Enclosure (1) - Response to DOH Comments on *Navy's Response to DOH Comments on Draft Groundwater Model Report, dated September 26, 2024* – April 6, 2026

1. Overview - Field investigations in support of the modeling effort have advanced our understanding of the geology and hydrogeology under Red Hill and made commendable progress in addressing knowledge gaps the DOH identified in responses to previous model versions. Until the model is technically and scientifically validated using this real-world data, it cannot be used for contaminant risk and remediation planning and decision-making. The model should seek to answer several critical questions, but these have yet to be satisfactorily answered with the model:

- ***What is the Red Hill Shaft (RHS) zone of capture?***
- ***Is there a flow path from beneath the Red Hill Bulk Fuel Storage Facility (Facility) to the northwest?***
- ***What are the groundwater flow trajectories within the Facility boundaries, and can the model define these flow trajectories to the level of detail needed to provide the requisite foundation for contaminant fate and transport assessment, and contamination remediation design?***

While the NCTF-RH asserts that the model currently answers these questions, and draws other conclusions, an important approach to evaluating how well the current model addresses these questions is to compare the simulated particle paths with the results of University of Hawai'i's (UH's) tracer test and borescope survey results when these data become available.

Navy Response –

The Navy agrees that model results should be compared to all the field data, studies, and tests, including those conducted by UH, and remains open to incorporating applicable data pending its availability. For example, the Navy has already incorporated the one finalized data set that has been provided by UH, which can be directly used in the modeling input – UH's geophysical transects near the north slope of Red Hill. The Navy engaged with UH on how best to interpret those study results and then input agreed upon changes into the Geological Site Model (GSM) which underlies all the other models.

2. Nature of DOH Technical Comments -The following includes some technical feedback. Due to the substantial volume of materials submitted, including nearly 13,000 pages of documentation and extensive modeling files, our review was conducted at a high level with emphasis on key elements.

Navy Response –

Comment noted. Some of the responses that follow identify portions of the report that are relevant to the comments.

3. Model Matching Observations - While the NCTF-RH's model does produce results that agree with many data points, there are critical differences between the modeled and measured results. Refer to Navy Response, page 1, item 2. Particularly, the NCTF-RH's model does not capture the variability of vertical and horizontal hydraulic gradients within the Facility boundaries, does not account for observed enrichment in chloride and nitrate, and produces flow trajectories that are not consistent with observed migration of contamination. Finally, the NCTF-RH's model results are not supported by the on-going tracer test that shows complex

tracer migration pathways with high concentrations of dye retained at the injection well (RHMW08) and much lower concentrations in some of the monitoring points.

Navy Response -

The Navy appreciates and agrees with DOH's acknowledgment that the 2024 models do agree with many data points. We recommend further discussion about vertical gradients (for which there are few data locations) and ion concentrations and recommend working towards practical agreement on the applicable metrics.

The Navy agrees with UH's recommendation that conclusions not be drawn on the tracer study until UH completes data collection late in 2026 and presumably analyzes the data and draws conclusions afterwards. The preliminary data we have seen suggests detections of dyes in monitoring wells that are compatible with modeled rates and directions of flow.

The Navy also notes that additional improvements have been made since the 2024 modeling effort.

4. Model Improvements - We would like to acknowledge some significant changes made in the flow model in the recent September 2024 iteration. There has been an attempt to capture the heterogeneity of the lava flows. This was in response to comments from the DOH's and U.S. Environmental Protection Agency's, hereinafter the Regulatory Agencies' (RA's), subject matter experts. The model better incorporates tilted layering to better represent the lava flow geometry.

Navy Response -

The Navy appreciates this acknowledgement, and thanks the RAs for the constructive recommendations that helped develop these improvements.

5. Subsurface Heterogeneity - While significant changes have been made to capture the complexity of the Moanalua/Red Hill/Halawa region, field measurements indicate that subsurface geology is highly complex at multiple spatial scales and its current representation in the model is only a best guess with significant uncertainty given the current amount of available data. For example, the current model grid structure has only a single strike and dip of the lava flows. Data presented in NCTF-RH and RA discussions show very high variability in the geometry of the lava flows with dip directions varying from near south to north northwest. The geometry of the lava flows will greatly affect the aquifer anisotropy, a parameter the NCTF-RH used to reconcile the apparent groundwater gradients with the simulated particle track trajectories. It may not be reasonable to assume the model can adequately capture the subsurface complexity at all spatial scales and the modeling effort would be better served by simplifying the model and focusing on questions that can reasonably be answered by the modeling.

Navy Response -

The Navy agrees that it is unreasonable to expect any model to capture the appreciable subsurface complexity at all spatial scales at this site. The current model uses the dip and strike that was discussed at length over the years and ultimately agreed upon between the Navy and the

RAs during a special purpose meeting conducted on January 10, 2023. In this sense, the GSM is necessarily a simplified representation of the complex and highly heterogenous site conditions, which is why some of the sensitivity studies attempted to evaluate aquifer properties. Sensitivity studies 6 and 7, for example, evaluate other dip and strike values. The Navy recommends retaining the agreed upon average value, in accordance with this recommendation of "simplifying the model". The Navy agrees that we should focus on questions that can reasonably be answered by the models and developing realistic expectations. The Navy welcomes discussion of any particular simplifications that the DOH might have in mind.

