 µm. The majority of research performed by NRL during 2009 and 2010 focused on the optimization of the protocols for organisms ≥50 µm (Nelson, et.al. 2010b) and determining which elements of these protocols may be applicable to automated analysis for organisms ≥ 10 µm to < 50 µm (Nelson, et.al. 2010a). This year's effort focused on analysis of complex samples that are representative of those that are encountered during ballast water management system (BWMS) evaluations (including tank filling operations, control tank drain operations, and test tank drain operations) (Nelson, et.al. 2011a and Nelson, et.al. 2011b). Significant effort has been made in establishing the properties of typical samples from each of these types of operations and in generating and analyzing representative samples of each of these types. This in turn has allowed for refinements to the protocol and automation algorithms. As more realistic samples were analyzed, the protocols and automation algorithms have been and continue to be adapted, refined, and modified to make for easier preparation of the samples and execution of the protocols as well as improved automation algorithm results.

The majority of the protocol for organisms ≥50 µm automation is directly applicable to an automation protocol for organisms ≥10 µm to < 50 µm. However, the difference in the size of these organisms requires different microscope magnifications and methods used to hold samples during the collection of image sets. Optimizing the method used to hold the sample under the microscope has been an area of focus in this year's research and development efforts. Work to further optimize and finalize the configuration that will be used in support the protocols for organisms ≥10 µm to < 50 µm is ongoing.

Two major elements of the protocol for organisms ≥50 µm needed to be evaluated in order to determine their applicability to the protocols for organisms ≥10 µm to < 50 µm. First, the need for operating the microscope in both epifluorescence and brightfield modes when working with organisms ≥10 µm to < 50 µm needed to be established. If all viable organisms could be identified in the epifluorescence image sets, there would be no need to collect both image modalities. Second, the need for using motility algorithms as a means for detecting viable organisms ≥10 µm to < 50 µm needed to be established. If all viable organisms produced a measurable fluorescence signal, then there would be no need to collect time resolved image sets and to apply motility algorithms to these image sets.

This report is meant to compliment the Protocols for Protist Automation (Nelson, et.al. 2011b) and provides technical details of the research efforts conducted in 2011 using these protocols and the automation algorithms described in that report to collect and analyze complex protist samples.

The next section of this report (Section 2.0) provides a recommended method for holding samples under the microscope and discusses the need for two microscope modalities and motility algorithms when working with organisms in the ≥ 10 µm to < 50 µm size class.

The complexity of samples makes the accurate enumeration of organisms difficult. The algorithms initially work independently on both the brightfield and epifluorescent imagery by analyzing 10 successive image pairs from each microscope modality collected twenty seconds apart. That is, the first image pair, collected at time 1 second (s), is compared to the image pair collected at time 20 s, and the second image pair, collected at time 2 s, is compared to the image pair collected at time 21 s, and so on, until 10 comparisons are made. In the case of motile organisms, this approach should result in 20 discrete signals for each organism in each microscope modality's output "motility" image, which is a binary image showing the motion paths of motile organisms. In both microscope modalities, many organisms are only observed

occasionally as they move through the sample. This can result from debris obscuring organisms (brightfield) and from weak signatures (both brightfield and epifluorescent) associated with certain organisms. In the case of weak signatures, it has been observed that decreasing the thresholds used in motility algorithms can result in an increased number of observed organisms. These issues make the accurate enumeration of organisms using automated methods an area of ongoing research and development.

Another important observation is related to determining the size of observed motile organisms in complex samples. In samples that have a significant number of motile organisms, it may not be possible to precisely determine the size of all of the observed organisms. First, during automated analyses, the microscope focus is fixed. Consequently, the apparent size of organisms can change as they move through the water column. This is especially problematic with organisms in the ≥ 10 µm to < 50 µm size class as these organisms are significantly smaller than the depth of the water column. This same problem occurs when using manual microscopic analyses, where size can only be accurately determined by focusing on every observed organism. This may not be feasible in reasonable observation times for a complex sample with many motile organisms. Second, the size of organisms in the epifluorescent microscope modality is impacted by the thresholds used to observe the data. Because many organisms are uniquely detected using this microscope modality, determining the size of these organisms may not be possible without brightfield corroboration.

Section 3.0 of this report provides a discussion on sample complexity and its impact on enumerating and determining the size of detected organisms.

Lastly, following the observation of a number of objects with strong epifluorescent signals in samples that were not motile, research was conducted to determine if the observed signals were the result of non-motile organisms or other sample artifacts. Work was conducted with a number of samples that were first heat treated (i.e., to kill living organisms) and then stained using the methods described in the Protocols for Protist Automation (2011b). This work has shown that there are a number of objects that produce epifluorescent signals at approximately the same signal strengths as those associated with viable organisms. This work has raised concerns that it may not be possible to properly assess samples in the ≥ 10 µm to < 50 µm size class without using motility using automated methods. A section of this report describes the work that was performed with "heat-killed" samples.

The final section of this report will provide the status of the automated analysis efforts for organisms ≥ 10 µm to < 50 µm and provides recommendations for performing additional research and development directed at finalizing both the automation algorithms and protocols.

2 SAMPLE IMAGING AND ANALYSIS

2.1 Unruled Sedgwick Rafter Counting Chambers

Standard sized unruled Sedgwick Rafter (SR) counting chambers are recommended as the primary method for holding samples of the ≥ 10 µm to < 50 µm size class under the microscope. This recommendation is made primarily because of the sample distortion and occlusion issues found in sample well trays that have been described the Protocols for Protist Automation (2011b). Initially, using custom-made, low-volume SR counting chambers was also considered for this application. This approach would have allowed the entire microscopic field of view to be viewed simultaneously, ensuring that organisms could not enter or leave the field of view. However, in support of this program, significant work was performed with complex samples in the ≥ 10 µm to < 50 µm size class, and this work has shown that organisms can be tracked and enumerated as they enter and exit the microscope's field of view. Based on this work, it was determined

that the cost for developing and utilizing a custom counting chamber configuration would not provide sufficient benefits to justify their added costs. Counting errors that are introduced through allowing organisms to enter and exit the microscope's field of view are not significant compared to other error sources and can be accounted for by tracking these organisms.

For this application, the microscope's objective lens and optical zoom are adjusted such that the SR counting chamber area associated with 20 microliter (µL) of sample is in the microscope's field of view. Up to 50 (but more likely 20 to 30) separate areas of this size can be analyzed sequentially by adjusting the microscope's automated X-Y stage to bring different 20 µL sample volumes into the microscope's field of view. In this way, successive samples can be evaluated in a similar manner as they would be evaluated using a multi-well tray.

A comparison that shows the superior data quality obtained using unruled SR counting chambers versus sample well trays is provided in the Protocols for Protist Automation (2011b). The high quality image data provided in this report provides additional corroboration of the use of unruled SR counting chambers as the preferred method to hold samples in the ≥ 10 µm to < 50 µm size class under the microscope.

2.2 Microscope Modalities

This report sub-section provides a summary of the results obtained with initial complex samples of organisms in the ≥ 10 µm to < 50 µm size class. The emphasis of this work was to determine the need for using two microscope modalities when evaluating samples in this size class. In addition, the large dynamic range associated with the epifluorescent signals produced by stained organisms is also initially discussed in this report section. This work has shown that many organisms without a visually ***observable*** epifluorescence signal produce a ***measureable*** epifluorescence signal. These weak signals can only be observed by utilizing a Look Up Table (LUT) or applying a threshold specifically designed to show extremely small epifluorescent intensity levels (looking at values between 0 and 1 % of an image set's maximum epifluorescent intensity values).

The data provided in Figures 1, 2, 3, 4, 5 and 6 have all been generated from the same protist sample. This sample was collected during fill operations in support of a simulated treatment experiment.

Figure 1 provides a brightfield (top) and an epifluorescence (bottom) image collected from a sample using an unruled SR counting chamber. This type of chamber consistently generated high-quality brightfield image sets that are not spatially distorted and are free from occlusions. This lack of spatial distortion and occlusions is the major reason that NRL has recommended this type of counting chamber be utilized for holding samples under the microscope. A review of the two images shows that one viable organism (a worm) circled in red in the brightfield image data did not produce an epifluorescence signal. Viable organisms that are only observed in the brightfield image data help support the requirement for two microscope modalities in the automation protocol. Other data that also support this conclusion are provided later in this report.

