UNITED STATES ENVIRONMENTAL
PROTECTION AGENCY
REGION IX
75 Hawthorne Street
San Francisco, CA 94105

STATE OF HAWAII
DEPARTMENT OF HEALTH
P. O. BOX 3378
HONOLULU, HI 96801-3378

March 17, 2022

CAPT Gordie Meyer, CEC, USN
Regional Engineer, Navy Region Hawaii
850 Ticonderoga St., Suite 110
Joint Base Pearl Harbor Hickam, Hawaii 96860 -5101

Subject: Disapproval of the Groundwater Flow Model Report

Dear Captain Meyer:

The U.S. Environmental Protection Agency ("USEPA") and Hawaii Department of Health ("DOH"), collectively the "Regulatory Agencies", have reviewed the *Groundwater Flow Model Report* ("GWFMR") dated March 25, 2020 submitted by the U.S. Department of Navy ("Navy") and Defense Logistics Agency ("DLA") to satisfy the requirements of Section 7.1 of the 2015 Administrative Order on Consent Statement of Work ("AOC SOW") for the Red Hill Bulk Fuel Storage Facility ("Facility") located in O'ahu, Hawai'i.

The Regulatory Agencies disapprove the GWFMR and its associated numerical models. The many deficiencies in the Navy's models have been discussed in detail throughout the modeling process and most recently in our May 2021 critique provided to the Navy and their consultants, followed by a summary critique by our subject matter experts given to the Groundwater Flow Modeling Working Group on October 18 & 19, 2021. These deficiencies are extensive and relate to foundational assumptions in the Navy's GWFMR that render the results unreliable for Agency decision-making regarding aquifer protection, as well as unreliable as an underlying basis to evaluate contaminant fate and transport (CF&T). A list of the deficiencies in the GWFMR is enclosed. We are providing the Navy and DLA an opportunity to cure the deficiencies identified and resubmit the GWFMR.

The Navy and DLA shall hold a meeting within 30 days of receipt of this letter with USEPA and DOH to discuss next steps as to which models should be carried forward and what model modifications should be incorporated. Next steps may include establishing specific technical groundwater flow model objectives and regular meetings with the Regulatory Agencies and other stakeholders to discuss modeling assumptions and approaches. Within 60 days of receipt of this letter, the Navy and DLA shall submit in writing the next steps with an associated timeline. Once the groundwater flow models have been refined, but no later than 90 days of receipt of this letter, the Navy and DLA shall revise and re-submit the GWFMR. The Navy and DLA shall summarize all changes made and relevant model run results in an addendum to the revised GWFMR.

If you have any questions, please contact us.

Sincerely,

Gabriela Carvalho
Red Hill Project Coordinator
US Environmental Protection Agency Region 9

Roxanne Kwan
Interim Red Hill Project Coordinator
State of Hawaii, Department of Health

Enclosures:
1. Attachment A Joint Agency Deficiencies on the Groundwater Flow Model Report for the Red Hill Bulk Fuel Storage Facility, dated March 25, 2020, delivered March 17, 2022
2. Attachment B – HDOH SME Deficiencies Identified, Red Hill Groundwater Flow Model Report, dated December 3, 2020, delivered to Navy March 17, 2022
3. Attachment C EPA SME Deficiencies Identified, Red Hill Groundwater Flow Model Report, dated November 10th, 2021
4. Attachment D - DOH Review: Navy Groundwater Flow Models & Related Issues with the Navy CSM for the Red Hill Facility, dated October 19, 2021

UNITED STATES ENVIRONMENTAL
PROTECTION AGENCY
REGION IX
75 Hawthorne Street
San Francisco, CA 94105

STATE OF HAWAII
DEPARTMENT OF HEALTH
P. O. BOX 3378
HONOLULU, HI 96801-3378

March 17, 2022

Attachment A: Joint Agency Deficiencies on the Groundwater Flow Model Report, Red Hill Bulk Fuel Storage Facility, dated March 25, 2020

<u>Background</u>
The primary objectives of the of the Groundwater Flow Model Report (GWFMR) is to refine the existing groundwater flow model and improve the understanding of the direction and rate of groundwater flow within the aquifers around the Facility, in order to evaluate risk to groundwater resources that may be posed by the Facility.

The Navy and DLA have expended considerable effort to further knowledge and understanding of the complex subsurface around the Facility since 2015, however, the resulting models in the GWFMR do not reflect site specific data and associated heterogeneity with sufficient accuracy to provide confidence in model predictions. During 2017, the Navy reviewed the previous groundwater model as reported in the Red Hill Bulk Fuel Storage Facility Final Technical Report (Rotzoll and El-Kadi [2007] as published in Navy [2007]). Based on discussions with Groundwater Flow Model Working Group (GWFMWG) members, development of an interim groundwater flow model commenced with the 2007 model, with refined and expanded lateral boundaries and other hydrogeologic data. However, the geologic detail associated with past and recent data collection, some of which was included in the Navy's 2019 Conceptual Site Model (CSM), was not incorporated into the 2020 Groundwater Flow Models (GWFMs) at an adequate degree of detail. For instance, CSM Figures 5-2 through 5-11 show interpretive geologic renderings of the heterogeneous subsurface geologic system; the GWFMs are geologically implausible as compared with this subsurface data.

In 2018, the interim flow model was developed, in part, to evaluate hydrogeologic system behavior and help identify data needs. In late 2018, in response to feedback from Regulatory Agency subject matter experts (SMEs) including an August 2018 "Top Ten Comments" summary regarding the CSM, interim model calibration, and representation in the interim model of the basalt dip and strike, saprolite, caprock, tuffs, and sediments –the Navy conducted several additional simulations using the interim flow model. Considering the complex and imperfectly known hydrogeologic setting and sparse data set, model development followed a multi-model approach to evaluate the potential importance of various features of the CSM on local flow patterns and enable the testing of alternative scenarios. Building on the interim flow modeling effort, the Navy presented several groundwater models in the 2020 GWFMR that incorporate various parameters and depict alternate potential groundwater flow patterns throughout the area of interest (AOI) encompassing Red Hill Bulk Storage Facility, Red Hill Shaft (RHS), and Halawa Shaft (HS).

Despite the expansion of the modeling efforts, the Navy's 2020 GWFMs exhibit many of the same limitations as did the 2018 interim models, and consequently do not provide the improvements sought. None of the models reflect site specific data and associated heterogeneity with sufficient accuracy to provide confidence in model predictions. Consequently, the GWFMR and the accompanying models that it describes require substantial improvement. Thus, the Regulatory Agencies' disapproval of the GWFMR.

Summary of Deficiencies

Some key concerns regarding the Navy models are summarized below and detailed in the attached Regulatory Agency SME technical memoranda. Any one of the deficiencies below would be sufficient grounds for disapproval, but taken as a whole, demonstrate the significant degree of model unreliability. While some of the deficiencies detailed in the attached SME memoranda stem from review of the Navy's 2020 GWFMR, the general themes were previously communicated to the Navy and its technical team during Technical Working Group (TWG) meetings, GWFMWG meetings, as well as in letters such as those dated October 29, 2018 (for the GWFM) and March 30, 2020 (for the CSM[1]). The DOH Safe Drinking Water Branch also detailed concerns on both documents in the *Assessment of Groundwater Flow Paths in the Moanalua, Red Hill and Halawa Regions, Revision 2* (July 11, 2019):

1. The hydrostratigraphic units, as represented in the GWFMs, are implausible and do not reflect the detail or characteristics of the system that will control Contaminant Fate and Transport (CF&T) and ultimately, risks posed by the Red Hill Bulk Fuel Storage Facility to the aquifer system.
2. Calibration and validation efforts over-emphasize drawdown and recovery matches at the expense of actual groundwater elevations. Consequently, while some of the Navy models demonstrate reasonable correspondence with drawdown and recovery data, they do not adequately represent groundwater elevations. Matching elevations is critical to ensuring the modeling represents water budgets and hydrogeologic behaviors to an adequate degree of certainty.
3. Further to (2), the GWFMR and GWFMs fail basic validation procedures, meaning they fail the testing intended to provide confidence in the modeling construction and results. The GWFMR uses non-standard techniques in the validation testing, and validation

---

[1] **Department of Health 3/30/20 Response to *Conceptual Site Model, Investigation and Remediation of Releases and Groundwater Protection and Evaluation, Red Hill Bulk Fuel Storage Facility, Joint Base Pearl Harbor-Hickam, Oahu, Hawaii***

charts misleadingly show a goodness of fit to measured groundwater elevations that cannot be replicated in the actual model results. The Navy's modelers admitted that their validation charts are in fact a superposition of transient modeled aquifer drawdown/recovery responses onto measured elevation data

4. The GWFMs are run in steady-state mode. As the Regulatory Agencies have pointed out on numerous occasions to the Navy and its modeling team, the conditions of interest are transient. Contaminant migration is transient, groundwater capture is transient, and other aspects of potential risk are transient. The Regulatory Agencies cannot accept steady-state modeling outputs for decision-making, although we recognize it can have value as a conditioning procedural step in the model construction and calibration processes.
5. Measured groundwater responses to pumping at RHS, which show differences between monitoring wells suggestive of hydraulic compartmentalization, are not represented in the Navy models. This is likely due to the absence of sufficient geologic heterogeneity and the unrealistic representations of the geologic system in the Navy models.
6. The Navy models do not adequately reflect local-area hydraulic gradient directions and magnitudes, tending to substantially over-estimate gradients along Red Hill Ridge as well as not reflect local gradients toward Moanalua Valley. These results imply a probability of hydraulic capture by pumping RHS that is inconsistent with actual observational data. The Navy GWFMs are non-conservative in this regard.
7. The Navy models use, with insufficient technical justification, parameter ranges and other inputs that are often outside the bounds of published Hawai'ian literature, including values used in the previous (Navy [2007]) groundwater model.
8. Likely flow paths, capture zones, and general flow patterns, that are simulated by the Navy models do not correspond with flow patterns that are implied by groundwater geochemistry. For example: chloride concentrations at several Red Hill monitoring wells are substantially higher than can be accounted for via the predominantly upslope source water that is simulated by most of the Navy models.

In summary, the current Navy models have not advanced the understanding of groundwater flow and dissolved constituent migration patterns within the AOI sufficiently to support risk management decisions. Similarly, the Navy models do not provide a basis, at present, for CF&T evaluations that are inherently more complex than groundwater flow. Calibration results for the collective set of Navy models implies that widely differing conditions cannot be distinguished from one another as more or less representative of actual conditions, and substantial uncertainties remain regarding overall groundwater flow directions across the AOI. Consequently, it is the Regulatory Agencies' position that capture zone predictions for RHS – one basis for the Investigation and Remediation of Releases (IRR) recommendations, tying to the Navy's conclusions about their ability to respond to releases and protect drinking water resources – are not reliable.

Next Steps

A premise of the Navy's multi-model approach was to explore a range of hydrogeologic conditions to improve understanding of groundwater flow and the underlying hydrogeologic system. At the present time no single model incorporates all features, events, or processes, that are likely to be important to accurately simulate groundwater conditions in the AOI. Given this, the Regulatory Agencies require a consolidation of CSM features into a smaller number of locally behavioral groundwater flow models with the intent that the consolidated model(s) will demonstrate sufficiently improved correspondence with field data. If successful, then the models can be carried into the next phase of modeling-based analyses to support CF&T modeling. In addition, to better improve on the reliability of modeling efforts, further site-specific investigation and characterization, such as in-well testing and inter-well tracer studies, should proceed. Modeling and site-specific investigation activities can occur concurrently so that remedial mitigation measures can be developed and implemented in a timely manner.