6. Appropriate Applicability and Utility of the Model - The next concern is the best use of the model at this point for Red Hill decision-making processes. The Underground Storage Tanks (USTs) have been defueled, reducing the risk of future large spills and off-site migration of contamination. The critical question facing NCTF-RH is remediation and the extent of the Red Hill Shaft capture zone (which is primarily important if the elevated contamination is still mobile). The modeling effort should identify those questions that are important to answer and are within the capability of a model to answer. Also, when the UH study results become available, the modeling effort should be directed to resolve the discrepancies between the UH Red Hill study and the current groundwater model. Below we have outlined more specific feedback.

Navy Response -

The Navy generally agrees.

Fortunately, the concentrations are decreasing and biodegradation is active. The extensive data collected during recent years does not indicate mobile elevated concentrations of data migrating offsite. As to UH's studies, the Navy intends to let UH complete its work before drawing any conclusions but is not aware of specific discrepancies at this time. As to capture, the Navy recommends discussing how precise the delineation of capture needs to be in light of the concentrations of petroleum-related analytes in the aquifer and entering the Red Hill Shaft at this post-defueling stage (generally very low and predominantly non-detect), and requests a clarification of what specific metrics or approaches the DOH recommends.

7. Cross-Ridge and Vertical Hydraulic Gradients

- ***a. The alluvium and underlying saprolite in South Hālawa Valley northwest of the tanks is a partial geologic barrier; however, there is uncertainty about whether this barrier prevents groundwater flow to the northwest or if there is flow beneath and around the upslope edge of this formation.***
- ***b. Assessing the risk that contamination at the Facility poses to offsite receptors is largely dependent on any groundwater that flows from beneath the Facility to the northwest. Graphs in Vol. 2, PDF pages 73 through 78 focus on gradients down the Red Hill Ridge. However, it is the gradient across as opposed to down the Red Hill Ridge and vertical gradients that will move water from beneath the Facility to the northwest either around or beneath the alluvial saprolite wedge. Table 2-13 shows that the model fails to adequately capture the magnitude of the vertical gradients, particularly in those West Bay wells along the southeast side of the Hālawa valley fill. Since flow to the northwest would be manifested as a downward vertical gradient and a horizontal gradient to the***

northwest, both need to be more critically evaluated and matched. In addition, the groundwater gradient between the axis of the Red Hill Ridge to the contact between alluvium and basalt to the northwest has not been graphically evaluated. Graphs like those on Vol. 2, PDF pages 73 through 78 should be developed for cross-ridge gradients.

Navy Response -

Challenges with the simulation of vertical gradients in 2024 were discussed at length during the 2025 Modeling Workshop. First, the Navy notes that most wells in the extensive monitoring network (which was developed in consultation with the DOH) are water table wells, and there are few locations with field data indicative of vertical gradients. We also note that all models include deviations from field measurements and contain areas that, through lack of detailed information, contain simplifications and idealizations. The Navy recommends a discussion of the vertical gradient data and modeling results and welcomes constructive suggestions.

In the next phase of modeling, the Navy will develop cross-ridge gradient graphs, as suggested.

8. Vertical Flow - The possibility of downward vertical flow is not adequately evaluated by the model or by findings resulting from field activities of the Red Hill investigation. The majority of the monitoring wells are screened across the water table and may miss contamination that follows the dip of the lava flows deeper into the aquifer. Table 2-13 shows that the model fails to capture the magnitude of the vertical hydraulic gradients measured in the West Bay and nested wells indicating the model also doesn't test this potential migration pathway. The degree of vertical groundwater flow may be an important consideration in future well design (e.g., place screens deeper) and contaminant migration beneath low permeable barriers. Greater insight into the potential vertical pathway can be gained by optimizing on-going investigative efforts. Currently UH's tracer test shows a strong tracer signal at West Bay Well RHMW15 Zone 5a. Only the water table zone of RHMW15 is currently being monitored. Collecting tracer samples from the deeper ports of RHMW15 provides a unique opportunity to develop a vertical profile of the tracer plume and increase our understanding of the groundwater transport processes in three dimensions. DOH recommends sampling all zones of RHMW15 on at least one sampling round followed by augmenting the water table sample with at least one deep sample that rotates through the deeper zones during successive rounds.

Navy Response -

Please see response to comment 7, above. The Navy is not sure whether the response at RHMW15 is stronger than the response at RHS, because significant dilution occurs with water drawn in from other areas by RHS, but agrees that the strongest preliminary signals are at these monitoring points. This is generally consistent with the modeled flow trajectories, but the Navy recommends letting UH complete its work before drawing any final conclusions. We also note that enhancements to the vertical analyses have been underway since the 2024 modeling reports and were a point of discussion during the 2025 Modeling Workshop. The Navy is available to discuss some of the improvements that have been made and are still being worked on.

As to the suggestion to sample multiple zones of RHMW15 for dye, the Navy has already conducted the sampling and would be happy to discuss these results and any specific recommendations DOH may have.