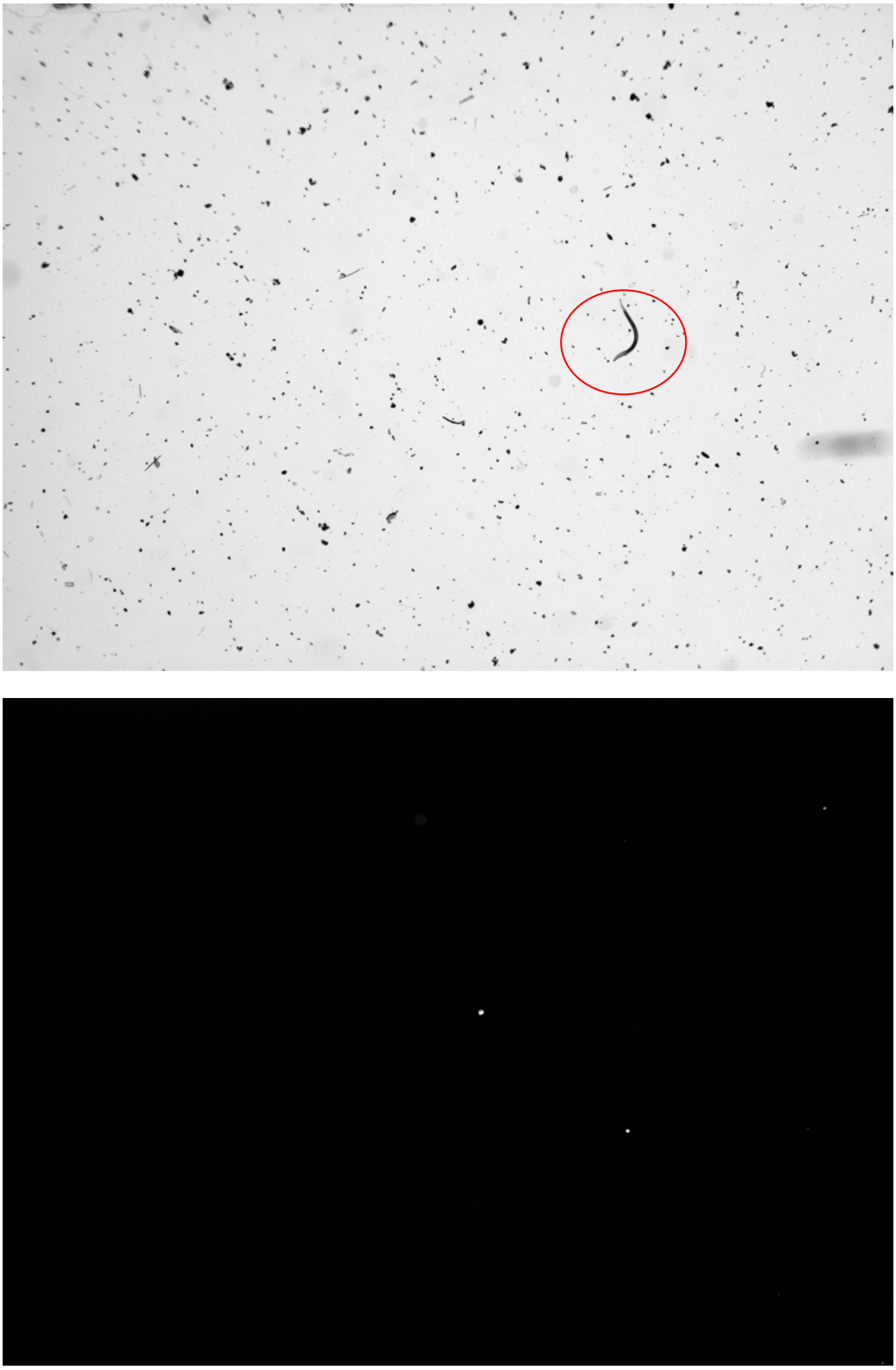

Figure 1. Brightfield (upper) and epifluorescence (lower) images collected from a sample using a simulated low volume Sedgwick Rafter Cell. Note that the worm in the brightfield image (red circle) did not produce an epifluorescence signal.

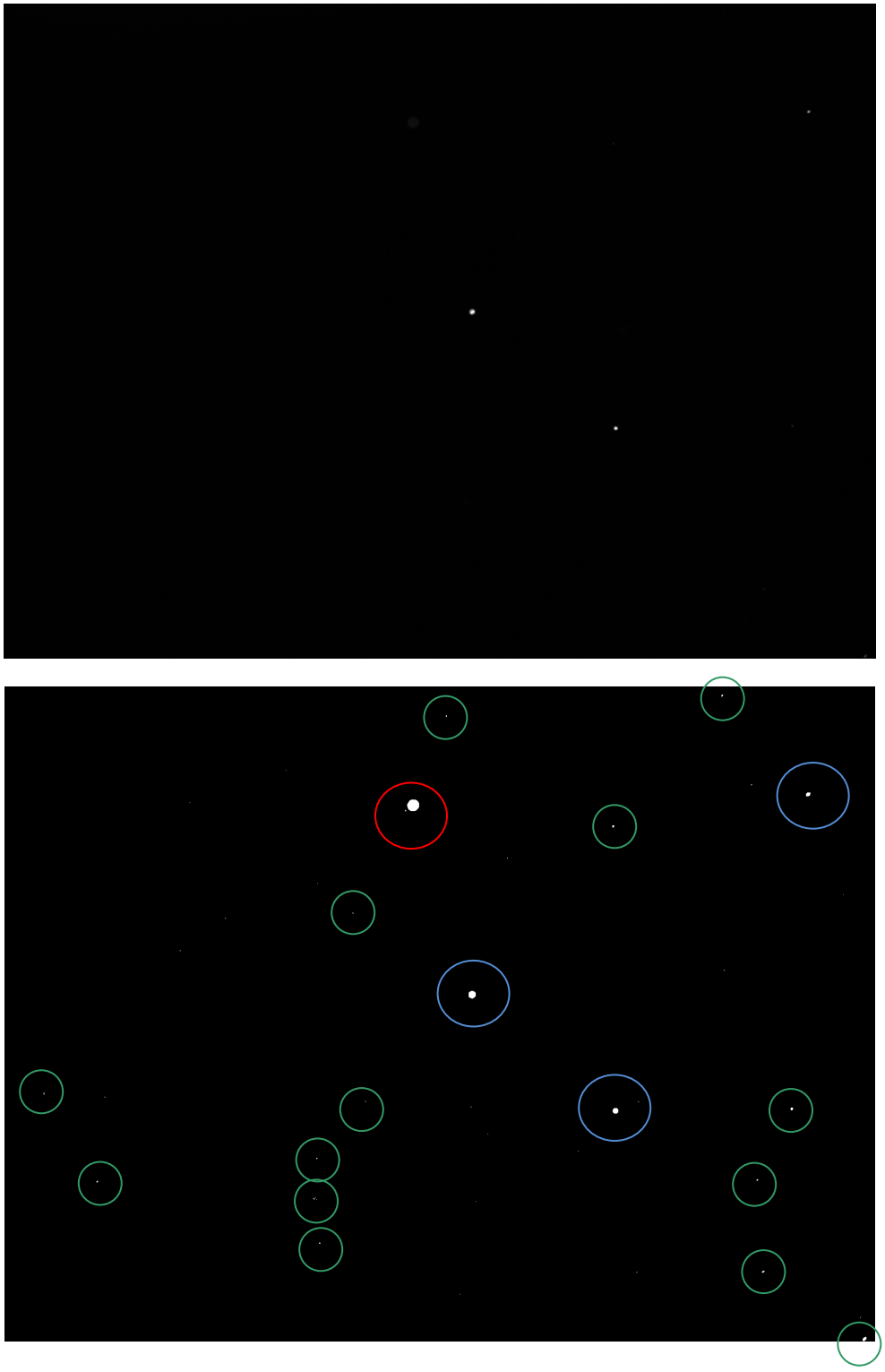

Figure 2. Threshold effects on epifluorescence images. Applying a threshold to an epifluorescence image (top) reveals many objects with a measureable epifluorescence signal (bottom) that cannot be directly observed in the epifluorescence image data. Blue circles identify organisms seen in original image. Green circles identify some of the larger objects with epifluorescence that can now be observed in this sample. It is important to note that there are many objects with epifluorescent signals that are not circled in this image.

In order to determine which organisms (or particles) in an epifluorescence image produce a measurable epifluorescence signal, a threshold is applied to epifluorescence image to create a binary image. The "on" pixels in the binary image are those that have intensity values above this threshold. Figure 2 provides two images: the image on the top is the same epifluorescence image that was shown on the bottom in Figure 1, and the image on the bottom shows the results of applying a threshold to the image on the top. A review of these two images shows that only three epifluorescence signals are readily observable in the top image while multiple fluorescent signals above the applied threshold value can be observed in the image on the bottom. The three signatures that can be observed directly in the epifluorescence image are circled in blue in the image on the bottom of Figure 2. None of the other signatures can be directly observed in the original epifluorescence image data. This is primarily the result of the large signal dynamic range associated with images from this microscope modality. The signature associated with an organism that is out of focus (due to water column height and depth of focus issues) is circled in red; other objects that produced signals greater than the threshold are circled in green in the lower image in Figure 2.

The data provided in Figure 2 show that many measurable epifluorescence signals cannot be directly observed in the epifluorescence image data. In this case, at least 14 (as there are some small signals that are not circled) out of the 17 objects with measurable epifluorescence signals could not be observed directly. Similar differences in the number of organisms with measurable epifluorescence versus observable epifluorescence signals have been seen in the majority of the data for organisms ≥ 10 µm to < 50 µm that have been analyzed to date. A similar increase was also observed in images sets collected from samples containing organisms ≥50 µm.

2.3 Motility

The initial work performed with complex samples of organisms in the ≥ 10 µm to < 50 µm size class was also directed at determining the need for motility algorithms when analyzing samples in this size class. The large dynamic range associated with the epifluorescent image data also impacts the results of motility analyses and is also discussed in this report section. The data provided in this report sub-section show that many of the objects that are observed with weak epifluorescent signals are produced by motile viable organisms.

The algorithms used to detect motile organisms in complex samples are fairly sophisticated. The algorithms initially work independently on both the epifluorescent and brightfield imagery. These algorithms analyze 10 successive image pairs from each microscope modality that are collected twenty seconds apart. For each image pair, binary images are generated from each of the two images that make up the image pair. These binary images are next subtracted and then squared (to make all values positive). This subtraction and squaring operation ideally results in two objects for every organism that moved (and no objects for particles that did not move) in the 20 sec interval that separates the collection of the first and second images associated with a given image pair. Summing the results across the ten image pairs creates a "motility" output image for a given microscope modality that tracks the positions of the motile organisms across the 10 image pairs analyzed. Ideally, 20 discrete signals for each motile organism in each microscope modality's output "motility" image should be generated. Additionally the results across the two microscope modalities can be combined to create a single "motility" output image. This image should ideally have 40 discrete signals for each motile organism.