Therefore, based on the Regulatory Agency comments and observations to date, the Navy will implement the following improvements, and consolidation of, the Navy models as detailed in the attachments:

- Refocus near-term modeling efforts, including calibration and verification, within the AOI and in particular the Red Hill Ridge area.
- Revise model layering to improve the representation of valley fill and saprolite incision within, rather than deformation of, basalts.
- Revise representation of the (un-weathered) basalt aquifer to improve realism and reflect the general character of documented subsurface heterogeneity. This includes, but is not limited to, geostatistical evaluations of the distributions of key hydrostratigraphic units and incorporation into the models within the AOI as demonstrated by Dr. Matthew Tonkin (EPA SME).
- After completing the foregoing:
    - Consolidate models to identify a smaller set of models representing the most probable conditions within the AOI and document the hydrogeologic distinctions between them.
    - Justify through technical analysis any deviations from more commonly used and accepted parameter values, inputs, and assumptions.
    - Refine the representation of geology to better reflect subsurface heterogeneity which may affect flow.
    - Re-evaluate the (transient) capture zone analysis using the updated, consolidated, models.
    - Provide standard transient model validation evaluation, with no changes to boundary or other hydrogeologic conditions and using aquifer heads as the criteria for goodness of representation. Map view plots of modeled versus measured flow fields should also be produced as part of the validation procedures.
- Work concurrently:

    - Seek regulatory concurrence approval on short term, in-well testing program to gather additional data about inter-well connectivities and local water conditions, using recent USGS transducer study results and other lines of evidence.
    - Work closely with 3rd party subject matter expert, Dr. Matt Becker, to design a tracer study to collect data to be used to assess GWFM flow predictions and match with natural tracer data studies conducted by DOH.
    - Final verification of Regulatory Agency prioritized well locations, permitting and installation of the new wells

Further details regarding the technical recommendations above, together with additional review comments, are included in Attachments B, C and D. Re-submittal of the flow models and accompanying report for approval will require revising the model assumptions to address Regulatory Agency concerns.

HAWAII DEPARTMENT OF HEALTH TECHNICAL REVIEW

Groundwater Flow Model Report, Red Hill Bulk Fuel Storage Facility – Joint Base Pearl Harbor-Hickam,
Oahu, Hawaii, Dated December 3, 2020 Revision 01

Delivered to Navy March 17, 2022

Executive Summary

The Hawaii DOH subject matter experts (SMEs) have reviewed the *Groundwater Flow Model Report, Red Hill Bulk Fuel Storage Facility – Joint Base Pearl Harbor-Hickam* (GWFM Report, March 25, 2020). This report and modeling work are a deliverable required under Task 7 of the Administrative Order on Consent – Statement of Work (AOC-SOW, 2014). Its purpose is to refine the existing groundwater flow model (Rotzoll and El-Kadi, 2007) and improve the understanding of the direction and rate of groundwater flow within the aquifers around the Facility. Underlying this purpose is the requirement for an improved understanding of the hydrogeology that controls groundwater flow. The improved models are intended to serve as a primary basis for subsequent contaminant fate and transport (CF&T) modeling and associated evaluations of current and potential future risks that may arise from fuel releases at the Site. Finally, those risk evaluations will provide guidance on response and remediation strategies for potential future releases at the Site and inform subsequent changes to the Groundwater Protection Plan.

The groundwater flow models (GWFMs), are well constructed and have the potential for representing the regional scale conditions at the Site and surrounding area. However, Site-area data directly indicate that the GWFMs fail to represent critical aspects of the system at a local scale around Red Hill, as detailed further below. Site-area data indicate a high degree of complexity in the aquifer system behavior and it is this local scale that is most relevant to CF&T and risk evaluations. In our review of the GWFM Report, we have identified multiple deficiencies that render the model(s) unreliable for increasing our understanding of the direction and rate of groundwater flow within the aquifers around the Site and for related decision-making. Similarly, these deficiencies make groundwater capture and plume containment conclusions in the GWFM report equally unreliable. DOH's overarching concern is the lack of verifiable metrics to ensure that the model replicates the hydrogeologic dynamics with sufficient certainty to support response mitigation planning. The most significant of the model deficiencies are the following:

- The GWFMs do not adequately reflect local area Red Hill groundwater gradients, elevations and individual well responses to pumping stresses. Further, the models indicate flow paths and rates that are inconsistent with background groundwater solutes, such as chloride, that are natural groundwater tracers.
- The GWFMs do not adequately utilize the available geologic and hydrogeologic data to interpret hydrostratigraphic conditions at and near the water table which is the interval most relevant to CF&T and risk concerns. Relevant information in the CSM and other Navy technical materials has not been used to refine that portion of the hydrogeologic model which, in turn, should feed into the numerical model framework.
- The deficiencies above and other associated issues make the groundwater capture conclusions by pumping Red Hill Shaft at 4.65 million gallons per day unreliable. The Navy GWFM Report uses particle tracking to develop Red Hill Shaft capture zones that the Navy concludes in other AOC documents (e.g. Tank Upgrade Alternatives and Release Detection Decision Document) to

demonstrate that a release from the USTs will be directed to the Red Hill Shaft under normal pumping conditions. These particle tracks are driven, in part, by the modeled hydraulic gradients that we find are not adequately reflective of site monitoring data. As described briefly below and in detail in the body of these comments, the disparity between the measured and modeled groundwater gradients beneath the USTs, and between the USTs and the Red Hill Shaft cast significant doubt that the simulated capture zone actually represents the flow field to the Red Hill Shaft, and, more critically, places doubt on the utility of pumping the Red Hill Shaft as a release response measure.

- Beyond the unreliability of the modeled groundwater capture zones are the implications drawn from the steady-state method of analysis. Future potential fuel releases will migrate most rapidly during their initial release period and so any associated risks to the groundwater system will be time-dependent (transient). All mitigation measures under consideration (inclusive of groundwater capture) will need to address the time-dependent considerations of plume release and transport and how those vary with release scenarios and the hydrogeologic conditions in and around the Site. Steady-state approximations are inappropriate for this level of groundwater protection considerations.
- The suite of models in the GWFM report represent a multi-model approach, which is typically used to resolve uncertainty or define the sensitivity of hydrogeologic assumptions that control the behavior of the aquifer system. In turn, that should lead to interpretations about the most likely suite of conditions that define and represent the aquifer system; i.e., an improved base model(s). The GWFM report does not lead to an updated set of interpretations that eliminates non-viable assumptions and validates those most likely present. That resulting base model (or limited set of models) should be a substantial improvement relative to the past modeling work that was the starting point of this effort. This improvement has not been achieved.
- The GWFMs use certain parameters and distributions that are not supported by site or area data, nor past modeling efforts. For instance, low porosity values are assumed (relative to past modeling), but no technical justification is provided. That results in groundwater flow that is likely too rapid and transient capture by pumping that is too large. This example assumption would also result in unrealistic estimates of contaminant degradation rates. These examples are non-conservative and are not useful to decision-making unless definitively shown to be more reflective of actual conditions than in past modeling work.
- The GWFMs do not appear to weight the area nearest the Site as the most significant from a calibration standpoint, in spite of having the highest data density and quality of any other areas in the model domain. As noted in the attached memo, the GWFMs do not adequately represent conditions in the local area of key regulatory interest and the remainder of the model domain is not particularly relevant in light of this deficiency.
- The GWFMs use a variable range of boundary conditions that, at present, are not verifiable and cannot be validated, and whose effects have not been adequately tested over a plausible range of values and, hence, the model results can't be considered unique or definitive. These amount to hypotheses without the associated technical evaluations to determine which (if any) is most likely representative.
- Based on our review of the Navy's draft numerical models, the verification model runs exhibit the same general issues as the precursor models. They do not adequately represent groundwater

elevations or gradients, which implies fundamental issues with the underlying conceptualization and parameter distribution framework. Although the Navy's validation models appear to match transient stresses (pumping and recovery), as did the precursor model (2007), groundwater elevations are equally critical to understanding the hydrogeologic behavior.

- The GWFM report claims to have selected conservative parameters and approaches. We do not find that to be accurate in many cases. But more important with regard to the AOC objective noted above, the models should refine hydrogeologic parameters and distributions that are most likely representative of the system to further our understanding of its characteristics to serve as a basis for future contaminant fate and transport (CF&T) and risk evaluations. The modeling does not achieve that goal.

Because of the uncertainties in the validity of the GWFMs local-area structure and parameterization, the DOH cannot reliably depend on the resulting flow rates, trajectories and capture zones generated by the models. Fundamentally, the models have not advanced our understanding of the aquifer system as compared to prior modeling work (e.g. Oki, 2005; and Rotzoll and El-Kadi, 2007) and the work does not meet the objective of the AOC.

The overarching deficiency that DOH recognizes is that the suite of models described in that document lack verifiability of the simulated groundwater flow trajectories and rates resulting in simulated drinking water source capture zones that are unreliable. DOH requires that the Navy provide a field verification plan to confirm or refute the representativeness of the simulated groundwater flow trajectories as indicated by modeled particle tracks. There are many field tests that can be employed to test the GWFM results. We have recommended a limited suite of field-scale testing to expand the local-area understanding of the aquifer system behavior including: i) a statistically robust suite of borehole measurements of flow rates and directions in the vicinity of the Facility; ii) borehole dye dilution-rate measurements; and iii) a controlled pump test of Red Hill Shaft at 4.65 mgd with transducers placed in all available local area monitoring wells, coupled with local-scale multi-well dye tracers, to provide clear evidence of the rate and trajectory of flow beneath the Facility and its capture by Red Hill pumping. The Navy can choose to follow these recommendations or propose other tests to evaluate the representativeness of the GWFM results. The tests the Navy proposes shall confirm or refute key aspects of the GWFMs output including flow rates, directions, and the ability of the Red Hill Shaft to capture groundwater and possibly contaminants around the Site at the modeled pumping rate of 4.65 mgd.

Introduction

The *Groundwater Flow Model Report, Red Hill Bulk Fuel Storage Facility – Joint Base Pearl Harbor-Hickam* (GWFM Report) is a deliverable required under Task 7 of the Administrative Order on Consent – Statement of Work (AOC-SOW). The purpose of the AOC-SOW Section 7 is to "Monitor and characterize the flow of groundwater around the Facility". The groundwater model is a key deliverable for accomplishing this task. Task 7.1 – Groundwater Flow Model Report - states "The purpose of this deliverable is to refine the existing groundwater flow model and improve the understanding of the direction and rate of groundwater flow within the aquifers around the Facility." The effort dedicated to the groundwater flow model (GWFM) has gone well beyond "refining" the existing model. But, in so doing, have yielded modelled flow rates and trajectories that conflict with many of the most critical field-measured groundwater parameters around the facility and, hence are not credible without further validation beyond the model. Further, because the Navy will rely on the GWFM to justify the Tank Upgrade Alternatives decision proposing that pumping Red Hill Shaft will be able to capture, or otherwise contain, any fuel release within the confines of the Facility (Department of the Navy, 2019a), **it is imperative that the modeled groundwater flow trajectories and velocities are defensible and can be validated by field measurements.**

The AOC-SOW mandate is that the model will provide a more refined understanding of the groundwater system. In turn, that refinement will allow us to ask pertinent area/tank-specific questions about what might happen under a range of plausible release conditions from individual in-service tanks and how receptors might be impacted. As the model currently stands, there is a substantial disconnect between the modeled relative groundwater elevations within the Facility monitoring wells and currently available field data; this results in an array of interpretational conflicts and uncertainties, described below, such that the model more likely obscures, rather than informs, actual risk conditions which we believe will lead to poorly informed decisions.