9. Hydraulic conductivity

a. DOH previously provided comments that the basalt horizontal conductivity is much higher than literature values (Vol. 1, PDF page 49). The NCTF-RH response to those comments conveys uncertainty about hydraulic conductivity at the Red Hill site and scale, citing Oki (2005), "Heterogeneity in the ground-water system likely exists but is currently poorly understood. Values assigned to model parameters generally were based on existing estimates. However, some of these parameter values may be poorly known."

b. The Navy Response also states, "In addition, in reviewing transmissivity rather than hydraulic conductivity, which eliminates the need for an assumption of saturated thickness, the estimates in Table 1 from Hunt Jr. (1996) are as high as 8.4E6 ft2 /d, with many values in the range of 0.3E6 to 2E6 ft2 /d." However, Table 1 from Hunt (1996) does not list transmissivity. This table lists hydraulic conductivity, with the highest value being 5,000 ft/day. Please clarify.

c. The Navy Response also states that using conductivity values closer to literature values, "did not fit the extensive local Red Hill dataset as well as the calibrated value." Please clarify and provide an explanation as to how adjustment of other modeling components would not be more appropriate and improve the accuracy of the model. Calibration may improve the fit of the parameter without gaining overall model accuracy. DOH requests that the NCTF-RH conduct a more in-depth evaluation to reconcile the currently selected hydraulic conductivity and anisotropy values with those used in previous scientific investigations.

Navy Response –

After stating that "some of these parameter values may be poorly known", Oki went on to suggest an approach that is almost identical to what the Navy employed:

> "Improved estimates of the distribution of hydraulic characteristics in the study area can be obtained using controlled aquifer tests as well as by careful monitoring of pumping and water-level conditions throughout the aquifer. Accurate pumping data in conjunction with water-level drawdown and recovery data can be used for calibration of a numerical ground-water flow model, particularly during periods when recharge does not vary."

The Navy has conducted several basalt horizontal conductivity modeling sensitivity studies and is available to discuss this further with the DOH.

The Navy does not recommend simply inputting other values assumed by other models that involved different models, scales, issues, data sets, and study questions, especially when the authors acknowledged the values they used were based on their judgement and not derived

directly from site-specific field studies. Please also note that the Navy has evaluated many of those other modelers' values in the sensitivity analyses. Red Hill has a ***substantial*** data set, covering many years, many site-specific monitoring locations, and several different pumping tests at Red Hill and in the vicinity over the years. The Navy does not recommend strict adherence to other estimated values, which among other things may not be compatible with the DOH's requests over the years to match the site-specific data. The Navy would be happy to discuss the sensitivity studies that have been conducted and any specific constructive suggestions the DOH may have regarding more in-depth evaluations.

Regarding the transmissivity, the reference should have been to Table 2 from Hunt Jr. (1996).

10. Sensitivity Testing - Section 2.5.5 Scenario 5 – Northwest Flow Direction Scenario – A scenario was run to evaluate the possibility of a northwest flow component. This is important as it evaluates the potential risk that contamination at the Facility may pose to Honolulu Board of Water Supply assets. The apparent method of testing was to reduce the weathered fringe beneath the saprolite/valley wedge, then execute model calibration runs to get the best match between the simulated and measured groundwater elevations. A significant increase in the longitudinal and vertical hydraulic conductivity and an increase in the horizontal anisotropy for the basalt were needed for the groundwater elevations to match. The results were evaluated as having a significant impact on the calibration due to a poorer match between the simulated and measured vertical head differences. We did not find a table that specifically stated the simulated vertical head differences for this scenario. A review of Table 2-13 shows that the base case scenario failed to capture the downward vertical gradients measured in all cases. It would seem logical that a simulation that is representative of a northwest flow component would produce downward gradients in the West Bay Wells along the southeast margin of South Hālawa Valley. We are unable to evaluate what the model produced. Regardless, to properly test the southeast to northwest flow hypothesis versus the down-ridge flow trajectory will require a much more rigorous model that incorporates the likely hydrologic parameterization that would result in a northwest flow trajectory. This parameterization would include decreased caprock hydraulic conductivity, increased flow from into the southeast boundary (as postulated by Mink, 1980), and potentially decreased flow into the northeast boundary. The scenario that tests for a northwest flow component is incomplete.

Navy Response -

The Navy is available to discuss the sensitivity studies that were conducted and will consider this comment in future modeling efforts. All the results of Scenario 5 can be found in Appendix C (PDF Page 1,109 of Vol. 2), in the form of simulated vertical head difference plots. Vertical gradients for the post-2024 modeling efforts were one of the specific topics of discussion in the 2025 Modeling Workshop, and have been the subject of further analyses, which will be continued and incorporated into the Enhanced Model and could be a future topic of discussion.

11. Anisotropy

a. Previous DOH comments and the Navy Response discussed anisotropy. The DOH questioned why the anisotropy values used in the groundwater flow model are an order

of magnitude higher than literature values. The NCTF-RH responded that the anisotropy values are the best calibrated values. DOH desires a more robust assessment of the disparity between the groundwater flow model anisotropy and the values used by previous modelers. The NCTF-RH's selected anisotropy is the key parameter that results in particle tracks going directly down the Red Hill Ridge and oblique to the simulated groundwater contours. Vertical anisotropy is a key model parameter for evaluating the effectiveness of the alluvial/saprolite barrier and potential deeper migration of a contaminant plume due to the dipping lava beds. Anisotropy cannot be directly measured, so a rigorous assessment using the available evidence is warranted.

b. Anisotropy estimates based on drawdown (Section 2.1.3.1) are problematic due to the uncertainty of where exactly water was being pulled from within the shaft. The estimates of hydraulic conductivity and anisotropy from drawdown data using analytical solutions are limited because they are applicable for the geometry of a pumping well and not a shaft. The limitation of these applied methods should be more fully considered for estimating values of hydraulic conductivity, horizontal and vertical anisotropy, and storage coefficients.