Figure 3 shows the automated algorithm's motility image output for the brightfield microscope data. The motility algorithm shows five organisms (and their paths) that were motile based on its analysis of the brightfield image set (note that one organism is depicted by only a few pixels). The two organisms that are circled in red were viable motile organisms that were only detected in the brightfield image data. Note also that the two objects that are circled in blue, which are also depicted by only a few pixels, are corroborated as being produced by a motile viable organism by the epifluorescent microscope modality's motility image output as shown in Figures 4.

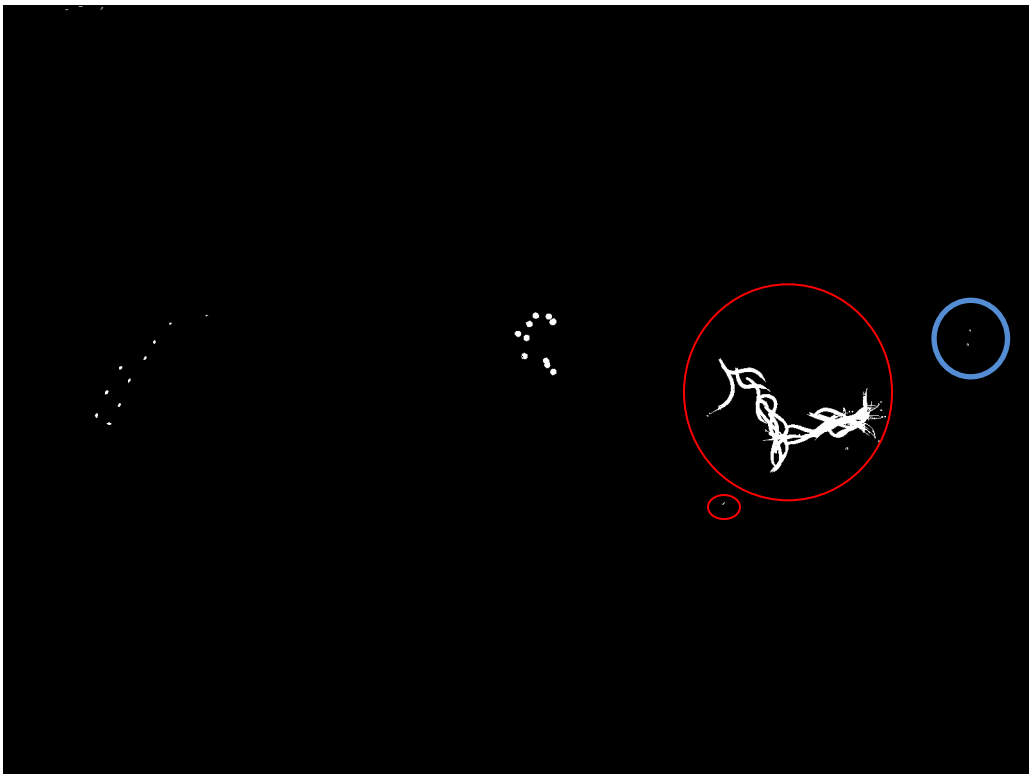

Figure 3. Motility algorithms applied to the brightfield image data detected five motile organisms. Note that two of these organisms were not detected in the epifluorescence image sets. These organisms are circled in red. Analysis of the epifluorescent data provided in Figure 4 corroborated that the two objects circled in blue were produced by a viable motile organism.

Figure 4 provides the automated algorithm's motility image for the epifluorescence microscope data for the same sample. The motility algorithm output image shows eight or nine (depending on interpretation) organisms and their paths that were detected in this epifluorescence image set. Again, note that some organisms are depicted by a single object in the binary image. In this case, five or six (depending on interpretation) organisms are uniquely detected by the motility algorithm in the epifluorescence image sets. These organisms are circled in red. One (or two) of these organisms appears to be in the same vicinity as motile worm observed in the brightfield image data.

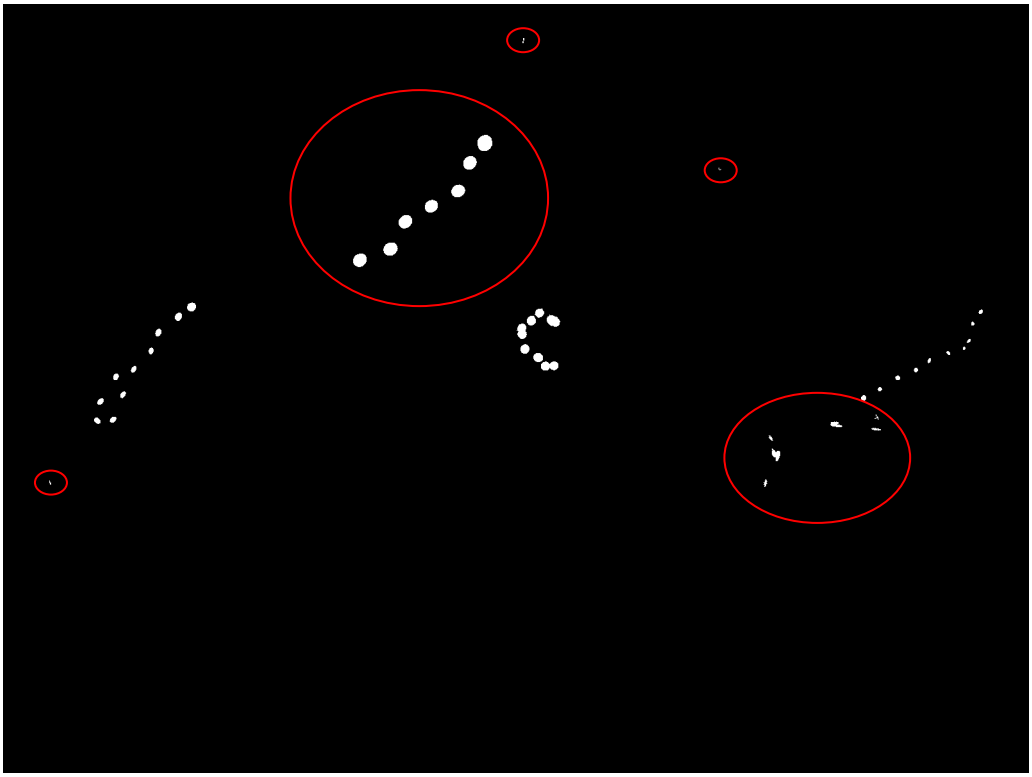

Figure 4. Motility algorithms applied to the epifluorescence image data detected a total of eight or nine motile viable organisms. Five or six of these organisms that are uniquely detected in these epifluorescent data are circled in red.

Figure 5 provides the algorithm's combined brightfield and epifluorescence motility image for the same sample. This image shows motile organisms that were detected in either the brightfield (Figure 3) or the epifluorescence (Figure 4) image sets. Analyses of images such as those shown in Figure 5 are used to spatially correlate the results obtained from the motility algorithm outputs in both the brightfield and epifluorescence image sets. It can also be used to demonstrate that the organisms that were circled in blue in Figure 3 are also associated with an organism path that was more clearly delineated in the epifluorescent image set. A review of the organism path circled in green in this image also shows the difference in apparent organism size that can be observed between organisms simultaneously detected in epifluorescent (larger) and brightfield (smaller) image data. This is discussed further in the next section of this report.

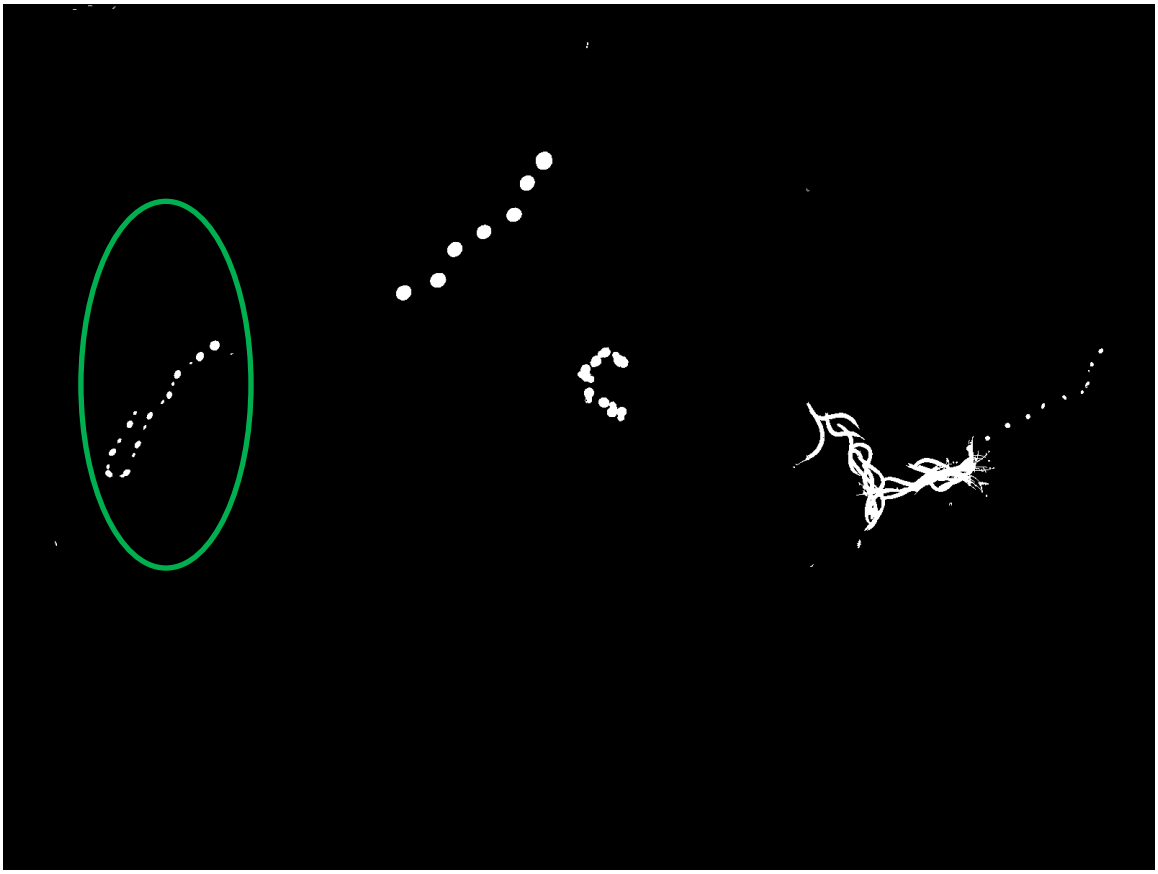

Figure 5. Composite image that shows motile organisms and their paths generated from both the brightfield and epifluorescence image sets. Note apparent size changes in organism circled in green. Also note the different organism size that is indicated by the brightfield (smaller) and epifluorescent (larger) microscope modalities.