General Review Comments

The models are well constructed. The engineers and geologists doing the modeling are among the best in the field. However, due to the complexity of the site, a critical weakness of the model is the absence of verifiability of many of the model inputs. To be informative for predicting contaminant migration and release response planning, the results must be shown to be consistent with actual groundwater flow trajectories. However, the model results show a number of conflicts with field measurements (e.g. water levels, water table gradients, concentrations of natural tracers, etc.) and, to date, there has been no independent assessment of groundwater flow trajectories against which the model can be tested.

The primary concerns that the Hawaii Department of Health (DOH) has with the GWFM are summarized below.

1. The model suffers from an over reliance on automated parameter selection and calibration-driven model zonation of hydraulic parameters. Greater attention should be paid to valuable field data collected that indicate groundwater flow trajectories other than those simulated. Examples are provided below and include: local gradients within the Facility monitoring wells that are inconsistent with the modeled groundwater flow trajectories; the diverse range of groundwater chemistry concentrations indicating a poorly mixed system with sluggish flow in the upper part of the aquifer; and the absence of a verified hydraulic barrier between the upper part of the Facility and the Halawa Shaft.

2. The modeling approach relies on boundary condition assumptions that can't be independently validated. Whereas there is a broad range of possible hydrologic conditions at the model boundaries, only a small fraction of those possible have been tested. There is currently no methodology available to determine the prevailing hydrologic conditions at the model boundaries and, because boundary conditions can't be verified with confidence, the model results can't be considered unique or definitive.
3. Insufficient attention is given to, arguably, the most important measured field data available: the relative groundwater elevations within the Red Hill Groundwater Monitoring Network. The Navy contends that relative water level elevations in closely-spaced wells are unreliable for determining groundwater flow trajectories. The Navy places greater emphasis on the groundwater elevation across the Moanalua/Halawa region, hypothesizing that the greater well spacing and differences in groundwater elevation provide a much more reliable water table map for flow trajectory analysis. This is counter-intuitive since a great deal of effort and expense has been applied to minimizing measurement errors within the Facility monitoring wells. Conversely, the observations wells in the Moanalua and Halawa regions used for model calibration are <u>not</u> tied to the same elevation reference point used for the Facility monitoring wells and no true vertical depth corrections have been applied. Hence, the level of confidence in the relative groundwater elevations between the Facility monitoring wells and the outlying observation wells is much less than that within the Facility monitoring network.
4. The hydraulic parameters used in the calibrated model diverge significantly from those used by experienced and respected hydrogeologists in previous South Oahu modeling efforts. Specifically, the values used for the vertical hydraulic conductivity of the basalt, the horizontal anisotropy of the basalt, and the basalt porosity in the Navy model differ, in some cases, by more than a factor of ten from those used previously in peer reviewed and published studies.
5. Qualitatively, the simulated GWFM trajectories fail to account for the distribution of chloride concentrations in the Facility monitoring wells. Chloride is a natural groundwater tracer and can be a more diagnostic indicator of groundwater flow trajectory than water levels if the chloride source zones and distribution are understood. The modeled particle tracks, which represent the simulated groundwater flow trajectories, conflict with the distribution of the elevated chloride concentrations in many of the Facility monitoring wells.

Specific Comments

Modeling Approach and Complexity

This evaluation considers the technical aspects of the GWFM as well as the philosophy behind the modeling approach. In 2011, Dr. Clifford Voss as Executive Editor of the Hydrogeology Journal wrote two essays on groundwater modeling (Voss, 2011 a and b). The Navy quotes from Voss (2011b) to provide independent support for their modeling approach. However, the quote from Voss (2011b) is incomplete and the entire paragraph from Voss (2011b) is provided below for greater context. The omitted phrases are underlined.

> "<u>In the view of this writer</u>, the best way to go forward with practical management is to rise above groundwater models as final products, and instead, empower hydrologists to provide advice by using groundwater models in simple ways that are intended to elucidate

> understanding. Pursuit of complexity in groundwater models intended for practical management is a diversion from the real work at hand."

Voss recommended against over-reliance on automated parameterization of groundwater models and arbitrarily assigning hydraulic parameter values for calibration point-matching as was done in Models 53, 54, and 55. Below are two paragraphs excerpted from Voss (2011a)

> "Whether warranted or not, whether useful or not, parameter estimation has become a major part of model creation and this evolution has been fueled by the recent wide availability of automatic estimation software. In some sense, this wide availability has promulgated greater fallacious use of groundwater models. Automatic estimation software is truly a wonderful convenience when used properly, but it is no more than a convenience—and it should not be the primary objective of a modeling analysis to use it.
>
> An error in zonation, assumed model structure, or in some value assumed for input parameters, will cause automatic fitting to generate errors in other parameter values. These erroneous values may be organized in a realistic-appearing spatial trend that some modelers naively accept as reality. How can reality of a trend or newly discovered model parameter zone be determined without further targeted collection of field data?"

The emphasis on automated parameterization/calibration should be replaced with a comprehensive review of the conceptual site model to better constrain scenarios for the various model runs. For example, Model 59 tests the model response to lateral inflow into the southeast boundary. It appears that the 10 million gallons per day (mgd) value was chosen arbitrarily and distributed uniformly along the southeast model boundary. An inflow of 10 mgd equates to the entire recharge from the adjoining Kalihi Aquifer. Is the Navy hypothesizing that there is no groundwater flow from the Kalihi Aquifer to the ocean? Further, any realistic assessment of inter-aquifer flow would be biased toward inland portions of the aquifer since the depth of valley fill and saprolite in the coastal plain is known to be much deeper than the bottom of the freshwater lens. This is one of several examples in which too little geologic thought and justification has been invested into the various model scenarios. A multi-model approach is typically used to test valid hydrogeologic hypotheses to tease out unexpected details and ask, "does this make more sense relative to the hydrogeologic system and behavior and how can it be demonstrated via the available data?" That, then, should lead to the most likely set of conditions that explain the system and agree with all available data. This does not seem to have happened in the current modeling effort and no specific model seems to rise above the rest.

The Representation of the Hydrogeology of the Shallow Aquifer Zone is Inadequate

The GWFMs do not adequately utilize the available geologic and hydrogeologic data to interpret hydrostratigraphic conditions at and near the water table. This interval is the most relevant to risk concerns and the follow-on contaminant fate and transport model because it is where released contaminants will first encounter the groundwater, whether from recharge dissolving hydrocarbons during transport through the vadose zone or from a non-aqueous phase contaminant plume resulting from a release. In the latter case, that plume will serve as primary, continuing source of the dissolved phase contaminants of concern. However, information relevant to conditions in the shallow water table, that were developed for the CSM and other Navy technical materials, has not been used to refine that portion of the hydrogeologic model which should be used to constrain the numerical model framework.

For example, the Navy's GWFMs depend on a high permeability layer in the shallow aquifer, that can move large amounts of water down the Red Hill Ridge resulting in rapid transit times from beneath the USTs to Red Hill Shaft. A review of boring logs for the monitoring wells within the Facility show no evidence of a spatially expansive clinker zone that lies at or just below the water table. Further, the dip azimuth stated by the Navy would preclude a preferential highly permeable path along the water table since the reported dips of the lava flows vary from 3 to 11 degrees (Department of the Navy. 2019b; Sections 5.1.2 and 5.1.3) and are much steeper than the dip of top of the water table. Rather than developing models that rely on a high hydraulic conductivity zone that extends from beneath the tanks to the Red Hill Shaft, the models should portray the much more likely scenario that require the groundwater flow in the shallow aquifer to move through or around lava flows that intersect and dive beneath the water table on the hypothesized path from the USTs to the Red Hill Shaft. To capture the importance of the shallow aquifer zone for contaminant trasport, the conceptual model of groundwater flow requires revision to reflect field measured conditions and that revised model should be reflected in the structural controls on groundwater flow in the GWFMs.

Fuel Transport and Potential Hydraulic Capture

Fuel transport following a release is driven by the gradients created by that release and their interaction with the hydraulic properties of the subsurface setting. Many of these geologic aspects parallel those used in groundwater modeling, but with significantly more complexity. In hard-rock settings, fuel transport can be further complicated by inter-connected pathways within the matrix that are likely to be present, but that cannot be easily characterized at relevant scales. The DOH SME's collective experience, coupled with literature studies, indicates that fuel releases will move rapidly and in directions that may differ from the prevailing geologic fabric of the subsurface materials. The DOH has observed at multiple on-Island sites that free product pathways may not be identifiable at common scales of sampling. This is further accentuated by the geographic sparsity of data points at the Red Hill Facility due to access limitations.

Fuel transport is more complex and heterogeneous than contaminant fate and transport in the dissolved-phase. The distance and directions of the dissolved-phase impacts depend on groundwater flow, dispersion, and other attenuative processes that will generally limit that migration distance relative to the fuel generating those impacts. The fuel transport, however, can be very rapid and heterogeneous. At present, the potential rates, directions and character of fuel transport through the vadose zone, to the water table and outward is undefined by any study or characterization work with which the DOH has concurred.

Before any fuel release mitigation measures can be considered by the agencies, the ranges of behavior of fuel releases must be defined and agreed upon by the agencies. Any mitigation action, such as the suggested hydraulic capture by pumping Red Hill Shaft, must first be put into context with that transient fuel migration behavior. For instance, if fuel transport is more rapid and or/distant than the short-term (transient) ability of pumping to capture that release, the mitigation measure will be ineffective at protecting distal receptors. The timing and dimensions of any mitigation measure must be placed in specific context with the rates, directions, and magnitude of potential releases from the Site. No final mitigation measures, inclusive of the proposed groundwater pumping and implied containment, will be considered by the agencies absent this linkage between fuel migration potentials and the associated transient effectiveness of any proposed measures.

Lastly, while some interim mitigation measures may be appropriate as first-steps in the protection of the area groundwater aquifers, all mitigation measures must be demonstrated through field-validation. The EPA Superfund program requires that all remediation measures must be demonstrated to be operating “properly and successfully” as a pre-condition to deed transfer (EPA, 2019). The DOH believes a similar level of validation is required for the Site to protect this sole-source groundwater resource. As noted by the EPA:

The phrase "operating properly and successfully" involves two separate concepts. A remedial action is operating "properly" if it is operating as designed. That same system is operating "successfully" if its operation will achieve the cleanup levels or performance goals delineated in the decision document. Additionally, in order to be successful," that remedy must be protective of human health and the environment.

As noted, the DOH has significant doubts that the remedy basis provided by the Navy’s GWFMs is reliable for decision-making, even if specifically applied only to groundwater flow and capture. Field demonstration is a relatively simple and straightforward endeavor to validate the concept and its design parameters, the modeled capture pumping rate of 4.65-mgd being key among those. Demonstration of groundwater capture as modeled is the first step, followed by demonstration that LNAPL migration will not escape that transient capture zone.