c. The estimated value of the vertical anisotropy (ratio of vertical to horizontal hydraulic conductivity) also needs additional justification. Table 2-18 lists the "Sensitivity Analysis Results" showing the Vertical Anisotropy Ratio of 50 was used in the base calibration and the vertical anisotropy did not impact the calibration significantly (Vol. 1, PDF page 79). Therefore, vertical anisotropy appears to be insensitive during flow model calibration. What are the possible reasons for this behavior in the current version of the flow model? One possibility that needs to be considered is that the model does not adequately capture groundwater movement downward and beneath the geologic barriers.

d. The simulated anisotropy has a dominant influence on the NCTF-RH's conclusions that RHS provides capture of contamination beneath the USTs when pumping, and that a conservative solute will pass down the Red Hill Ridge, flowing beneath Aiea Bay, and eventually being captured by Kalauao Springs (Figure 2-20 and others) when RHS is not pumping. These flow trajectories depend heavily on a very high horizontal anisotropy that is much greater than what has been historically used in groundwater models of the area. For this reason, it is important to critically evaluate and accurately estimate values of the anisotropy and hydraulic conductivity to ensure the selected values are valid instead of only using PEST-generated best fit values. It is well known the parameters estimated by models are not unique and different sets of parameter values including anisotropy can provide the same calibration match. The values used in the NCTF-RH model are quite different from those used in previous models (Oki, 2005; Rotzoll and ElKadi, 2007; Gingerich and Voss, 2005). It is not DOH's position that the selected anisotropy values are necessarily incorrect, but due to its importance, the stated modeling conclusions must be better supported.

Navy Response -

Please see response to comment 9. Modeling of anisotropy for the post-2024 modeling efforts was one of the specific topics of discussion in the 2025 Modeling Workshop, and has been the

subject of further analyses, which will be continued and incorporated into the Enhanced Model. We also note that several sensitivity studies evaluated the effects of anisotropy, and the Navy is available to discuss the details and welcomes any constructive suggestions.

12. Borescope data

a. Defining groundwater flow paths is critical to understanding contaminant transport and for designing remediation approaches. It is very challenging to constrain groundwater flow paths with a high degree of certainty, so all available approaches need to be seriously considered. The NCTF-RH emphasizes the model matches observations (Navy Response, pages 6, 8, and others) and the limitations of the colloidal borescope data collected by UH and U.S. Geological Survey (USGS), suggesting the borescope surveys provide little meaningful insight into groundwater flow paths. DOH agrees that caution needs to be exercised in using borescope data to estimate spatial groundwater flow trajectories and velocities. However, we disagree with the statement "it is not technically valid to use these highly variable local velocities to infer groundwater flow directions over a larger region." The borescope results need to be evaluated as a suite of data rather than individual datapoints. The application of borescope data to the Waimanalo Gulch Landfill, when viewed as a suite of data, produced results that agreed with a regional model and captured the deflecting influence of the coastal caprock.

b. The Navy Response further states that, "Significant differences in groundwater flow measured by a borescope can occur over short vertical distances in a well, sometimes over as little as 2 feet" (PDF page 6) and "Significant differences in apparent flow can occur at the same location over the course of 1 day" (PDF page 7). These statements suggest that the NCTF-RH does not plan to incorporate the borescope results in future model development. DOH agrees that borescope measurements show significant spatial variability. However, it is DOH's position that the borescope survey data in conjunction with other indicators of groundwater flow direction may provide valuable insight into groundwater flow paths in the Red Hill and Halawa Regions. DOH would like clarification from the NCTF-RH on how, if at all, they plan to incorporate the borescope data into the groundwater model development and verification process.

Navy Response -

The Navy agrees that colloidal borescope measurements show significant spatial variability but might still provide some insight in conjunction with other indicators of groundwater flow. Unfortunately, the Red Hill environment is likely far more complex than the near-shore site in Waimanalo, which likely accounts for the higher variability in the limited preliminary UH data that the Navy has been provided to date. In any case, the Navy recommends letting UH complete its interpretations and analyses before drawing any conclusions.

13. Concerns with Validating the Model using Post May 2021 TPH-d Data

a. The NCTF-RH improperly uses TPH-d (total petroleum hydrocarbons – diesel range) data collected following the May 2021 release to validate the model results showing particle pathways going directly down the Red Hill Ridge with 100 percent capture by RHS (Navy Response, page 7). The DOH previously commented that 2021 fuel releases indicated a contaminant migration to the northwest. In reference to Section 5.2.4 –

Forensic Analysis of TPH-d in Tank Farm Wells RHMW01R, RHMW02, and RHMW03 (Vol. 1, PDF page 138), the NCTF-RH uses highly processed TPH-d breakthrough curves (Charts 5-3 and 5-4) to conclude that these data show that groundwater flow and contaminant transport is directly down the Red Hill Ridge following the May 2021 fuel release. The NCTF-RH does acknowledge that there are many inconsistencies between the observed data and hypothesis that TPH-d data collected after the May 2021 release shows a contaminant plume moving down the axis of the Red Hill Ridge. This hypothesis seems to be counter to the stated conclusion that all but 100 gallons of fuel from that release were recovered (Vol. 1, Section 5.3.1, second paragraph). The major issue in this section is that, in the last paragraph, the TPH-d at RHMW03 is removed from consideration. DOH contends that it is improper to remove this important line of evidence. It is also important to note that the label for the Y-axis in Chart 5-3 is incorrect. The Y-axis shows normalized concentrations and thus would be unitless. Chart 5-3 makes it appear that the TPH-d concentration at RHMW03 was about the same as at RHMW02 which is misleading. The TPH concentration at RHMW03 was much less than at RHMW02. DOH's major concern is that so little is known about where the TPH contamination reached the water table and what role infrastructure, washwater, and vadose zone transport had on the contaminant migration, which makes using these data to validate the model inappropriate.