Figure 6 provides a binary image that was created from the same epifluorescent image that was previously shown in Figure 1. As was the case in the bottom image shown in Figure 2, this image was created by applying a threshold to create the binary image shown in Figure 6. In this case, objects in the binary image have been size filtered, such that only objects made up of five or more pixels are displayed (and circled). The four objects circled in red were all motile and contributed to the algorithm's motility output image generated from the epifluorescent image set (Figure 4). The 15 objects that are circled in green in Figure 6 were not motile during the 31 sec observation window and therefore not depicted in the motility image provided in Figure 4. Note that there are more than 15 non-motile objects in this image, but some of these are associated with very weak signals. Initially, it was thought that these non-motile objects were the result of non-motile organisms that were stained (and hence viable) using the combination of FDA (Fluorescein diacetate) and CMFDA (5-Chloromethylfluorescein diacetate). However, work that has been performed on heat-killed samples strongly indicates that these types of signatures are not produced by viable organisms in the sample (See Section 4). This, in turn, makes motility significantly more important to determining organism viability when evaluating complex samples in the ≥ 10 µm to < 50 µm size class using automated methods.

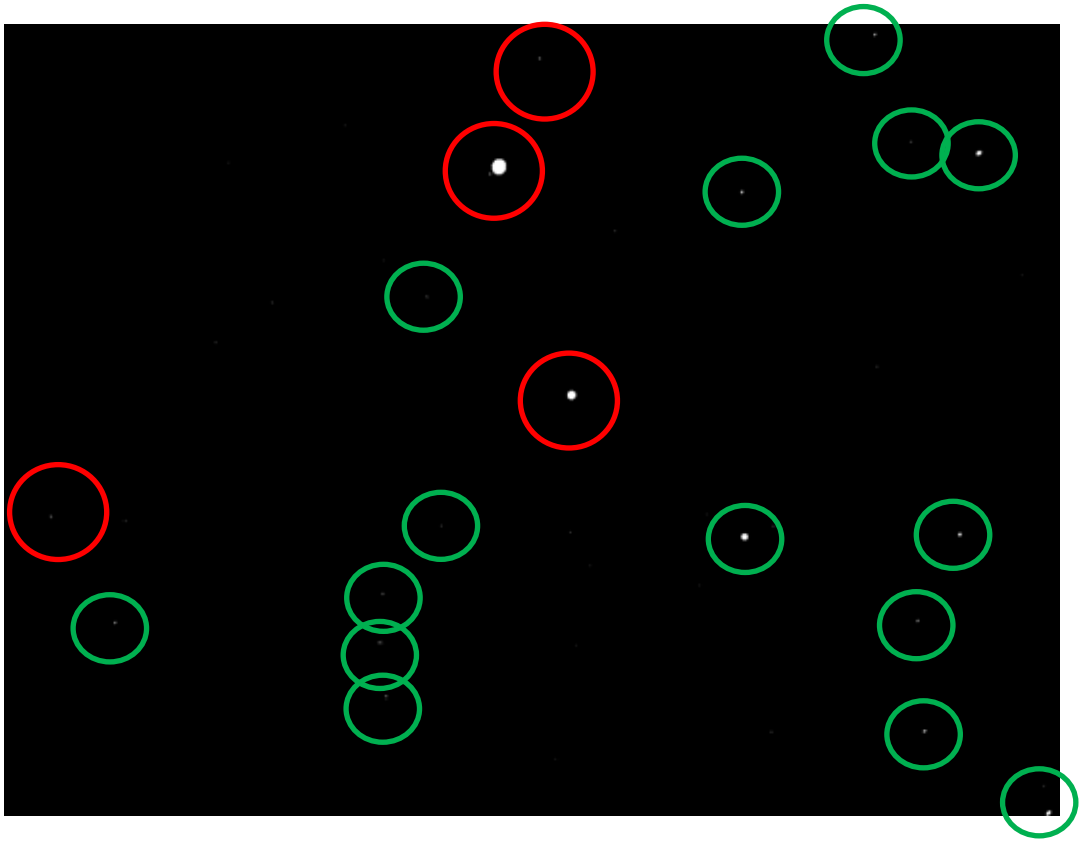

Figure 6. Binary image generated from the epifluorescent image previously shown in Figure 1. This image was generated by applying a threshold to the Figure 1 image and size filtering the binary image to display objects made up of five or more pixels. The four objects circled in red were motile during the 31 sec observation window (and contributed to the organism paths shown in Figures 4 and 5). The objects circled in green were not motile during the observation window.

The data provided in Figures 7 and 8 were generated from the same protist sample during tank fill operations associated with a simulated treatment experiment (a different sample than that which generated the data provided in Figures 1,2,3,4,5 and 6).

As has been mentioned, there is a large dynamic range associated with the epifluorescent signal levels produced from motile viable organisms. This makes it difficult to not only properly visualize the epifluorescent image data but also to properly specify thresholds for use in the motility algorithms. The difficulty in properly visualizing the epifluorescent image data is demonstrated in the three images shown in Figure 7. The image on the left is created by utilizing a LUT based on the maximum epifluorescent signal value to display the image data. There is only one faint object (circled in blue) that can be observed in this image. The image in the center is created by utilizing a LUT that is based on 10 % of the maximum epifluorescent signal value to display the image data. There are at least four additional objects, circled in red, that can now be observed by using this more sensitive LUT to display the image data. The image on the right is created utilizing a LUT based on 1 % of the maximum epifluorescent signal value to display the image data. A close review of this image shows that there are now at least 50 independent objects that can be seen in this image. This is the result of using an even more sensitive LUT to display the image data.

A review of the three images provided also shows that the size of objects in the displayed image varies with the choice of the LUT used to display the image data (Fig. 7). This is also true when binary images are generated from epifluorescent image data. In this case, lowering the threshold (the equivalent of using a more sensitive LUT to display data) increases the size of objects in the binary images. This provides another example of the difficulty in determining the size of detected organisms using automated methods.

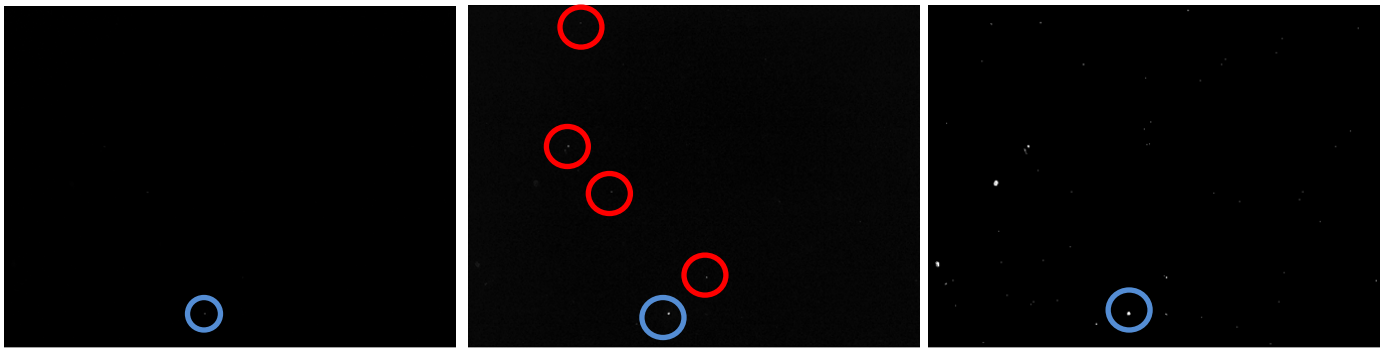

Figure 7. The Look Up Table (LUT) used to display data affects both the number and size of objects that can be observed in epifluorescent image data. The image on the left uses the maximum signal level in the LUT used to display the image data. The images in the center and right use 10 % and 1 % of this maximum signal level in the LUT used to display the image data, respectively. The one object that can be observed in all three images is circled in blue. Four additional objects that can be observed using the LUT based on 10 % of the maximum signal are circled in red. In general, as the LUT is made more sensitive, both the number and size of observable objects increases.

The LUT used to display epifluorescent image data and the thresholds used to create binary images from these data can significantly impact the specific results that are obtained using the motility algorithms. Figure 8 provides two algorithm output motility images generated from the epifluorescent microscope modality from a complex sample (same as Figure 7). The image on the left and right used 2 % or 1 % of the maximum epifluorescent signal, respectively, as a threshold to create the binary images used in the motility algorithm. A review of the data provided in Figure 8 shows that an increased number of organisms, many of which moved along well defined paths (such as that circled in blue in the image on the right), are detected in this epifluorescent image set generated with the smaller threshold value. This shows that even a relatively small change in the threshold used to create binary images can have a significant impact on the number of organisms that are detected in complex samples. A review of the image data also shows that the apparent size of objects in these images increases as binary thresholds are lowered.