The Relative Groundwater Elevations in the Facility Monitoring Wells do not Support the Modeled Groundwater Flow Trajectories.

Development of an accurate understanding of the groundwater flow trajectory beneath the underground storage tanks (USTs), along with any pathways to the Red Hill Shaft, are critical to any risk assessment and contaminant plume containment plan. In the absence of physical tests such as a tracer test or borehole flow vector survey, the relative water level differences in the wells beneath the USTs and in the hypothesized migration path to the Red Hill Shaft Infiltration Gallery are the prime metrics for an evaluation of groundwater flow trajectories. These differences provide the hydraulic potential to move groundwater, and dissolved contaminants, from areas of higher hydraulic head to areas of lower hydraulic head. However, the Navy has characterized the small differences in groundwater elevations across the Facility monitoring wells as unreliable and chose instead to use the drawdown response in the individual wells as the primary calibration parameter. The key questions for evaluating the groundwater flow model becomes: are the relative elevations between wells across the Facility monitoring wells so unreliable as to be dismissed; and should the priority for groundwater flow path analysis be placed on the relative differences in groundwater elevations between the Facility monitoring wells and the outlying observation wells in the Moanalua, and Halawa/Aiea area? It is DOH’s position that the relative groundwater elevations within the Facility monitoring well network are most important for groundwater flow path analysis as it relates to risk assessment and plume capture evaluation.

The regulatory agencies have expressed concern about the Navy’s failure to meaningfully address the local gradients within the Facility monitoring well network multiple times in last two years including in an agency letter to Navy in 2018 (EPA/DOH, 2018a), a presentation to the Red Hill Groundwater Modeling Working Group (EPA/DOH, 2018b), and a DOH report on probable groundwater paths in the Red Hill region (DOH, 2019). The weaknesses in the Navy’s interpretation of the relative groundwater elevations and implied gradients are still present in the current GWFM Report and are an unacceptable deficiency that precludes DOH’s reliance on any conclusions drawn from the models.

A review of Model 54 results offers informative insights into the models' deficiencies. Model 54 is an alternate parameter model developed to determine whether greater flexibility in the assignment of hydraulic conductivity to the basalt aquifer could capture localized variations in the observed water levels. Model 54 was selected because the overall calibration of this model was very good and similar to the other models presented in the GWFM report. Model 54 closely replicates the elevations across the model domain, but within the Facility monitoring wells, the model results are <u>not representative</u> of the observed groundwater elevation gradient changes that occur in response to changing pumping stress at the Red Hill Shaft.

Figure 1a and 1b show the observed groundwater elevations compared to those simulated by Model 54. The observed groundwater elevation values for Stress Period 1 (SP1, Red Hill Shaft pumping at an average rate) and Stress Period 3 (SP3, Red Hill Shaft not pumping) were taken from Figures 3.1-2 and 3.1-3 of the GWFM report respectively. The simulated groundwater elevation values were computed by subtracting the Model #54 residual mean error in Figure 5.4-5 from the target hydraulic heads in Figures 3.1-2 (SP1) and 3.1-3 (SP3). Figures 1a and 1b reflect the Navy's priority of calibrating the model to match the measured groundwater elevations across the model domain from the Moanalua Ridge to Aiea. This agreement between modeled and measured groundwater elevations is indicated by the high coefficient of correlation with data points falling along the 1:1 observed versus simulated line shown by the green dashes. However, as described below, the model performs poorly when simulating the response of the Facility monitoring wells to changes in pumping stresses at the Red Hill Shaft.

The Navy emphasizes that under normal pumping conditions at the Red Hill Shaft, the model indicates capture of water from beneath the USTs (Page 5-34, Lines 17 and 18). Using this conclusion, the Navy proposes pumping the Red Hill Shaft as means of capturing any fugitive contamination that may be released from the Facility. DOH used data from the GWFM Report to evaluate the Navy's hypothesis. For capture to occur, a hydraulic gradient needs to exist along a line from the USTs to the Red Hill Shaft, a line that includes wells RHMW03, RHMW02, RHMW01, and RHMW05. Figure 2 shows the measured and modeled groundwater elevations for the wells along the centerline of the Red Hill Ridge (a) and along the northwest boundary of the Facility (b). The groundwater elevations for SP1 (Red Hill Shaft pumping) are shown in dark grey while the groundwater elevations for SP3 (Red Hill Shaft off) are shown in violet. The measured groundwater elevations are shown as diamonds whereas the modeled groundwater elevations are shown as squares. Best fit lines are shown for measured (solid lines) and modeled (dashed lines).

The key observation for Figure 2(a) is that the measured response to changes in pumping stresses in the groundwater elevations is markedly different from that modeled. The slope of the best fit line for the **measured** groundwater elevations is **the same** for the **Red Hill Shaft off (SP3)** and the **Red Hill Shaft pumping (SP1)** showing no increase in the hydraulic gradient between a non-pumping and pumping condition. By contrast, the best fit lines for the **<u>modeled gradients</u>** for **both SP1 and SP3 show**: **a much greater slope than that for the measured groundwater elevations**; and**, a much steeper slope** for the Red Hill Shaft pumping stress period (SP1) than for the Red Hill Shaft off (SP3). The modeled groundwater elevations show the gradient that is necessary to move groundwater down Red Hill Ridge along the flow trajectory indicated by the particle tracks. Because there is **no change in the slope** of the best fit line for the **measured water levels between the Red Hill Shaft pumping and non-pumping conditions**, we conclude, contrary to the modeling results, that **the groundwater beneath the USTs and along the path to the Red Hill Shaft is not significantly mobilized by normal pumping of the Shaft.** We further conclude that the modeled particle tracks are not representative of actual groundwater flow trajectories during pumping and non-pumping conditions of Red Hill Shaft.

Figure 2(b) performs the same evaluation for the wells along the northwest boundary of the Facility (RHMW04, RHMW06, RHMW11, and RHMW08). **There is a definitive response in the measured gradients observed** in the northwest wells between Red Hill Shaft pumping (SP1) and not pumping (SP3). The **critical observation for these wells is the apparent reversal in the observed gradient from downslope when the Red Hill Shaft is pumping to upslope when the Red Hill Shaft is off**. By contrast, the **modeled groundwater elevations show the gradient going downslope under both pumping and non-pumping conditions,** and show no reversal, but only a steepening of that gradient, when the Red Hill Shaft is on.

The divergence in the measured and modeled responses to changes in pumping conditions at the Red Hill Shaft cast serious doubt on the ability of the model to predict a capture zone for the Red Hill Shaft. Hence, these models, as presented, can’t provide a reliable groundwater flow field for the contaminant fate and transport models. **These discrepancies indicate the modeled groundwater flow trajectories beneath the USTs are neither valid nor sufficiently reliable to guide response and remediation strategies or for use in assessing risk associated with future releases.**

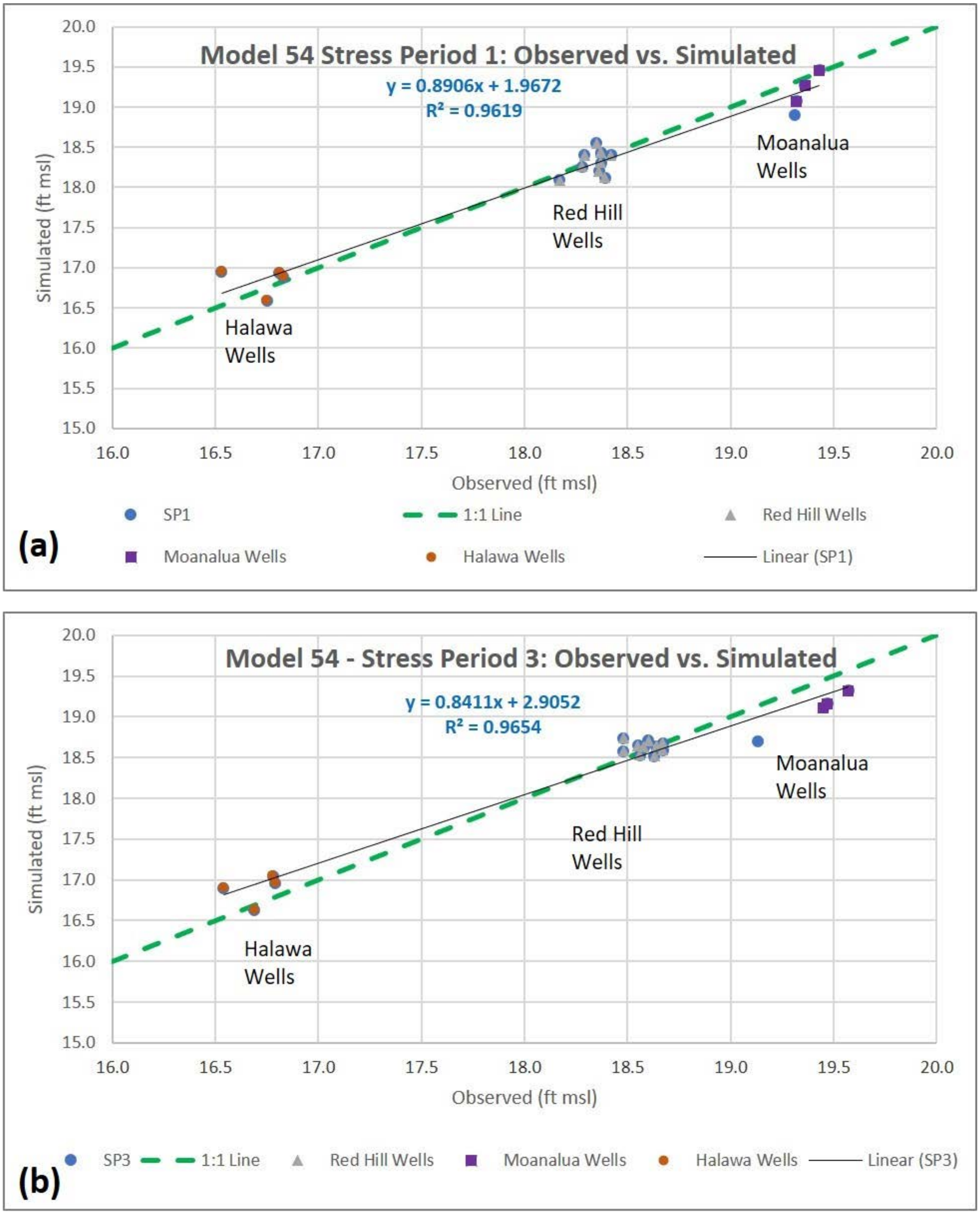


Figure 1. A comparison of the observed and modeled water levels across the model domain for Stress Period 1 (a) and Stress Period 3 (b) as simulated by Model #54.