b. The NCTF-RH also responds that many distal TPH detections are not valid (NCTF-RH Response, page 7) but does not provide convincing evidence to support this claim. Positive detections interspersed by non-detections do not invalidate detections. A period of months between release and detections also does not invalidate detections, rather this likely shows the velocity of transport. With respect to the recent USGS carbon study, an uncontaminated well was estimated to contain 38% ancient carbon. Diesel range organics (DRO) and C-14 (of NVDOC [non-volatile dissolved organic carbon]) were sampled from Red Hill wells and measured separately. NVDOC and DRO were compared, and there was a relationship between the two. However, this does not mean that 38% of DRO is background. The NCTF-RH has not performed adequate forensics to show that TPH detections are not fuel related. It is DOH's position that until sufficient forensics analyses are completed, all detections need to be considered when evaluating how well the model results replicate actual conditions at the Red Hill site.

Navy Response -

During a Special Purpose Meeting between the RAs and the Navy conducted on May 31, 2023, potential CF&T calibration data sets were discussed, and the RAs were specifically asked for feedback and recommendations on potential calibration datasets. The only potentially suitable (but imperfect) dataset identified by any of the three parties was TPH-d following the May 2021 release. This was discussed again in the 2025 Modeling Workshop. The Navy is not aware of any better datasets to use but still welcomes discussion of any specific recommendations DOH may have. We also note that normalizing concentrations and using moving averages is an accepted way to visualize patterns through the noise that is often present in environmental data. The raw data was also shown, such that nothing was being hidden. The Navy recommends further

discussion if there are other specific datasets showing breakthrough curves that DOH recommends using.

The Navy acknowledges statements made about migration to the northwest, and occasional TPH detections in certain wells, but is not aware of any data sets showing potential breakthrough curves confirming migration to the northwest. The Navy is also not aware of detections in the BWS' Halawa shaft over the past several decades. It is unclear how one would model sporadic low-level detections of TPH interspersed in an overwhelming majority of non-detections, or what the calibration metric would be. The Navy is willing to discuss potential datasets or approaches which the DOH may consider suitable for model calibration.

14. Water chemistry data contradicts modeled flow trajectories

a. Measured chloride concentrations cannot be accounted for by the modeled groundwater flow trajectories (NCTF-RH Response, page 10, item 3.b., last paragraph). DOH has pointed out that the chloride enrichment can come from down dip in the lava flows. This is important because, if the RHS is receiving a significant fraction of its water from the southeast or downslope, this decreases the ability of the RHS to capture or otherwise contain contamination in the upper part of the Facility.

b. DOH's concern is that the chlorides measured at RHS are much higher than would occur if the dominant groundwater flow to the RHS infiltration gallery was from upslope. Section 2.3.6, Evaluation of the Source Water with Chloride and Temperature Data, attempts to address the DOH concern by adding a layer to the bottom boundary and assign a General Head (head dependent flux boundary) and Prescribed Concentration Boundary conditions that set the chloride concentration to one-half that of seawater chloride concentration. Specified concentrations for chloride were also assigned to the southeast and northeast boundaries. The concentrations assigned are not representative. For example, assigning 104 mg/L to the south recharge zone is not realistic. It appears the water balance calculation for the southern recharge zone was based on rainfall and recharge values for the very southern part of this Zone. The value assigned to the North Recharge Zone is reasonable. However, the chloride concentration of 30 mg/L assigned to the northeast boundary is much higher than that reported to DOH and CWRM (Hawai'i Department of Land and Natural Resources, Commission on Water Resource Management) for wells in the upslope recharge area. Enclosed Figure 1 is a map of the distribution of naturally occurring chloride in the groundwater recharge based on chloride mass balance calculations and recharge estimates of Engott (2017). DOH's concerns are two-fold; first, the results shown in Table 2-17 are heavily dependent on the model accurately simulating the chloride contribution from the theoretical mid-point of the transition zone. If the model overestimates that contribution by even a small amount, that has a significant effect on the calculated contributions from the other boundaries. The actual relationship between RHS and the mid-point of the transition is very uncertain and an analysis that depends heavily on that relationship is not valid. Second, this approach does not capture the full freshwater/saltwater circulation system leading to an inaccuracy in the contribution from the theoretical mid-point of the transition zone.

c. DOH is also concerned with the applicability of the modeling packages utilized to accomplish this. Buoyancy packages in MODFLOW can simplify density effects by assuming constant or sharply separated densities and often rely on hydrostatic or steady-state conditions, which can lead to significant errors when salinity varies in space and time. These approaches cannot simulate mixing zones, dispersion, or density-driven instabilities, and they poorly represent vertical flow and transient responses to stresses such as pumping. As a result, they do not conserve salt mass and can poorly predict interface movement, upconing, and flushing behavior. In contrast, fully densitydependent packages like the Sandia Waste Isolation Flow and Transport (SWIFT) model explicitly couple flow and transport, avoiding these fundamental limitations.