As a result of the sensitivity of the motility algorithms to the specific threshold used to generate binary images, the epifluorescent motility algorithm now utilizes an automated method for determining the specific threshold to use when creating the binary images. This approach has reduced the previous requirement to iteratively identify the proper threshold to create binary images in the epifluorescent motility algorithm.

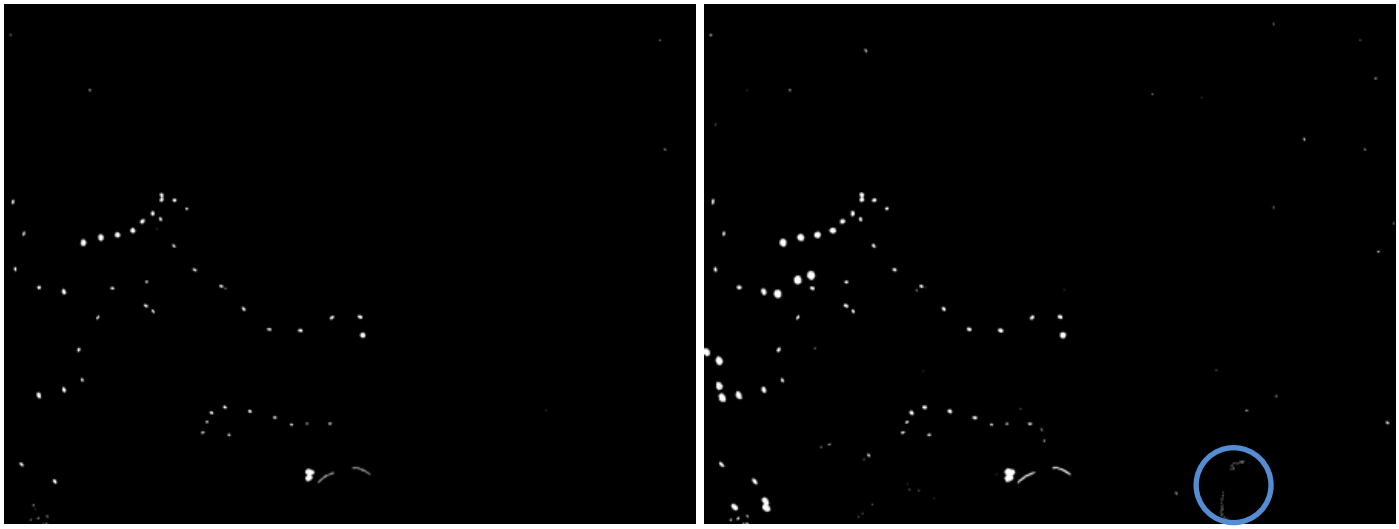

Figure 8. The threshold used to create binary images can impact the results of the motility algorithms. The image on the left used 2 % while the image on the right used 1 % of the maximum epifluorescent signal to generate binary images. The choice of threshold can both increase the number and size of detected organisms. Path of a motile organism that is not visible using the 2% threshold is circled in blue.

2.4 Sample Imaging and Analysis Summary

The results provided in this report section demonstrate that clear and non-distorted image sets can be collected from both microscope modalities when using unruled SR chamber to hold samples under the microscope. It is for this reason that this is presently the preferred device for a samples platform when collecting image sets.

The results also show that there are benefits associated with using both microscope modalities, as certain organisms are uniquely detected in data generated by either brightfield or epifluorescence illumination. This provides a strong argument for continuing to work with both microscope modalities when imaging samples of organisms in the ≥ 10 µm to < 50 µm size class.

Lastly, the results strongly indicate the need for utilizing motility algorithms when analyzing samples using automated methods. This provides a strong argument for continuing to collect time resolved image sets of samples in the ≥ 10 µm to < 50 µm size class.

3 COMPLEX SAMPLES

3.1 Sample Complexity

Samples associated with BWMS in the ≥ 10 µm to < 50 µm size class are inherently complex. This is the result of particulate, organic, and mineral matter that are added to the challenge water. The complexity of samples is also a result of the significant concentrations of viable organisms and the variety of organism types in this size class. This is true for both initial test and control fill operations (where challenge levels of organisms must be present) but also for evaluating samples from the control tank (which have not been treated with a BWMS).

During this year's ongoing research and development effort, the effect of sample complexity on the ability to both accurately enumerate and determine the size of detected motile organisms in the ≥ 10 µm to < 50 µm size class was explored.

3.2 Sample Enumeration and Organism Size Determination

The automation algorithm that operates on the epifluorescent image data now uses both image normalization and adaptive thresholding methods to create the binary images that are subtracted. This approach affords the ability to identify motile organisms with weak epifluorescent signals. The algorithm that operates on the brightfield microscope modality uses image normalization but does not currently utilize adaptive thresholding methods to create binary images. In both microscope modalities, many organisms are only observed occasionally in the 10 image pairs that are analyzed. This can result from debris obscuring organisms (brightfield) and from low contrast or weak organism signatures (both brightfield and epifluorescent). In the case of weak epifluorescent signatures, as was described in the previous section, in many cases, decreasing the thresholds used in the motility algorithms can result in an increased number of observed organisms. Also, if an organism is a virtually non-motile or a "slow mover", it may not move sufficiently to create two discrete signals. All of these types of issues make it more difficult to accurately enumerate the number of organisms based on analyses of the algorithm's output "motility" images.

Accurately determining the size of detected motile organisms can also be difficult. First, with organisms in the ≥ 10 µm to < 50 µm size class, the relatively small organism size combined with depth of focus limitations and the 1 mm height of the water column in unruled SR chambers can cause the apparent size of organisms to increase by greater than a factor of two in both the epifluorescent and brightfield image data as organisms travel vertically through the water column. This can also be an issue when using manual microscopic methods, as size can only be accurately determined by focusing on every observed organism. In a complex sample with many motile organisms, this may not be feasible within a reasonable observation time.

Second, there is quite often a difference in the apparent size of observed organisms in the brightfield versus epifluorescent microscope modalities. Further, as discussed in the last section, the size of objects in the epifluorescent microscope modality is impacted by the thresholds used to either create a binary image or to display the image data. Many organisms are uniquely detected in the epifluorescent microscope modality and, as such, it may be impossible to accurately determine their size without brightfield microscope corroboration. Camera settings (primarily exposure and gain) can also affect the apparent size of objects in epifluorescent image data.

The following discussion of three examples of motility analyses performed on samples with organisms in the ≥ 10 µm to < 50 µm size class illustrates issues associated with accurately enumerating and determining the size of organisms in complex samples.

Figure 9 provides the output motility images generated during a motility analysis of data from a simulated treatment experiment conducted at the Naval Research Laboratory in Key West, FL (NRL KW) in April 2011. These data were associated with the tank fill operation and are generated from the same sample as the one that produced the data provided in Figures 7 and 8. The image shown in the upper left is the output motility image from the brightfield image data. The image shown in the upper right is the output motility image from the epifluorescent image data, and the bottom image is the combined output motility image. A review of these images shows that unique organisms were detected in both the brightfield and epifluorescent microscope modalities. However, a significantly greater number of organisms were detected in the

epifluorescent microscope modality. In these data there are few if any cases where 20 objects were produced in either microscope modality (i.e., where the same organism was detected in each of the 20 pairs) or where 40 objects were produced in the combined output motility by any single organism. In the output motility images from both microscope modalities there were also many cases where a motile organism was only detected occasionally (and rarely) during the evaluation of the ten image pairs. Lastly, there were several organisms that were "slow movers" that produced line or blob like signatures rather than discrete signatures. Examples are circled in red in the combined output motility image. All of the issues make enumerating the number of viable organisms more difficult.

A review of the images in the upper right and bottom of Figure 9 provides a good example of the difficulty in determining the size of objects. The blue circled region in the upper right image shows the path of an organism that at least doubled in size as it moved towards the left. The image on the bottom shows that there may be as much as a factor of ten difference in the apparent size of organisms detected in the brightfield (smaller-sized) and epifluorescent (larger-sized) microscope modalities.

Figure 10 also provides the output motility images generated during a motility analysis of data generated during a simulated treatment experiment conducted at NRL KW in April 2011. These data were associated with the test tank following filling. The image shown in the upper left is the output motility image from the brightfield image data. The image shown in the upper right is the output motility image from the epifluorescent image data, and the bottom image is the combined output motility image. A review of these images show that there were unique organisms detected in both the brightfield and epifluorescent image data. The data shown in Figure 10 is actually less complicated than the data that was shown in Figure 9. However, many of the organisms are only detected in a limited number of the image pairs in both microscope modalities. Further, there are several slow moving organisms that produced amorphous line-like signatures rather than discrete signatures when the results across image pairs and across microscope modalities were combined. All of these issues make accurate enumeration of the organisms in this sample more difficult. The apparent difference in size of organisms detected in both the brightfield and epifluorescent data can also be seen in these data.

Figure 11 provides another set of output motility images generated during a motility analysis of data from a simulated treatment experiment conducted at NRL KW in April 2011. These data were associated with the control tank after a 1 day hold. The image shown in the upper left is the output motility image from the brightfield image data. The image shown in the upper right is the output motility image from the epifluorescent image data, and the bottom image is the combined output motility image. A review of these images show that there were no unique organisms detected in the brightfield image data and that the majority of organisms are detected in the epifluorescent image data. Even with very few organisms detected in the brightfield imagery, this is the most complicated image set of the three presented in this report section.