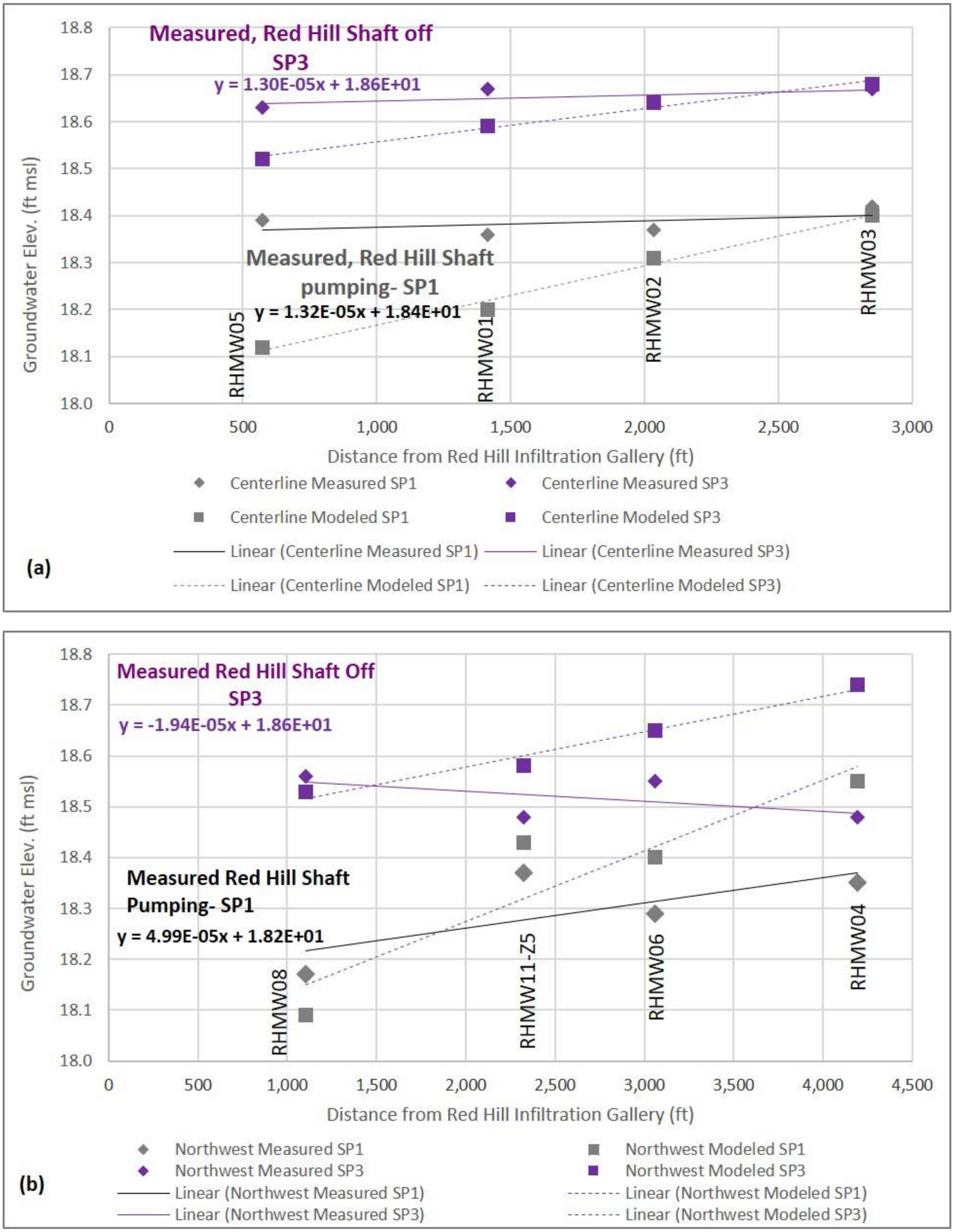


Figure 2. Groundwater elevation measured, modeled, and best fit lines for the wells along the centerline of the Red Hill Ridge (a) and along northwest side of the Red Hill Ridge (b)

Whereas the Navy has claimed that these gradients should be disregarded, no rigorous analysis is provided to support that position. The rationale given for disregarding local gradients is that the "datum or borehole alignment inaccuracies and the low precision of the gyroscopic corrections" (Page 3-2 lines 33 to 34) render the water level measurements unreliable. A review of the gyroscopic directional survey data sheets (Wellbore Navigation, Inc., 2017) show the resolution of the gyroscopic corrections for true vertical depth is 0.01 ft. A review of the top of casing leveling survey (Department of the Navy, 2018) shows that mean misclosure was 3.9 x 10^{-7} ft per ft of survey loop length. The greatest misclosure (Leveling Loop 4 from RHMW07 to RHMW06) was -0.002 ft over a loop distance of 2,372 ft or 5.6 x 10^{-7} ft per ft of loop length. The apparent uncertainty in the relative top of casing elevations is extremely small relative to the differences in groundwater elevations within the Red Hill Groundwater Monitoring Network. For example, the apparent gradient from RHMW03 to RHMW05, about 2,300 ft, is 1.3 x 10^{-5}, about 20 times larger than the uncertainty in the top of casing elevations.

The simple conclusion is that the modeled gradient going down the axis of the Red Hill Ridge accurately shows what is required to move groundwater in this direction (under the modeled hydraulic conductivity conditions described in the model) to meet the demands of the Red Hill Shaft pumpage and the assumed mauka-to-makai groundwater flow. The fact that the **measured gradient** going down the axis of the Red Hill Ridge **is nearly an order of magnitude lower** shows that the underlying assumptions of the model are incorrect. Further, the fact that the **measured** gradient going down the axis of the Red Hill Ridge changes minimally between pumping and non-pumping conditions shows that pumping the Red Hill Shaft can't be depended upon to contain any fugitive contamination from the Facility. The groundwater gradient in the wells along the northwest and southeast side of the Red Hill Ridge do show a response when the Red Hill Shaft transitions from a normal pumping condition to no pumping. This suggests that rather than capturing water from beneath the USTs, the water flowing to the Red Hill Shaft is drawn more from the periphery of the Red Hill Ridge rather than from beneath the tanks. Therefore, the proposed strategy of pumping Red Hill Shaft to capture a future release is uncertain/unlikely and will be an ineffective strategy for contaminant management.

Evaluation of Modeled Groundwater Flow Paths using Chloride Distribution

DOH has suggested on numerous occasions (e.g. DOH, 2018b; DOH, 2020; and the Technical Working Group No. 24 Webinar on March 5, 2020) that the distribution of chloride in the groundwater beneath the Facility can be informative when evaluating the groundwater flow trajectories simulated by the GWFM. It appears that the Navy may have given this approach some consideration because, on Page 1-4, Lines 6-8 of the GWFM Report, the Navy indicates a chloride calibration was done at select wells. However, no further mention is made of the simulation of chloride concentrations or what the results were.

DOH has completed a conceptual assessment to demonstrate that it is difficult to reconcile the measured groundwater chloride concentrations with the simulated groundwater flow trajectories as indicated by the particle tracks. According to Visher and Mink (1964) the sources of chloride to south Oahu's groundwater are rainfall, deposition of sea spray, and chloride from the saline water beneath the basal lens. A simple box model, shown in Figure 3(a) and 3(b), will suffice for this demonstration with representative chloride concentrations applied to the box model boundaries based on literature or measured values. The groundwater chloride concentration in the recharge zones upslope is very low at about 16 mg/L (Visher and Mink, 1964). The particle tracks displayed in GWFM Report and the water budget (Table 3-4) show that 40 percent of the groundwater comes from the northeast boundary, which should have a chloride concentration of about 16 mg/L. The freshwater lens is very thick at that northeast boundary and, hence, no mixing with deeper brackish water would be expected. Water recharging along the groundwater flow path to the Facility will have higher chloride concentrations than that in the upslope recharge zones: the chloride concentration in rainfall increases closer to the coast, as does evapotranspiration, and therefore, the chloride concentration in the infiltrating water will be greater than that of local rainfall. Visher and Mink (1964) state that the chloride concentration in Honolulu coastal rainfall varied from 3.0 to 29 mg/L for an average value of 16 mg/L. Average chloride concentration in the upland areas in the Kipapa drainage basin was 6.5 mg/L making 11 mg/L a representative rainfall concentration. Assuming that chloride is a conservative species, the chloride concentration of the recharging water can be approximated by:

$$[Cl]recharge = 11\ mg/L * Recharge/(Rainfall-runoff)$$

The spatial distribution of rainfall, runoff, and recharge can be taken from the USGS recharge coverage for Oahu (Engott et al., 2017). This simple approximation returns a chloride concentration for recharge directly into the subsurface of the Facility of about 45 mg/L. However, if the GWFM is correct, the groundwater chloride concentration beneath the facility would be dominated by upslope recharge and would be much less than 45 mg/L. Figure 3a uses identical color schemes to show the chloride concentrations in the Moanalua and Halawa area wells, within the Facility monitoring wells, and false color shading showing a hypothetical chloride distribution within the Facility based on the GWFM "mauka to makai" particle tracks. The hypothetical chloride distribution assumes the groundwater flowing into the northeast boundary of the model has a chloride concentration of 30 mg/L and reflects the increase in chloride concentration due to increased evapotranspiration as elevation decreases. Within the box model the chloride concentrations continue to rise due to increasing evapotranspiration and the general trend of increasing chloride concentrations in the Moanalua/Halawa Region wells going down slope (refer to Figure 3(a)). The regional chloride values are from the 2004 USGS National Water Quality Assurance study (Hunt, 2004) and in samples collected for the UH geothermal resources study that were provided to the Navy in November of 2017 (Lautze et al., 2017). The high chloride concentration at the southwest part of the facility assumes some chloride is brought up from depth and from zones nearer the coast due to the large pumping rates of the Halawa Shaft and the Red Hill Shaft. This is indicated by the relatively high chloride concentration of 152 mg/L at the Halawa Shaft. In summary, the false color shading approximates the chloride concentrations that would be expected in groundwater beneath the Facility wells if the Model 52 particle tracks are representative of the actual groundwater flow trajectories. Differences in color shading between the background and the color shading representing measured chloride values in the monitoring wells show significant conflicts between the modeled and measured chloride values implied by the GWFM flow trajectories.

Figure 3(b) focuses on the Facility and shows the particle tracks simulated by Model 52. As with nearly all the GWFM simulations, the flow trajectory is only slightly oblique from going down the axis of the Red Hill Ridge. Most of the path lines don't pass beneath any developed area or known source of chloride prior to reaching the Facility boundaries. This suggests that chloride concentration within the Facility should be closer to that of the recharge areas than to the downslope production wells. Compared to the hypothetical chloride distribution, wells RHMW04, RHMW06, RHMW07, RHMW08, RHMW05, and OWDF-MW1 stand out as having chloride concentrations significantly greater than those implied by the GWFM. The chloride concentrations in wells RHMW06, RHMW07, and OWDF-MW1 are an order magnitude or more than what would be expected given the large simulated flux of upslope recharge water down the axis of the Red Hill Ridge. The cause of these elevated chloride anomalies is currently unknown, but clearly conflict with the simulated groundwater flow trajectories.

The key point of this analysis is to show that the modeled groundwater flow field and the measured groundwater chloride within the facility wells can't be reconciled unless a source of chloride can be identified that falls within the modeled zone of contribution to the Red Hill Shaft. If the footprint of the particle tracks shown in Figure 3(b) indicates the zone of contribution to the Red Hill Shaft, there must be a source of significantly elevated chlorides within that zone of contribution. Figure 4(b) also shows the location of elevated chloride hypothesized by the Navy. Both the Halawa Quarry and the area "north of South Halawa Valley" (CSM Page 6-31, Line 44 and 6-32, Line 1) are well outside of the simulated zone of contribution for flow down the Red Hill Ridge.