d. Compared with SWIFT, the Groundwater Modeling System (GMS) Density Dependent Flow (DDF) package is an improvement over simple buoyancy formulations but remains a simplified representation of density effects. DDF accounts for variable density in the flow equation, improving simulation of pressure gradients and density-driven flow, but density is not fully coupled to solute transport and salt mass is not rigorously conserved. As a result, mixing zones, dispersion, and transient feedback between salinity and flow are only approximated or externally prescribed. Models like SWIFT, by contrast, fully couples variable-density flow and transport, providing a more physically complete and defensible representation for transient, pumping-influenced, or strongly saline systems. Please discuss the assumptions and limitations of the DDF utility and density-dependent simulations in MODFLOW6 for simulating particles traces in this area and plans for improving the modeling approach.

e. The question that remains unresolved for model validation is whether the additional chloride measured in the RHS is due to a small contribution from down dip of the lava bedding plane or from the southeast boundary suggesting a significant flow component laterally from the southeast rather than down the axis of the Red Hill Ridge.

f. The best resolution for this question is to revise the Red Hill conceptual model as more UH data comes in. Current UH results, particularly the tracer test and borescope data, provide an emerging picture that differs from much of what the current groundwater model indicates. Once the evaluation of the UH data is complete, strategies can be developed to resolve the differences between the model output and the hydrologic conclusions of the UH Study. This may require significant modifications to the model.

Navy Response –

The Navy agrees that the models can be reevaluated as appropriate once UH completes its work.

The Navy understands that the current modeling software and code, which has long been agreed upon, is commensurate with the groundwater impacts present at the de-fueled Facility. The density dependent flow package in MODFLOW-USG transport is a tightly coupled flow and transport model that iteratively solves the flow and transport equations and accounts for transport processes such as diffusion, dispersion, and advection (Panday, 2025). Mass is conserved in the calculations through the iterative coupling process. It is not a sharp interface model like the

Seawater Intrusion package (SWI2) in MODFLOW. While the Sandia Waste Isolation Flow and Transport (SWIFT) model simulates density changes as a fully-coupled flow and transport model, there are many limitations that would make this model less useful for this application such as lack of flexibility in grid design to simulate flow, heterogeneity, and transport at various scales; no connected linear network package to represent water supply tunnels; and inferior Graphic User Interface (GUIs), particle tracking, PEST, and python integration. As to GMS, that is a GUI, not a modeling code, and the DDF package is not part of the GUI. We are also not clear regarding the relevance of MODFLOW 6. The Navy generally agrees that deeper flow patterns with respect to pumping wells are uncertain, but any model used would likely have non-unique solutions in the regions below most of the monitoring well screens and observed groundwater impacts.

With respect to chloride near the model boundaries, the Navy will revise the concentrations used in future modeling with reference to these and other DOH comments.

15. Simulated Flux at the Northeast Boundary

a. The constant flux boundary condition at the northeast should be more fully examined. The estimated value of the flux has significant uncertainty and additional sensitivity analyses should be performed including the use of other types of boundary conditions. The use of a constant flux boundary condition can cause the flow model calibration to have inaccurate values of estimated parameters including, not only anisotropy, but also vertical and horizontal components of hydraulic conductivity. The flow model calibration should include additional parameter estimation simulations based on different types of specified boundary conditions for the northeast boundary.

b. There is insufficient data to know with certainty what the flow into the northeast boundary of the model domain is, so a range of flows need to be evaluated. The NCTF-RH assumes that all recharge that occurs between the northeast model boundary and the topographic divide flows into the model. The actual location of the Ko'olau Rift Zone groundwater divide is unknown. Also, groundwater may move laterally along the strike of the dikes in the marginal dike zone. Either condition could result in the actual groundwater flux into the model being significantly different than that currently simulated. Sensitivity model runs should be done to evaluate the impact of a range of the northeast boundary fluxes on the model results.

Navy Response –

The Navy agrees that there is (and likely always will be) insufficient data to know with certainty what the flow into the northeast boundary of the model domain is, which is why several sensitivity studies were run, including an evaluation of no inflow from the dike region. In the latest version of the GWFM, the flux at the boundary has been allowed to vary during calibration with the starting value derived from Engott (2017). Boundary conditions for the post-2024 modeling efforts were a topic of discussion in the 2025 Modeling Workshop, and have been the subject of further analyses, which will be continued into the Enhanced Model. The Navy

continues to conduct additional boundary condition sensitivity studies and welcomes further discussion or specific recommendations.

16. Simulated and Observed Water Levels - Observed and simulated water levels do not match key wells, including RHP02, RHP07, RHP04A, RHMW2254-01 (Vol. 1, PDF page 59). PDF page 58 states, "The lack of full recovery at these three wells is indicative of a subdued hydraulic connection compared to wells with full recovery. It is likely that the lack of full recovery and higher heads are due to a local-scale heterogeneity." We question whether the model is able to sufficiently represent the broader system without being able to represent local scale heterogeneity. RHMW02, RHMW03, RHMW04, and RHMW05 also do not show good agreement (see Vol. 2, PDF page 6560). Please provide Residual Mean Error and Mean Absolute Error to provide a measure of goodness-of-fit.