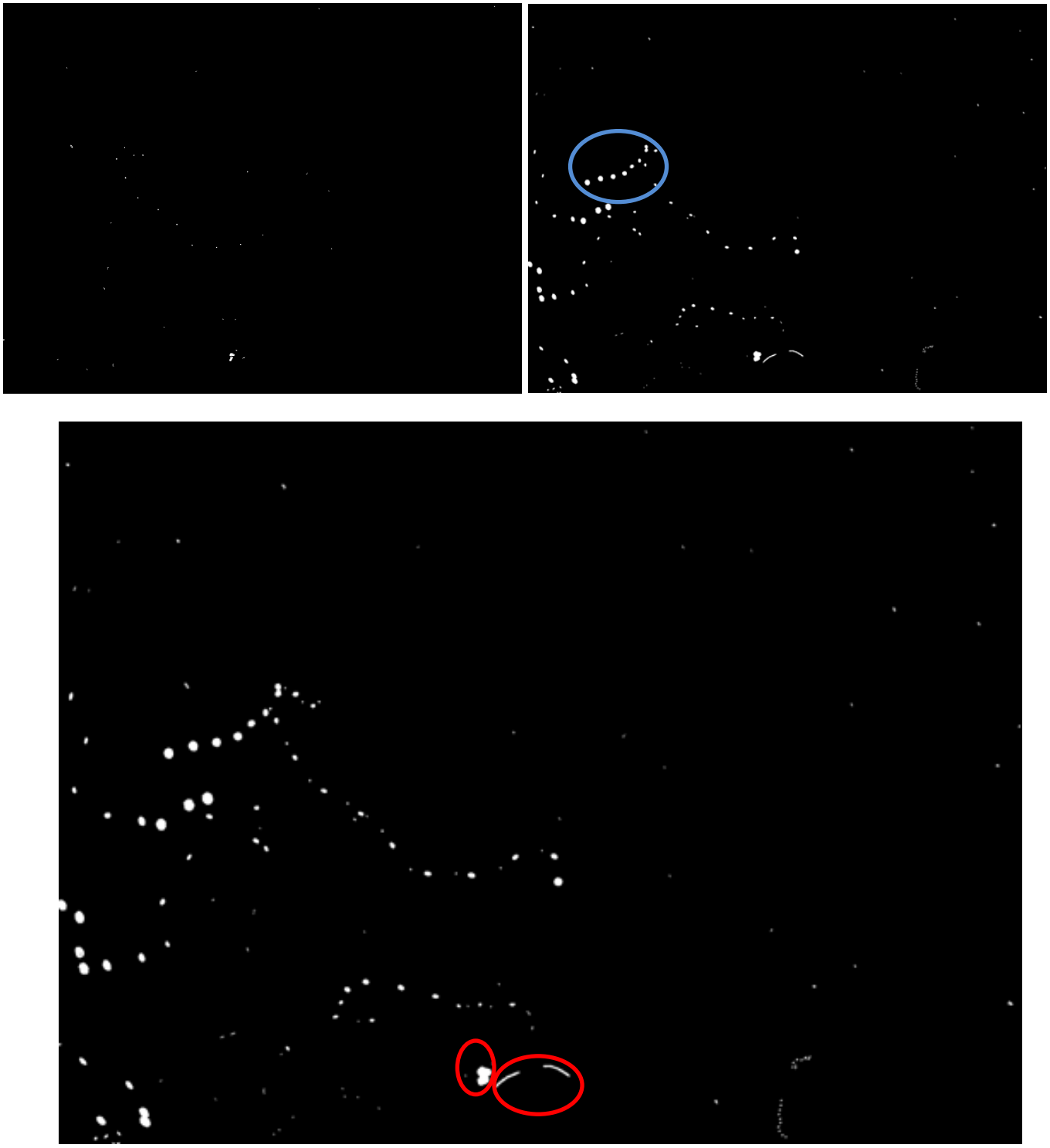

Figure 9. Motility output images generated from a complex sample generated during a simulated treatment experiment conducted at NRL KW in April 2011 (Control Tank T = 0). The image on the top left is the brightfield output motility image. The image on the upper right is the epifluorescent output motility image, and the image on the bottom is the combined output motility image. Note the apparent size change of organism in blue circle. The two organism paths circled in red were produced by slow moving organisms.

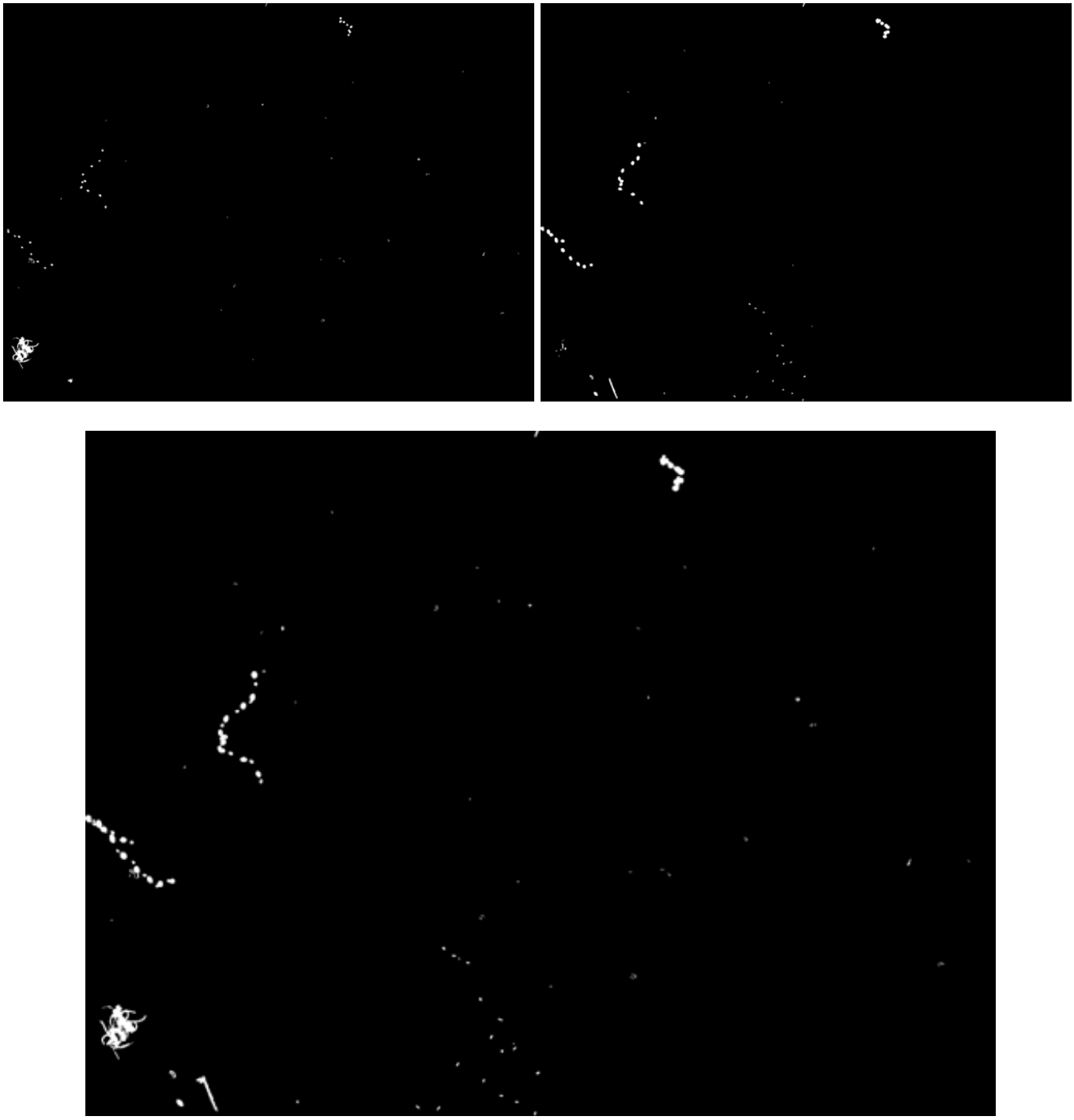

Figure 10. Motility output images generated from a complex sample generated during a simulated treatment experiment conducted at NRL KW in April 2011 (Treatment Tank T = 0). The image on the top left is the brightfield output motility image. The image on the upper right is the epifluorescent output motility image and the image on the bottom is the combined output motility image.