DOH recognizes that the GWFM and its particle tracks are not the same as a contaminant transport model. However, particle tracks do show the simulated groundwater flow field from areas of recharge to the point of capture by the Red Hill Shaft, essentially identifying that part of the aquifer that the model indicates contributes water to beneath the USTs and to the Red Hill Shaft. Nowhere within that flow field is there a source of chlorides that could account for the great disparity between measured and expected groundwater chloride concentrations in the monitoring wells. The GWFM is intended to provide the groundwater flow field for the follow-on contaminant fate and transport model. If the GWFM can't account for the measured chloride distribution, the contaminant fate and transport model will not be reliable for simulating the migration of fuel-related contaminants or for planning release response measures.

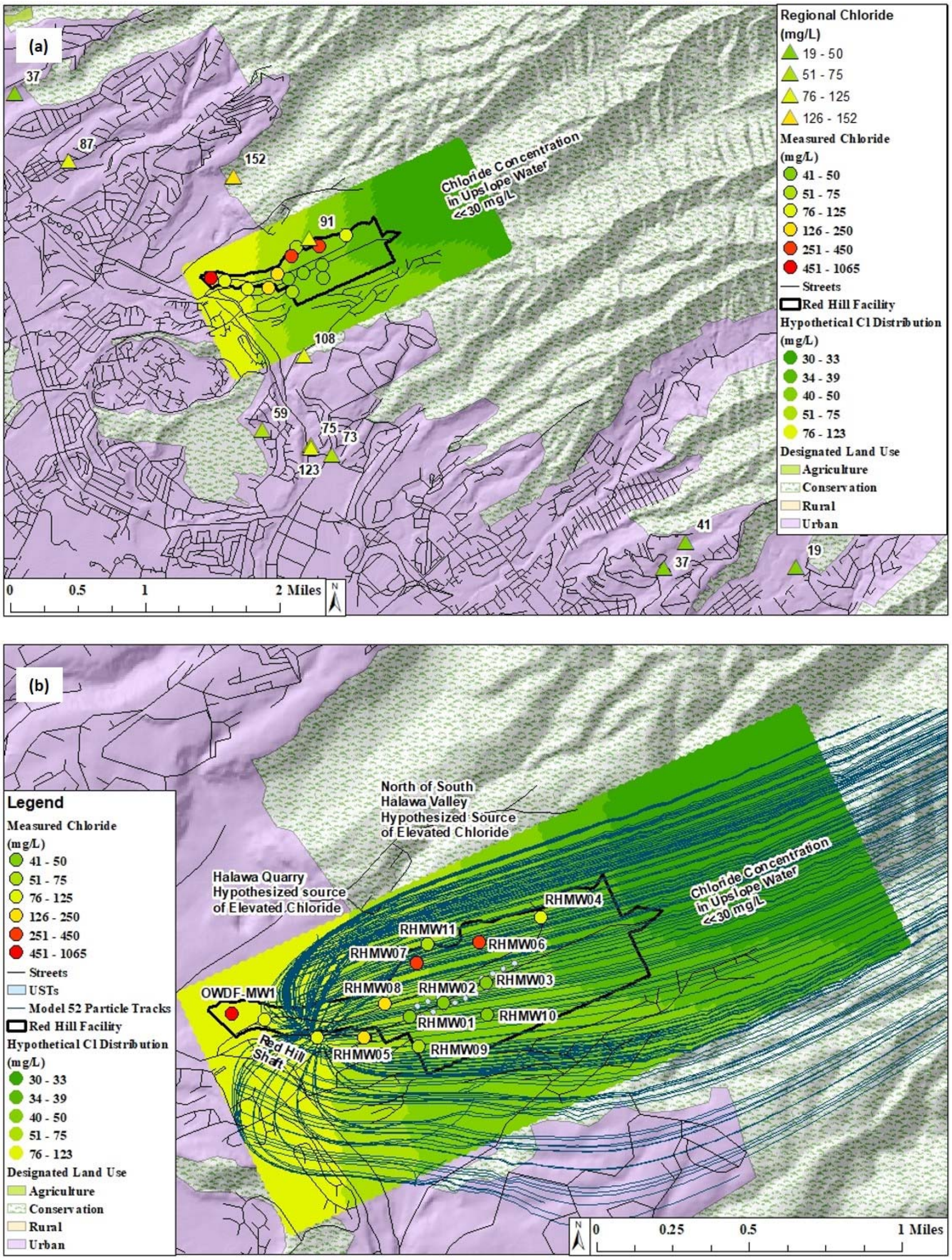


Figure 3. (a) regional chloride concentration in Moanalua/Halawa Area Wells and a hypothetical distribution of chloride in the groundwater beneath the Facility based on regional chloride concentrations. (b) Particle tracks from Model 52, and the Navy's proposed chloride source areas are added

Deficiencies of the Multi-Model Approach

The DOH agreed in concept with the use of a multi-model approach in the Navy's groundwater flow modeling efforts first discussed in July 2019. As discussed in several technical working group meetings since that time, the DOH's expectation was that this approach would systematically test plausible and defensible hydrogeologic conditions to arrive at the most likely set of conditions that represent the hydrogeologic system (i.e., the "best" model or models). The model suite presented in the Navy's GWFM report does not drive toward any definitive conclusions about the nature and behavior of the groundwater system. In the language of the AOC, it does not *"improve the understanding of the direction and rate of groundwater flow within the aquifers around the Facility."* (Site AOC, 2015). To achieve that goal, the multi-model approach needs to test and support or eliminate various assumptions in order to arrive at defensible conclusions about the nature of the hydrogeologic system that controls the direction and rate of groundwater flow around the Facility.

The conclusions section of the GWFM report (section 5.10) does not provide any significant interpretations regarding the actual in situ conditions. Rather it discusses observations of modeled outcomes and differences in that modeled behavior; those are closer to findings, not the conclusions anticipated under the AOC and groundwater modeling extensions. A number of divergent models were tested using non-traditional parameter values; all calibrated equally well; and, hence, provided little insight into which of the approaches best represent the processes occurring in the groundwater system below the tanks. Further, that the quite different models all yielded the same results using the divergent parameters suggest that the results are more reflective of the parameters selected than they are of the different models tested. A discussion of the parameters follows in the next section.

A Comparison of the Red Hill Groundwater Flow Model Hydraulic Parameters with Those Used by Other Modelers

Model parameters vary among past individual modeling efforts that simulate groundwater in the Red Hill region. Reasons for the differences in parameter values include differences in modeling codes; varying hydrogeologic assumptions; the extent of field data that are available at the time the model was developed; and the preferences of the different modelers. However, the parameters used by the various modeling efforts should fall within a reasonable range of each other and, if they don't, the reasons for the differences should be explained. Table 1 compares the USGS model of the Pearl Harbor Aquifer (Oki, 2005), the 2007 Navy model of Red Hill (Rotzoll and El-Kadi, 2007), the Board of Water Supply model of the Honolulu Aquifer (Honolulu Board of Water Supply, 2005), with the current Navy groundwater flow model for Red Hill.

This comparison of hydraulic parameters, such as the basalt horizontal hydraulic conductivity values, shows that the GWFM falls well within the range used by other modelers. However, the values the Navy chose to use for other hydraulic parameters, as highlighted in Table 1, are much different from those used in prior efforts. These parameters include vertical hydraulic conductivity, horizontal anisotropy, and porosity. The vertical hydraulic conductivity values used by the Navy are much higher than those used by Oki (2005), and Rotzoll and El-Kadi (2007). Conceptually, the effect of using a high value for vertical hydraulic conductivity is to reduce the influence that the alluvial/saprolite wedge exerts over cross-valley groundwater flow. A high vertical hydraulic conductivity will increase the ease with which the groundwater can move downward through the basalt layers and flow deeper into the aquifer, bypassing the poorly permeable alluvial/saprolite wedge. The effect of the valleys' alluvial/saprolite wedges is further reduced by modeling layers 4 through 9 as continuous, and passing beneath the alluvial/saprolite wedge, rather than a more conceptually correct approach of having the basalt layers terminate at the contact between the basalt and alluvial/saprolite wedge and resume at the contact on the opposing side of the wedge.

Table 1. Comparison of Model Parameters Used by Various Modelers to Simulate Groundwater Flow in South Oahu

Geology	Parameter	Units	Oki, 2005	Rotzoll & El-Kadi, 2007	Honolulu Board of Water Supply, 2005	Mdl 51a	Mdl 51b	Mdl 51d	Mdl 51e Zone 3
Basalt	Kh	ft/d	4500	4428	1500	2828	5316	8280	2152
	Kv	ft/d	7.5	7.4	150	200	66.3	54.9	1355
	K_L:K_T	(ft/d)/(ft/d)	3	3	1	3	10	17	
	Porosity	ft^3/ft^3	0.04	0.05	0.3	0.01	0.01	0.01	0.01
Caprock Limestone	Kh	ft/d	2500	na	50	5000	5000	5000	5000
	Kv	ft/d	25	na	50	9.45	11.87	11.9	10
	Porosity	ft^3/ft^3	0.2	na	0.4	0.073	0.095	0.095	0.07
Caprock Sediments	Kh	ft/d	0.6	115	1	20	20	20	20
	Kv	ft/d	0.6	115	1	20	0.1	0.1	20
	Porosity	ft^3/ft^3	0.1	0.1	0.4	0.03	0.022	0.022	0.03
Valley Fill	Kh	ft/d	0.058	0.066	10	1	1	1	1
	Kv	ft/d	0.058	0.066	10	0.001	0.001	0.001	0.001
	Porosity	ft^3/ft^3	0.1	0.15	0.4	0.02	0.02	0.02	0.02

Abbreviations:

Kh – horizontal hydraulic conductivity
Kv – vertical hydraulic conductivity
K_L:K_T – horizontal anisotropy ratio
ft – feet
ft^3 – cubic feet

This page is intentionally left blank

Horizontal anisotropy is another important hydraulic parameter where there is a significant difference between some of the values used in the Navy GWFM and that used by other modelers. The base horizontal anisotropy of 3:1 agrees well with values used by other modelers. However, the Navy's GWFM calibrates more closely to measured heads when horizontal anisotropies of 10:1 and 17:1 (the limit imposed during the PEST runs) were used. Other modelers found better model calibration with anisotropies of 3:1 (Oki, 2005; and Rotzoll and El-Kadi, 2007) and 1:1 (Honolulu Board of Water Supply, 2005). While it can't be determined definitively what the most appropriate horizontal anisotropy values are, the model that uses the most extreme values should include a concise physical explanation to justify that the more extreme values are in fact real and not just an artifact of the automated parameterization that produces the very high horizontal anisotropy.

The aquifer porosity selected for the Navy's GWFM is much lower than that used by any other comparable model. The USGS uses aquifer porosity in their density dependent flow models to reach agreement between the measured and modeled profile of the freshwater/saltwater transition zone (e.g. Gingerich, 2008) giving their selection of porosity a physical basis. The simulated aquifer porosity will affect the calibrated value for other hydraulic parameters such as hydraulic conductivity when simulating aquifer drawdown in response to changes in pumping stresses. A lower porosity will increase the aquifer drawdown when pumping is increased. Since the simulated drawdown in response to changes in pumping stresses was a calibration parameter, the selection of an inappropriate porosity will have a compounding effect as the values for other hydraulic parameters will need to be adjusted for the model output to match the measured drawdowns.