Navy Response –

The Navy respectfully maintains that the models do represent the broader system and there is reasonable agreement with field data. Other recognized modeling experts who conducted detailed reviews have commented that the "September 2024 groundwater flow model (GWFM) is the most comprehensive, best supported version of the model submitted to date and represents significant progress in developing an accurate, functional model of groundwater flow and contaminant transport at Red Hill."

As to the reference to page 6,560 of Vol. 2, that shows a hydrograph for the Ft. Shafter MW well in Realization 23, which does not seem to be relevant to this comment.

Residual Mean Error and Mean Absolute Error are summarized in Table 2-12 of the report. Minimum, maximum, average residual, and root mean square error are presented on Chart 2-3. Additional details are documented in Appendix D. The calibration metrics are generally well within acceptable values, but the Navy welcomes any specifics as to what values the DOH reasonably expects.

One of the benefits of a physically-based numerical groundwater model is that it must adhere to certain fundamental principles such as flow continuity and mass balance. Such principles constrain the model and prevent it from making physically inconsistent predictions, such as groundwater flowing in opposing directions within a small volume of the subsurface, or during a small interval of time. Fluid mass must be fully accounted for: mass that flows into a region must equal fluid mass that flows out of that region. The site setting is unusual in many ways: hydraulic gradients are exceptionally flat, hydraulic conductivities are unusually large, and variable between material types, and detections of petroleum chemicals in groundwater are infrequent in most locations away from the tank farm footprint. This setting may be described as having a low "signal-to-noise ratio" which makes it challenging to uniquely identify patterns and trends and leaves significant room for multiple interpretations of the same data. All models include deviations from field measurements and contain areas which, through lack of detailed information, contain simplifications and idealizations. Additionally, environmental systems often

have "outlier" data points, the existence of which do not invalidate a model. Given the extremely high number of data points as emphasized by DOH in the recent Fuel Tank Advisory Committee meeting, which may not all be able to be matched at once, the Navy suggests further discussions aimed towards setting realistic and achievable expectations that meet the overall needs in light of current and anticipated site conditions.

17. Calibration - The use of ranks for the model calibration targets may bias the model to estimate parameters primarily on the higher rank targets. Using ranks with a wide range likely makes the lower rank targets irrelevant. Additional justification should be provided for the ranking methodology. Provide additional justification as to why pilot points were not apparently used during model calibration. Given the very low hydraulic gradients between the tank area and RHS, a more accurate match of the measured hydraulic heads should be obtained during calibration. The use of pilot points may locally improve the match of the model in this area. The spatial distribution of the residuals (observed minus simulated head) should not be spatially correlated. We recommend using PEST with pilot points. Based on the Groundwater Vistas file for the calibration model, PEST with pilot points were not used to calibrate the model.

Navy Response –

Previous model iterations used PEST with pilot points. However, in Attachment B of the March 2022 joint agency letter reviewing the 2020 model, the RAs rejected the use of pilot points, requesting that the Navy:

"evaluate and implement alternate methods to represent subsurface heterogeneity. The subsurface in the vicinity of RHBSF is neither homogeneous such as represented in several of the models, nor does it demonstrate radially symmetric heterogeneity such as generally produced using the pilot point method. Alternative, structure-imitating, methods for representing subsurface heterogeneity in basalt settings should be considered that, while not deterministic, provide more realistic parameter fields and can be calibrated. Examples include multiple-indicator and multi-point (geo-)statistical methods that can be conditioned on local stratigraphic data such as that recorded in the barrel logs; and methods derived from sequence-stratigraphy to stack and accumulate lava sequences."

The Navy subsequently developed a method satisfying these prior recommendations. If the DOH now recommends otherwise, the Navy requests that the parties meet and agree upon the specific details of an approach that DOH and the other parties would consider reasonable. As such, the Navy welcomes a discussion of specific recommendations DOH may have regarding calibration and pilot points.

18. Reasonableness Evaluations

a. Navy Response, page 9, item 3.a.: During the 2015 Administrative Order on Consent Statement of Work meetings, it was expected that the model would not progress to the point that density dependent modeling would be needed. However, unexpected

groundwater flow trajectories that are currently simulated will require a density dependent flow model to evaluate adequately. With the RHS not pumping, a groundwater flow trajectory is simulated as going underneath Pearl Harbor, then making a "U" turn to Aiea. This modeled flow trajectory should be further assessed in terms of reasonableness. DOH's concern is that this unexpected, modeled flow path is an artifact of the modeling techniques used rather than an actual representation of groundwater flow in the Pearl Harbor area. The use of the DDF utility is dependent on fixed model boundaries and does not adequately test the validity of the unexpected flow trajectories. A thorough assessment is needed to evaluate whether the groundwater dynamics can be adequately simulated using a single density flow model.

b. Figure 3-4, Red Hill Shaft Flow Distribution shows the assumed flow distribution along the length of the RHS Infiltration Gallery (IG). It is based on a 1942 drawing by Harold Sterns that showed a significant increase of flow into the RHS IG after a shothole was blasted in the northeast end of the RHS IG. The NCTF-RH and many others assume that most of the flow into the RHS IG comes from the northeast, and the model was set up to reflect this. However, UH's tracer test brings this assumption into question. The dye concentration at RHMW15 – Zone 5a located near the northeast end of the RHS IG is nearly two orders of magnitude higher than at the discharge of the RHS pumps. If most of the flow captured by the RHS pumps originated from the northeast end of the RHS IG, the reduction in dye concentration between the RHMW15 – Zone 5 and the RHS pump discharge should be much less than two-orders of magnitude. Also, when the Navy sent an ROV into the RHS IG, the ROV operator stated in their presentation that the flow down the RHS IG ceased prior to them reaching their termination point, which was significantly short of the northeast end. The default modeling scenario should be groundwater withdrawal distributed evenly along the length of the RHS IG rather than concentrated at the east end of the RHS IG.