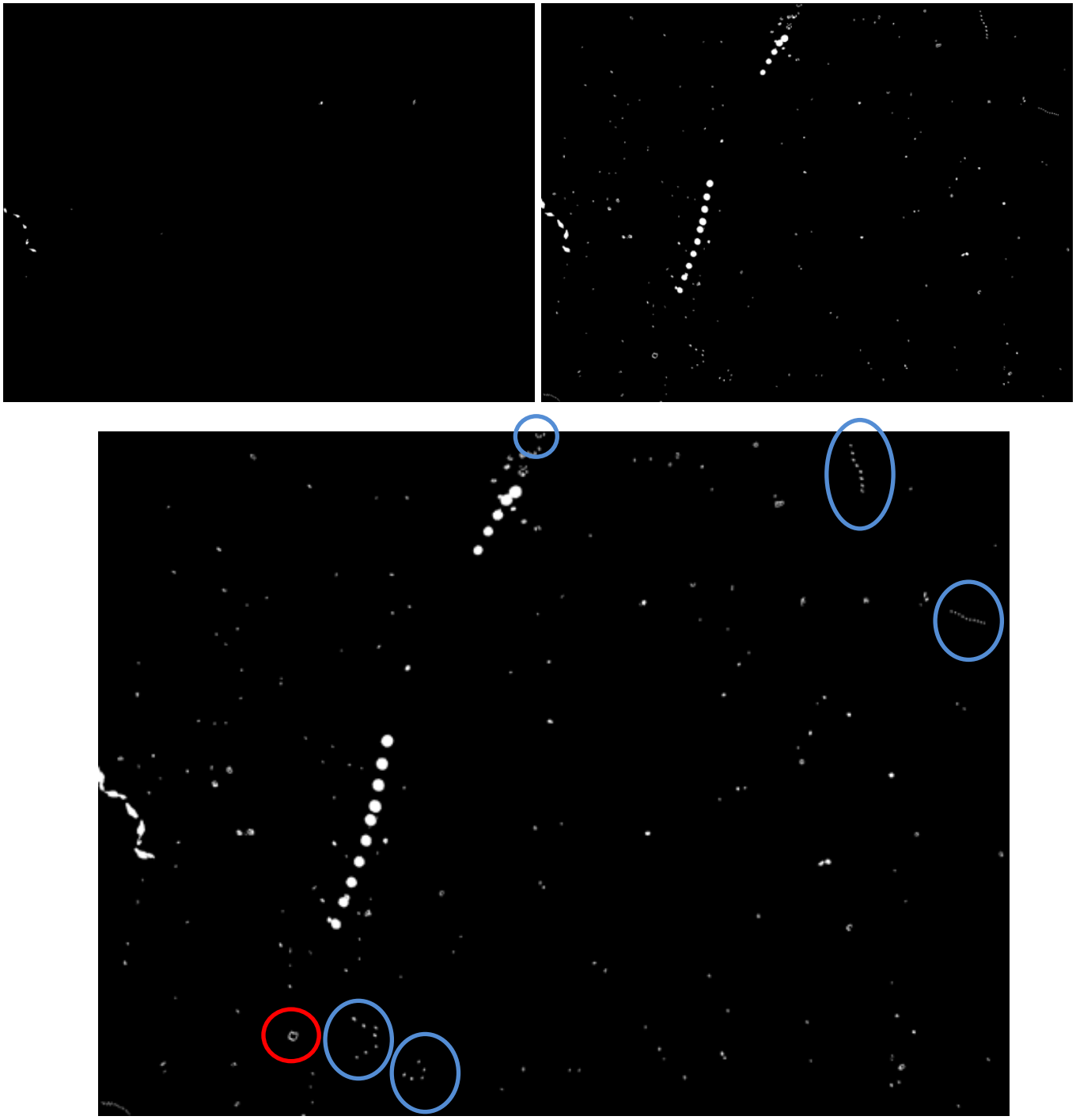

Figure 11. Motility output images generated from a complex sample generated during a simulated treatment experiment conducted at NRL KW in April 2011 (Control Tank T = 24 Hours). The image on the top left is the brightfield output motility image. The image on the upper right is the epifluorescent output motility image and the image on the bottom is the combined output motility image. Small organisms with well defined paths are circled in blue. The signal produced by an air bubble is circled in red.

In order to detect many of the motile organism ≥ 10 µm to < 50 µm, a relatively low threshold was used to create the binary images (the same value used to create the data shown in Figure 9). This low threshold was required to detect the comparatively small organisms (or organisms with spatially small epifluorescent signals) that created the well-defined paths that are circled in blue in Figure 11. At this lower threshold value, many objects that do not appear to be associated with motile organism paths can be observed in these data. A more detailed review of the imagery associated with this analysis has shown that many of these objects are associated with viable organisms that were only detected in a few of the ten image pairs that are analyzed. There were also a number of objects that were not associated with viable organisms but were associated with sample artifacts (such as the signature of a slightly moving air bubble circled in red). These types of issues make accurate sample enumeration more difficult. The data in Figure 11 also show the same sort of disparity in the apparent size of organisms that were detected in both the brightfield and epifluorescent image data. This makes it difficult to definitively determine the size of the organisms that are detected in this motility analysis.

3.3 Complex Samples Summary

The data provided in this report section demonstrate the difficulties associated with developing accurate counts of motile organisms from the algorithm's output motility images. A more complex method than counting the number of objects and dividing by either 20 (individual microscope modalities) or by 40 (composite) is required for accurate motile organism enumeration. An enumeration algorithm that develops a motile organism count but which also provides upper and lower bounds is under development. The upper and lower bounds are developed by considering all of the possibilities for what might create motility signals (e.g., two "objects" or blobs can be from either the same object or two separate objects) in the output motility images at both the image pair and full observation window levels. The organism count will be generated by determining the most consistent data interpretation based on results obtained at both the image pair and full observation window levels as well as across both the individual and combined microscope modalities. It is believed that this type of approach will lead to a robust enumeration algorithm for working with these data.

The data provided in this section also demonstrate some of the issues associated with accurately determining the size of detected organisms and clearly show that the apparent size of organisms can be very different in the brightfield and epifluorescent microscope modalities. The data also show that, as a result of microscope depth of focus issues combined with the 1 mm water column depth, the apparent size of organisms using the epifluorescent microscope modality can double as organisms move across the field of view. Although not presented in this section, similar increases in organism size have been observed in the brightfield microscope imagery as organisms move up and down the water column as they move through the microscope's field of view.

4 RESULTS FROM HEAT-KILLED SAMPLES

Prior to April 2011, it was thought that samples with organisms ≥ 10 µm to < 50 µm would be analyzed such that both motile and non-motile organisms would be identified by the automation algorithms. The basic concept was to first identify all of the motile organisms in both microscope modalities and then identify the non-motile organisms in the epifluorescent image data. The non-motile organisms would be identified as having an epifluorescent signal above the threshold value used to create the binary images used in the automation routines. In Section 2 of this report, Figure 6 provided an example of how this type of analysis might be performed. First the four motile organisms (circled in red) would be identified and then the non-motile organisms (circled in green) would be identified and included in the overall counts. For this

type of approach to be valid, it is essential that the signatures that were circled in green in Figure 6 be produced by viable organisms that had been stained by the CMFDA and FDA and not be the result of other types of sample artifacts. In order to determine if the non-motile signatures produced in images such as those shown in Figure 6 were from viable organisms, additional work was conducted with "heat-killed" samples that were stained with CMFDA and FDA.

Heat killing was performed by submerging the sample vessel in a water bath. Samples, which contained a mixed community of ambient protists, were placed in a water bath that was preheated to either 50 °C or 65 °C. The temperature of the sample was monitored with a thermometer, and once the sample temperature reached the target value, the sample was held in the water bath for 15 minutes. Following exposure to the water bath, the sample was cooled to room temperature, stained and analyzed.

Figure 12 provides six images. Each row of images consists of a binary image generated from one of the images from the epifluorescent image set on the left and the algorithm's output motility image (generated from the epifluorescent microscope modality) on the right. The binary image is generated by applying a threshold (whose value is defined by the adaptive thresholding algorithm) to the epifluorescent image. The top row of images is associated with a low temperature (50 °C) heat-killed sample. The middle row of images is associated with a higher temperature (65 °C) heat-killed sample. The bottom row is associated with an ambient sample (not heat-killed) that was collected on the same day as the low temperature heat-killed sample.

A review of the algorithm's output motility images produced by the epifluorescent image sets (on the right in Figure 12) show that organism motility was only observed in the sample that was not heat-killed. There are a number of "on pixels" in the motility images associated with the two heat-killed samples (especially the higher temperature heat-killed sample), but none of these show the types of paths that are indicative of motile organisms being present in the sample.

It is believed that the "on pixels" in the heat kill output motility images are the result of camera noise in the epifluorescent imagery. Camera noise results in an apparent intensity variation at the individual image pixel level. In cases where there are true epifluorescent signals in a sample, this variation is small with respect to the epifluorescent signal and generally only impacts the algorithm when smaller thresholds (≤ 1 %) are used to create the binary images. In cases where there is not a true epifluorescent signal in the sample, the image intensity normalization routines that are utilized in the automation algorithms combined with camera noise induced intensity variations can result in image pixels exceeding the threshold values used to create binary images. With a camera noise induced intensity variation, these image pixels can be above the threshold in some images and below the threshold in other images that make up the 31 image epifluorescent image set. As these objects are turned on and off in different images that make up the image set, they can be mistaken for motile organisms by the automation algorithms and are displayed in the motility images. As stated above, in these cases, there are no paths in these motility images that are indicative of being produced by motile organisms.