The Navy contends that using a low porosity is conservative from a risk evaluation perspective (Page 5-9, Lines 18 through 25). A non-conservative effect of the low porosity value is an artificial increase in the plume attenuation rate due to the erroneously high groundwater flow velocity. If the contaminant fate and transport model is calibrated with an erroneously high groundwater flow velocity, an unrealistically high contaminant attenuation rate will also be simulated. DOH assumes that GWFM results will be used in a manner like that presented in Section 3.5 of the conceptual site model report (Department of the Navy, 2019b). As described in the DOH conceptual site model review comments (DOH, 2020), the plume attenuation rates using the modeled groundwater velocities resulted in high attenuation coefficients that were inconsistent with the anoxic conditions between RHMW02 and RHMW01.

The groundwater velocities stated in the GWFM Report are much higher than any previously reported velocities. Based on travel times listed in Table 5-6, the groundwater particle velocity from the USTs to the Red Hill Shaft varied from a low 16 ft/d to a high of 110 ft/d with an average for the model runs of 51 ft/d. These velocities are much higher than generally accepted values for Oahu groundwater. For example, Lau and Mink (2006) state the average groundwater velocity in Hawaii is on the order of 1 ft/d. Whereas the geometry of the alluvial/saprolite wedges may increase the groundwater flow velocity by constraining the seepage face to between these barriers, a groundwater velocity of more than 50 ft/d is not supported by any evidence independent of the model and is inherently non-conservative.

The only direct measurement of Hawaii groundwater transport velocity on the scale of the distances present within the Facility is the Lahaina groundwater tracer study (Glenn et al., 2013). While there are distinct differences between wastewater injection in West Maui and the movement of groundwater beneath the USTs, there are also clear similarities. The wastewater injection rate is comparable to the Red Hill Shaft long term pumping rate, with both at about 4 million gallons per day. The travel distances are also similar, about 3,000 ft. The average groundwater flow velocity measured by the Lahaina groundwater tracer study was about 10 ft/d. Modeled groundwater velocities that are several times that value must be viewed with a significant amount of skepticism and require additional explanation. DOH contends that unrealistically high groundwater velocities are non-conservative and non-informative since erroneously high contaminant natural attenuation rates will be estimated.

Boundary conditions

Boundary conditions set hydrologic conditions at the perimeter of a numerical model because simulating an entire hydrologic system is not always reasonable. As described on pages 226 and 227 of National Academy Press (1990):

> “Groundwater and associated contaminants can also enter or leave the region across boundaries. The boundary conditions imposed on a model’s solution can have an important impact on the predicted flow and transport behavior. Parameters included in the boundary conditions (such as specified heads, concentrations, and fluxes) can sometimes be inferred from field observations. They are more often simply postulated.”

While there is physical logic for the boundary conditions used in the GWFM, the important details of how these boundary conditions influence groundwater flow within the model domain are largely postulated. The variability seen between the model run results are much less than would be expected given the significant changes to the values and distribution of hydraulic parameters and features within the model domain. This suggests that the boundary conditions are driving the model results.

The model’s upper boundary condition has the best scientific basis. This boundary condition represents recharge and should not be varied. The next best supported boundary condition is the bottom boundary that is the mid-point of the freshwater/saltwater transition as simulated by Oki (2005). All other boundary conditions are subject to a great deal of uncertainty.

Figure 4 shows southeast Oahu, the GWFM domain and boundaries, and water levels measured in each of the aquifer systems in southeast Oahu.

Northeast Boundary Discussion

The northeast boundary is the assumed contact between the flank lavas and marginal dike zone. There is a great deal of uncertainty about where this contact occurs and the relationship between the groundwater flow in the marginal dike zone and flank lavas. The Navy assumes that groundwater recharged upslope of the northeast boundary flows into the model. While this assumption is reasonable, whether this condition exists is currently unknown/unproven. Dikes within rift zones generally align with the axis of the rift zone (Walker, 1987), which is perpendicular to the assumed groundwater flow direction. This divergence in the directions of maximum hydraulic conductivity between the dike system and the flow basalts could impart a large anisotropy to the movement of groundwater from the dike system into the basal groundwater system. Further, adding all upslope recharge to the bottom layer of the model is questionable.

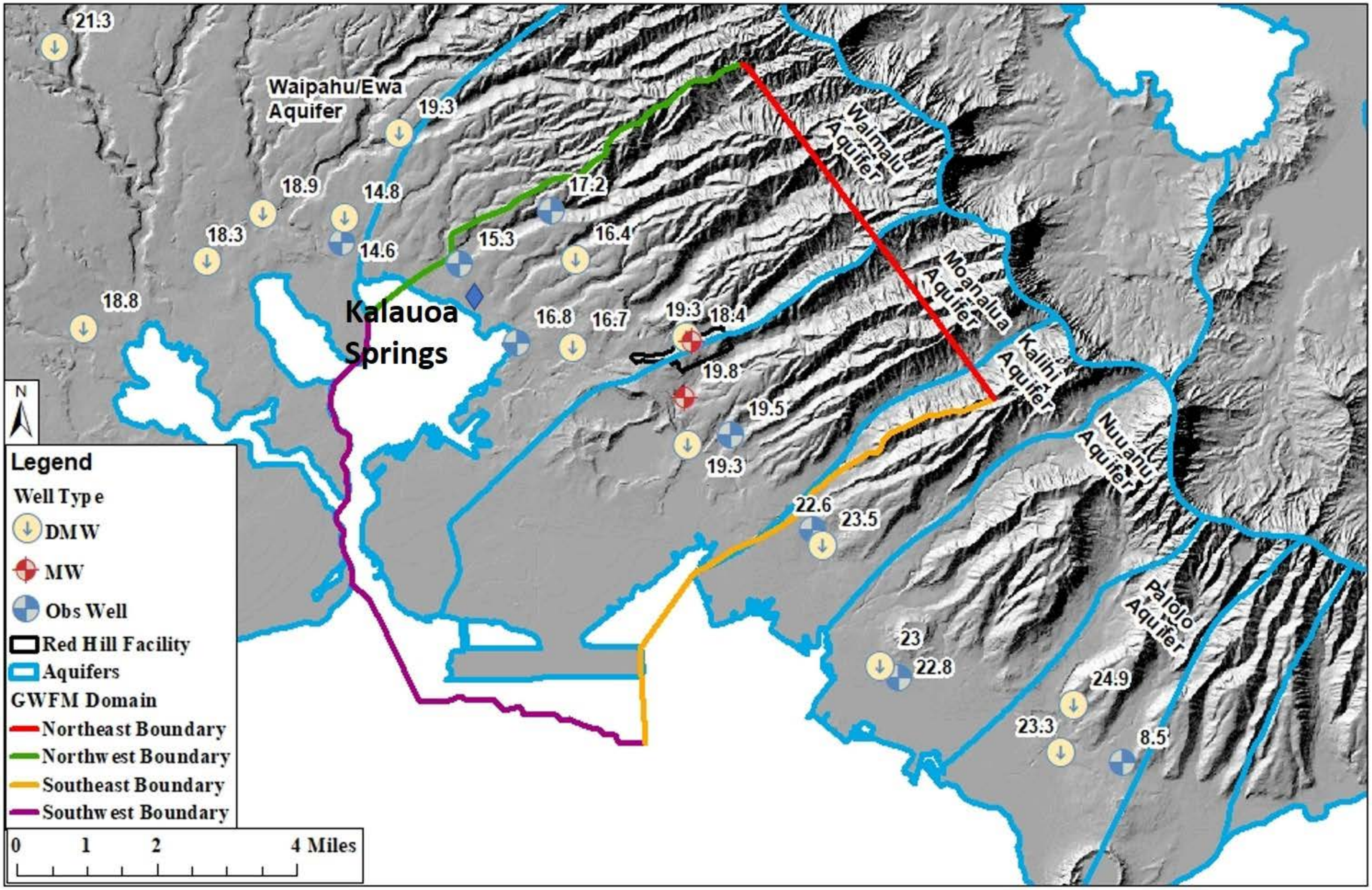


Figure 4. Southeast Oahu showing average 2017 to 2018 groundwater elevations and the model domain.
DMW – Deep monitoring well; MW – Red Hill Monitoring Well; Obs Well – HBWS water level observation well

This page is intentionally left blank

Southeast Boundary Discussion

The southeast boundary assumes no flow from Kalihi Valley into the model domain for all model runs except Model 59 that simulated 10 mgd inflow at this boundary. As with the northwest boundary, assuming little to no flow across this boundary is reasonable, but the actual flow relationship beneath and around the alluvial and saprolite wedge is not known. As Figure 4 shows, the groundwater elevation on the Kalihi (east) side of Kalihi Valley is about 2 feet higher than on the Moanalua side (Honolulu Board of Water Supply, 2018). The geometry of the alluvium and saprolite wedge in Kalihi Valley is unknown. It is also not known how far upslope from the coast the valley fill and saprolite act as an effective barrier to cross aquifer flow. Investigations in Halawa Valley indicate that the depth of the alluvial/saprolite wedge decreases rapidly as the topography transitions from the coastal plain to upslope valleys. If this is true in Kalihi Valley, differences in groundwater elevations would result in groundwater flux into the model domain. Model 59 was run to test the effect of inflow across the southeast boundary and the Navy's assessment was that model results were very similar to Model 51a. That is, any water from the Red Hill area that would be captured by the Halawa Shaft, would pass down the Red Hill Ridge on its way to Pearl Harbor. However, if other boundary conditions are simultaneously changed, the model results could be altered dramatically. As will be discussed later in this review, it is likely that testing boundary conditions individually will not adequately evaluate groundwater flow trajectories within the vicinity of the Facility.

Southwest Boundary Discussion

The southwest boundary is a general head boundary meant to represent groundwater flow from the terrestrial model domain to Pearl Harbor and the Pacific Ocean. A general head boundary assumes a hydraulic head at some distance from the physical model boundary with a bulk hydraulic conductance parameter value being assigned to represent the distance and permeability of the intervening aquifer material. As with the previous boundary conditions, the assumptions made by the Navy are reasonable but not verifiable. How accurately this important boundary represents actual hydrologic conditions is difficult to verify. The is no definitive way to determine if groundwater flow from the beneath the Red Hill Ridge flows to Pearl Harbor or to Kalauao Springs as many of the model runs indicate, or takes another pathway. The implications for groundwater flow beneath the Facility are quite different depending on which of the terminal flow possibilities is most correct.

Northwest Boundary Discussion

The northwest boundary assumes no flow across Waimalu Valley either into or out of the model. As with the other boundary conditions used, this assumption is reasonable based on the groundwater flow trajectory illustrations that dominate USGS publications such as Hunt (1996) and Oki (2005). But again, the conditions at the northwest boundary are unknown. The alluvium/saprolite wedge barrier in Waimalu Valley, located near the northwest boundary, is likely not an effective barrier resulting in groundwater flow across this boundary. For example, the large amount of discharge from the Pearl Harbor and Kalauao Springs will undoubtedly result in some groundwater flow across this boundary into the model domain.