Navy Response –

In light of groundwater conditions at the defueled Facility, and agreements made over many years between the Navy and the RAs, the Navy does not recommend creating a new, density-dependent model at this point; the current contract task order does not support such an allocation of remaining resources. Moreover, any such modeling would likely be highly speculative given the size of the model domain and the location of the screened intervals of most of the monitoring points.

As to the groundwater withdrawal distribution throughout the RHS, the Navy did not restrict the distribution in the homogeneous GWFM used for calibration and it was therefore not used as a calibration target. The well boundary condition was placed at the western end of the shaft, where the pump is, and flow enters the shaft based on the gradients and properties between the aquifer and the shaft. Flow into each segment of the shaft is tracked for presentation purposes. For example, page 13 of Volume 2 of the 2024 report indicates that flow into the shaft for the homogenous model is distributed throughout most of the water development tunnel (including 25% from the western segment, 44% from the eastern segment, and 31% from the eastern clinker segment). The distributions are also shown, for example, for the sensitivity studies, which confirm contributions of varying magnitudes throughout much of the water development tunnel

(see, for example, PDF page 222). Concentration of flow towards the eastern end of the shaft occurs in the heterogeneous realizations because the realizations are conditioned on boring logs and the drawing by Stearns (1942) which shows more clinker towards that end of the shaft. The Navy does not recommend forcing an even distribution throughout the water development tunnel but is open to discussions if the DOH does not recommend modeling the stratigraphy mapped by Stearns. If necessary, the Navy recommends a constructive discussion between the Navy, DOH, EPA, and their modeling experts to try to come together and agree upon reasonable and achievable expectations to meet the needs of the project.

19. Conclusions Beyond Demonstrated Certainty

a. The Navy's Response draws conclusions before the model is complete such as:

- ***The models indicate that most constituents originating at the Red Hill Bulk Fuel Storage Facility Tank Farm migrate to RHS when the shaft is pumping, but do not migrate to Honolulu Board of Water Supply's Halawa Shaft under foreseeable pumping conditions. This information is of great value, suggesting that Halawa Shaft is not threatened by past releases at the Tank Farm.***
- ***This assertion cannot be made from the groundwater flow model in its current form due to its inability to represent observed conditions based on corroborating data.***

b. Additionally, Vol. 1, PDF page 4 states, "Results generally showed particles migrating to RHS with upwards of 90% capture when pumping at either 4 or 1.8 million gallons per day (mgd), which was largely independent of conditions at other pumping wells." As UH's dye tracer study proceeds and additional data becomes available, the flow paths do not appear as straightforward as this statement implies. Furthermore, if this was the case, then we would expect higher concentrations of the dye to appear at RHS sooner. Until the model shows reasonable agreement with key observations, conclusions like these are premature.

Navy Response –

The Navy respectfully disagrees that there is an inability to represent observed conditions, as discussed in Response 16, and the modeling results do not conflict with any data from the Board of Water Supply's Halawa Shaft that the Navy is aware of. The Navy welcomes a more detailed discussion of this assertion, which should be evaluated with specific reference to, and reasonable expectations regarding, the vast dataset available.

With respect to UH's studies, please see prior comments. Preliminary data made available to the Navy from UH's tracer test may support flow in generally southwest directions, similar to those modeled. In any case, the Navy recommends letting UH complete its analyses before drawing any conclusions based on the ongoing tracer study.

20. Closing - We outlined critical questions at the beginning of this letter. In addition, some important questions remain, including:

- ***What is the 3-D groundwater velocity field that fluctuates over time in primary areas of contamination that are beneath the USTs and lower Facility area including the area impacted by the November 2021 release and Oily Waste Disposal Area?***
- ***What is the advective velocity of contaminant transport within the Red Hill area and natural biodegradation rate?***
- ***Does pumping RHS hydraulically contain the contamination present at the Facility?***

Navy Response –

The Navy expects to address the first two questions in the next phase of modeling. As to the third question, the Navy agrees with DOH's experts that concentrations are decreasing over time and biodegradation is reducing concentrations at the site. The data suggests that contamination is centered near the tank farm and not migrating offsite. It is conceivable (but not known for certain) that this condition might continue even without RHS pumping; in any case the data does not suggest a lack of containment of impacts.

References

Panday, S., 2025, USG-Transport Version 2.5.0: Transport and Other Enhancements to MODFLOW-USG: GSI Environmental, March 2025.

Engott, J. A., A. G. Johnson, M. Bassiouni, S. K. Izuka, and K. Rotzoll. 2017. *Spatially Distributed Groundwater Recharge for 2010 Land Cover Estimated Using a Water-Budget Model for the Island of O'ahu, Hawai'i* (Version 2.0, December 2017). Scientific Investigations Report 2015–5010. Revised December 2017. Prepared in cooperation with the State of Hawai'i Commission on Water Resource Management and the City and County of Honolulu Board of Water Supply. U.S. Geological Survey