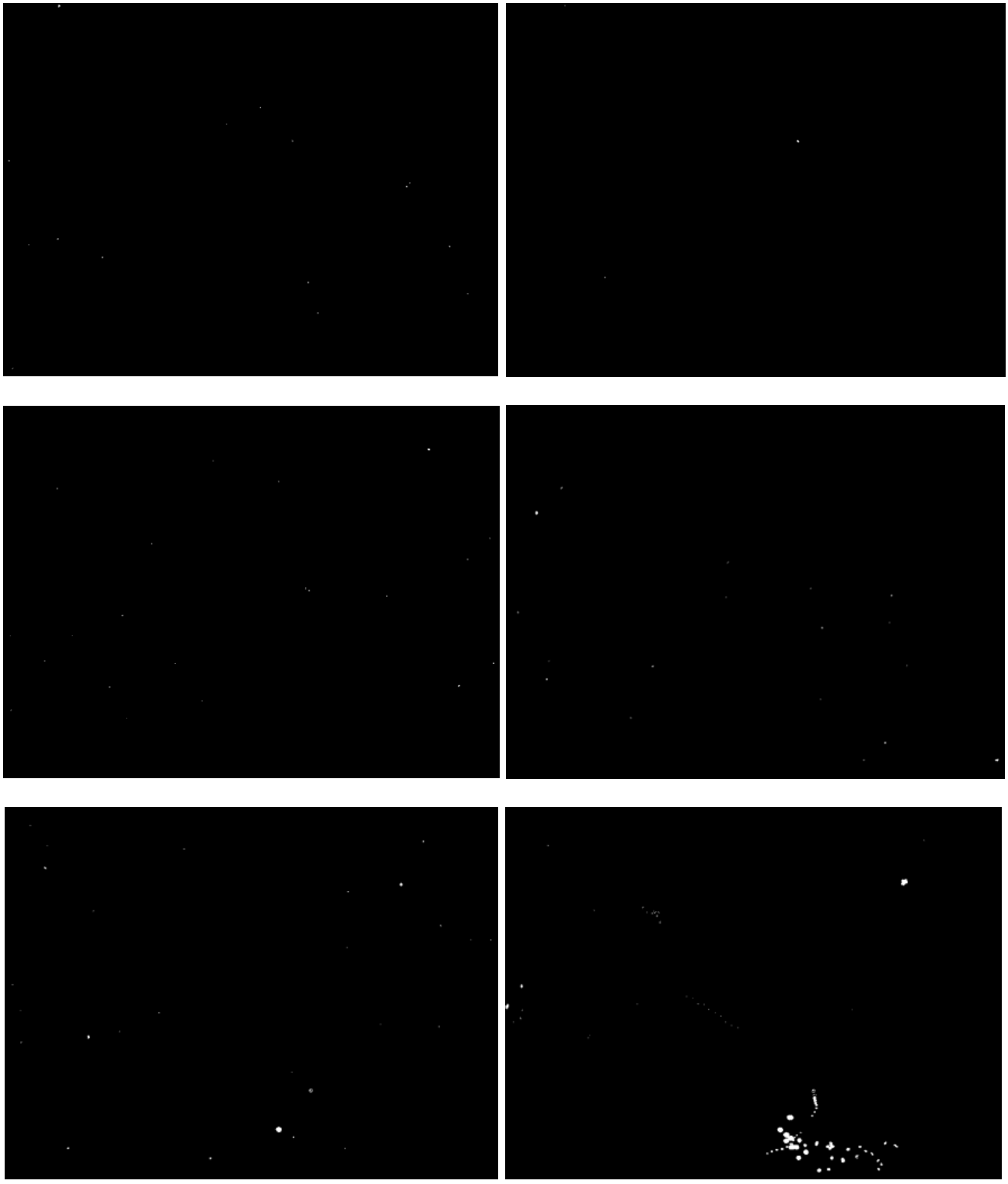

Figure 12. Heat kill experiment epifluorescent images. In each row a binary image generated from a single image is on the left and an output motility image on the right. The top row is associated with a low temperature (50 □C) heat-killed sample. The middle row is associated with a higher temperature (65 □C) heat-killed sample. The bottom row is associated with an ambient sample (not heat-killed) that was collected on the same day as the low temperature heat-killed sample.

In the case of the high temperature heat killed data shown in the middle of Figure 12, it is believed that the majority of objects in the motility image are the result of camera noise. In the case of the low temperature heat killed data shown at the top of Figure 12, it is believed that one object is the result of camera noise and the second object (the larger signature) results from a FDA and CMFDA stained particle that is moving through the water column (the larger signature). In all 18 of the low and high temperature heat killed samples evaluated, a manual review of the epifluorescent image sets as "movies" using a very sensitive LUT that emphasizes any and all objects with epifluorescent signals provided no indication of any motile organisms.

A review of the binary images shown on the left in Figure 12 is more problematic. There is no appreciable difference in the number of objects with epifluorescent signals between the two heat-killed samples or between either of the heat-killed samples and the non-heat-killed sample. These images show that, in general, the size of some of the objects with epifluorescent signals is larger in the non-heat killed sample than those in the heat-killed samples, but this was not the case in all of the 18 heat-killed samples that were evaluated. Further, all the binary images in this figure have been filtered such that they only include objects that are made up of five or more pixels. Many of the motile organisms that produce well-defined paths also produce signals in this size range. Work with filtered seawater samples showed some indication of objects with epifluorescent signals, but none of those objects produced a signature that was five pixels or larger in size. This makes it difficult to exclude the types of signals such as shown in left panels of Figure 12 based on their size.

The overall signal levels and their variation combined with information used in the image normalization and adaptive thresholding algorithms appear to provide a better indication than object size that an observed object with an epifluorescent signal is not produced by non-motile stained organisms or sparsely detected stained motile organisms. This was true for all twelve high temperature heat-killed samples and for two of the six low temperature heat-killed samples. In these fourteen heat-killed samples, the dynamic range of the epifluorescent image data was at least a factor of four lower than that associated with samples that had viable organisms. The table of numbers that is used by the adaptive thresholding algorithm also produces a very different pattern in these fourteen heat-killed samples as well as in other samples that have no viable organisms.

The heat-killed image data provided in this report section provides strong indications that non-motile epifluorescent signals that are observed in samples should not be attributed to non-motile organisms in the sample. Based on the work performed to date, motility is the only modality that can be used to uniquely identify viable organisms in complex samples when using automated methods. As discussed in the next report section, samples with a large diverse population of "non-motile" organisms in the ≥ 10 µm to < 50 µm size class should be evaluated to determine how (and if) these organisms impact the automation algorithms and the results that they generate. These data can also provide a basis for modifying the automation algorithms so that they can better account for "non-motile" organisms in this size class.

5 STATUS AND RECOMMENDATIONS

As a result of this year's ongoing research and development activities, two major changes were implemented in the protocols described in Nelson, et.al. 2010a. First, it is now recommended that unruled SR chambers with the microscope set up to observe a 20 µL sample volume be used for organisms in the ≥ 10 µm to < 50 µm size class. The justification for making this change is provided in Nelson, et.al. 2011b. Making this modification is also supported by the high quality image data that is provided in this report.

Second, it has been determined that when using automated methods, viable organisms in the ≥ 10 µm to < 50 µm size class can only be identified in complex samples using motility. This conclusion is based on the results performed on heat-killed samples that were presented in Section 4 of this report. Data presented in that section show that many of the observed non-motile objects with epifluorescent signals are not produced by viable organisms. Consequently, until samples are collected with a large and diverse assemblage of "non-motile" protists, the automated methods will not consider these objects when determining the number of viable organisms in samples in the ≥ 10 µm to < 50 µm size class.

During this year's ongoing activities, a significant number of complex samples were analyzed using automated analysis software. As a result of these evaluations, the automation algorithms were significantly advanced. There is an immediate need to develop and incorporate a robust enumeration algorithm for both determining the number of motile organisms in the ≥ 10 µm to < 50 µm size class, but also for generating upper and lower bounds for these counts. This requirement is driven by the complexity of the samples that will be analyzed using the automation algorithms and has been described in this report in Section 3.

Work performed on complex samples has also shown that it is difficult to accurately determine the size of organisms that are detected in complex samples. For organisms in the ≥ 10 µm to < 50 µm size class, the ability to accurately determine organism size is affected by the height of the water column and microscope depth of focus issues, by the LUTs and thresholds used to visualize the epifluorescent image data, and the difference in the apparent size of objects that are observed in both the brightfield and epifluorescent microscope modalities. This has also been described in Section 3 of this report.

It is likely that some modifications will be made to the motility detection algorithms. As described here, the motility algorithms used with the epifluorescent image data use both an image normalization and adaptive thresholding method to generate binary images. It is believed that changing the threshold increment step used in this algorithm might improve its overall ability to detect motile organisms that are only weakly stained. The motility algorithm used to analyze the brightfield image data currently uses image normalization and fixed thresholding methods. It is believed that the use of an adaptive thresholding algorithm in this algorithm will improve it with an enhanced ability to detect low contrast (compared to the debris) motile organisms in the brightfield imagery. It is recommended that ongoing incremental improvements continue to be made to the existing algorithms based on the analysis of additional complex sample data.

A critical next step will be to compare the results obtained using automated methods to manual counts generated on the same samples. As a result of the extensive amount of samples that were generated and analyzed in support of this year's program, the data are currently in place to accomplish this once a more robust enumeration algorithm is developed. It is highly recommended that these comparisons be made as part of next year's program.

It is also recommended that samples with a large and diverse population of "non-motile" organisms in the ≥ 10 µm to < 50 µm size class be evaluated. There is some indication that relatively non-motile organisms are detected by the motility algorithms as they move up and down in the water column and rotate orientation. Further, working with this type of sample may lead to additional methods to identify and account for these organisms when performing automated analyses.

It is highly recommended that complex samples from other facilities and potentially other applications be generated and analyzed using the Protocols for Automated Protist Analysis. It is also recommended that anytime a complex experiment is conducted at NRL KW, that samples are generated and analyzed using the Protocols. The ability of the automated analysis algorithms to work well on data generated at other test facilities as well as from data generated from more complex experiments conducted at NRL KW is important to the developed methods obtaining broad acceptance by the larger community involved with BWMS testing.

(This page intentionally left blank.)

6 REFERENCES

Nelson, B.N., Steinberg, M.K., Robbins, S.H., Lemieux, E.J., Herring, P.R., and Drake, L.A. (2010a) Feasibility of Developing a Protocol for Automated Protist Analysis. Coast Guard Letter Report: CG-D-02-11, Accession Number: ADA535257

Nelson, B.N., Steinberg, M.K., Lemieux, E.J., Herring, P.R., and Drake, L.A. (2010b) Analysis Protocol for Automated Zooplankton Analysis. Coast Guard Letter Report, Report: CG-D-01-11, Accession Number: ADA535255

Nelson, B.N., First, M.R., Steinberg, M.K.; Riley, S.C., Robbins-Wamsley, S.H.; Herring, P.R. and Drake, L.A. (2011a) Protocols for Automated Zooplankton Analysis. Coast Guard Letter Report in final publication.

Nelson, B.N., First, M.R., Steinberg, M.K.; Riley, S.C., Robbins-Wamsley, S.H.; Herring, P.R. and Drake, L.A. (2011b) Protocols for Automated Protist Analysis. Coast Guard Letter Report in final publication.

(This page intentionally left blank.)