Boundary Conditions Concluding Remarks

In summary, all the model boundary conditions are reasonable, but reasonable does not equate to accurate or even correct. Regardless, the hydrologic conditions at the model's lateral boundaries are currently unknown and therefore not verifiable. In the interim modeling phase, and to a limited extent in the final model phase, the Navy has attempted to test various boundary conditions. However, this was done in the same manner that sensitivity analysis is typically done: change is made to a single parameter (or boundary condition) and the model response is evaluated. In the case of Red Hill, it is likely that multiple boundary conditions must be changed simultaneously for the model to best represent actual conditions. This combination of boundary condition changes will likely have a compounding effect on the model output. The uncertainty about the boundary conditions is so great that it is not possible to adequately test their full range in a model, yet it is the boundary conditions that exert a large influence on modeled groundwater flow trajectory. What is needed to gain confidence in the groundwater flow model and to identify specific areas for revision are field tests that have been asked for by DOH that include, but are not limited to: a borehole flow vector survey and a tracer test of water flow velocities and vectors beneath the Facility.

Model Verification Simulations

Model verification runs are done to validate that models calibrated to the primary data sets adequately reflect groundwater conditions for a different time period. In the GWFM report, the verification simulations appear to adequately reflect the measured groundwater elevations over the period from January 10, 2018 to February 10, 2018. This is a period when there were distinct changes in pumping regimes at the Red Hill and Halawa Shafts. The DOH concurs that this is a suitable verification test period.

However, in the draft numerical models provided, the DOH SME team is unable to duplicate the results that are presented in the Navy GWFM report. This was discussed in one or more working meetings. It appears to the DOH reviewers that the GWFM drawdown results were superimposed on measured groundwater elevation data, as opposed to the modeled groundwater elevations. The DOH is unable to replicate the model verification results presented by the Navy in the GWFM report.

In our execution of the Navy's draft numerical groundwater flow models, there is substantial variance between the measured groundwater elevations versus those modeled. Groundwater elevations are much more important than drawdown/recovery because they reflect aspects of the total water budget and system behavior. Figure 5 shows the head variance between various verification model runs and the actual heads from the synoptic water level data. Except for Model 56 the simulated hydraulic head is significantly higher than the measured head hydraulic head. Transient drawdown/recovery is relatively easy to match by modeling, groundwater elevations are typically more difficult to match, but are more important. Unless the draft models were updated relative to those the agencies received, the verification simulations are unacceptable. Rather than showing strong concurrence with measured data, they are highly in error. If these observations are the result of model changes not available to the DOH, then we request those model updates so that we can again compare the verification simulation against measured groundwater elevations. If the models have not been updated in any substantial manner, then the verification runs directly demonstrate that full suite of GWFMs are demonstrably flawed.

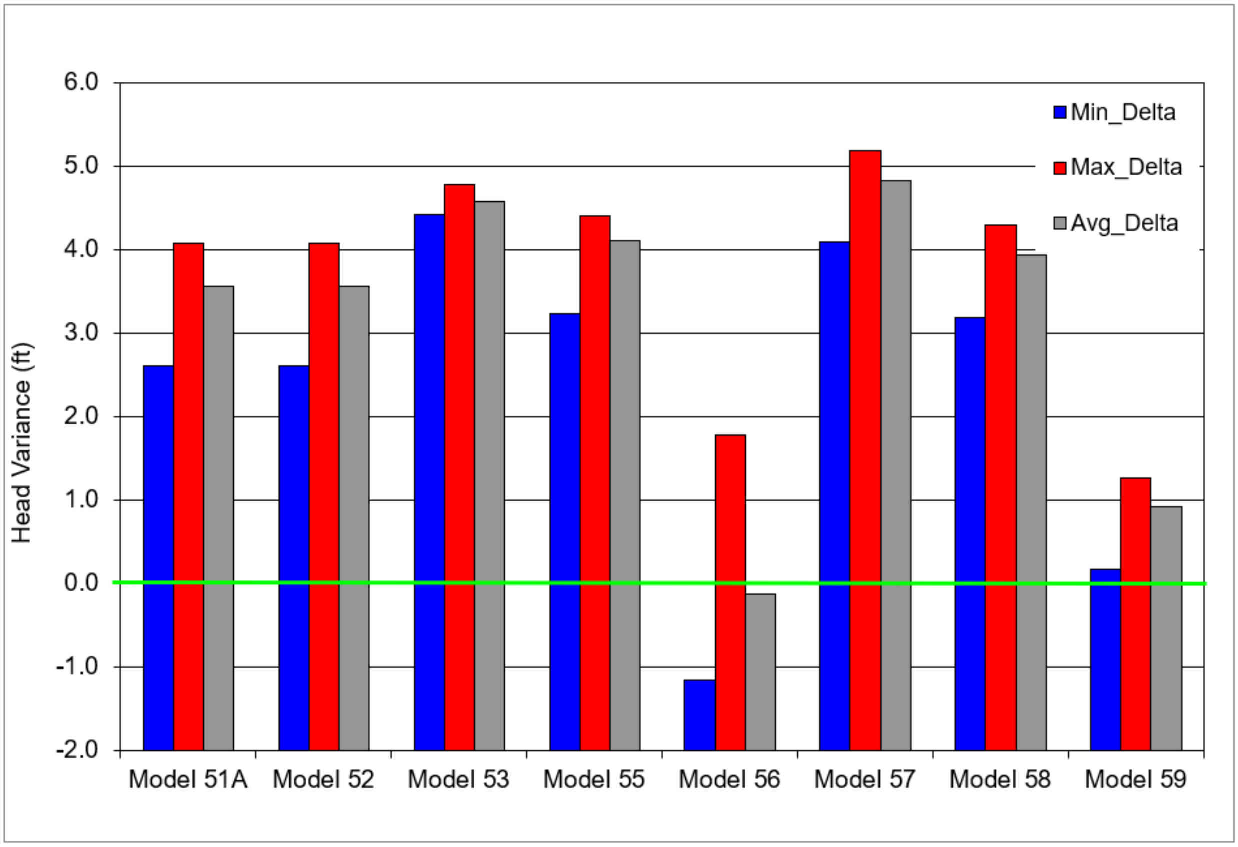


Figure 5. Modeled Groundwater Elevations Compared to Actual Synoptic Data Verification Model Variances to Measured Red Hill Area Well

Risk Conservatism

The word "conservative" is used frequently throughout the GWFM report. But rarely is there a distinct technical framing for that conservatism, or, as the degradation rate example discussed above shows, a model result may be conservative from one perspective, it can also be equally non-conservative from a different perspective. The GWFMs have elements that are conservative in some ways and non-conservative in others. But the core issue is that none of these models drives us toward a more refined understanding of the system characteristics and flow behavior. For these reasons, DOH does not find this suite of models to be adequately consistent with the area data to meet the primary AOC objective of refining our understanding of the aquifer system and its related parametric characteristics.

Further, results that the Navy labels as conservative may be an artifact of model weaknesses. An example is the degree of simulated hydraulic connectivity between the Facility and the Halawa Shaft (Page 5-33, Lines 13 through 17) the Navy concludes that because the modeled response in the Facility monitoring wells to changes in pumping stresses at the Halawa Shaft is greater than that observed, the model conservatively overestimates the connectivity between the Facility and the Halawa Shaft. However, this apparent overestimation of connectivity is almost certainly due to the high vertical hydraulic conductivity value used for the basalt and that the basalt layers are continuous from the Red Hill Ridge to the Halawa side of Halawa Valley. This overestimation of hydraulic conductivity is non-conservative since the hydraulic factors responsible for the increased response allow groundwater to flow more easily under the saprolite wedge where LNAPL floating on the water table would be blocked by the saprolite and valley fill. What is not adequately tested is the possibility that, while using values of vertical hydraulic conductivity more similar to those used in prior models, groundwater from beneath the upper part of the Facility flows around the upper toe of the saprolite toward the Halawa Shaft, a potential pathway for LNAPL.

Concluding Statements

In response to this evaluation of the GWFM the Navy will correctly point out that multiple model runs have been done to test the potential boundary conditions and hydraulic parameter values. However, in nearly all of these runs a single parameter or boundary change was tested. It is highly probable that simultaneous changes to multiple aquifer parameters and boundary conditions will be needed to properly represent the groundwater flow trajectories in the vicinity of the Facility. For example, modification to the flux from the southeast boundary (Model 59), a more extensive coverage for the Honolulu Volcanics (Model 56), a more representative geometry for the alluvial/saprolite wedge; all combined with a more realistic vertical hydraulic conductivity, and basalt layers terminating at the contact with the alluvial/saprolite wedge would likely produce a much different outcome than the individual sensitivity runs. The problem is that the possible combinations are endless and, without a definitive method of testing the model, there is no way of knowing if the models can provide useful guidance in the development of release response plans or for developing a realistic risk assessment of the threat that future releases may pose to Oahu’s drinking water supplies.

The key weakness of the GWFMs is the lack of verifiability. The Moanalua/Red Hill/Halawa region is complex, likely more complex than the Navy currently realizes. This complexity leads to simplifying assumptions about critical model components such as the hydrologic conditions at the model boundaries. Boundary conditions exert great influence on the model results and the similarity between the numerous model runs suggests that it is the boundary conditions, rather than the conceptualization within the model that are driving simulated groundwater flow trajectories. Because the boundary conditions can’t be verified, yet appear to constrain the model results, it is difficult to have confidence in the model results. Further, in many cases cited in these comments, the modeled groundwater flow trajectories don’t comport with the measured data within the Facility monitoring wells. Because one of the goals of the GWFM is to have a tool that can be used for risk assessment and release response planning, the simulated groundwater flow trajectories beneath the USTs need some form of independent verification. Currently the most likely groundwater flow trajectories beneath the USTs and elsewhere in the Facility are unknown due to the disagreement between the modeled and measured relative water levels and that both the relative water levels and natural groundwater tracers present a confusing picture. Verifiability will only come from an interactive process of physical tests and model modification based on the results of the physical tests. For example, a borehole flow vector survey that has been recommended by DOH could be done in the Facility monitoring wells. The flow vector survey results could be evaluated for correlation with measured groundwater gradients, groundwater chemistry gradients, and the modeled groundwater flow trajectories. Then modifications made to the model to bring the simulated groundwater flow trajectories in alignment with what the physical evidence indicates is most probable. DOH understands that there is no guarantee that the borehole flow vector survey will produce definitive results. However, based on success in evaluating groundwater flow trajectories at the Waimanalo Gulch Landfill, the probability of success justifies the modest investment for this approach. Ultimately, a well-designed and executed tracer test is needed to provide verifiability of groundwater flow trajectories. While a tracer test is a complex undertaking, the Lahaina Tracer Test demonstrated that this approach combined with evaluation of natural groundwater tracers is very effective at removing the ambiguity of groundwater flow trajectories. Dyes like Fluorescein are measurable at very trace concentrations and have been shown to be stable in contaminated groundwater for a period of years (United States Air Force, 2001 and 2007; and Glenn et al., 2013). Due to the dye stability in the aquifer and the very low detection limit, a properly designed tracer test can be done without risk of fouling a drinking water source.

Our review of the GWFM, as detailed above, finds that the modeling approach used by the Navy has sufficiently serious deficiencies that the DOH cannot accept the simulated flow rates and trajectories without further field verification. We believe that reliance on the model results for assessment of contaminant fate and transport, or for guidance for planning response and remediation actions in the event of future releases would be so uncertain that the threat to public drinking water supplies is currently not determined. DOH arrived at this conclusion due to model weakness that include, but are not limited to:

1. A modeling approach that favors an emphasis on automated hydraulic parameter estimation and statistical/analytical methods, but with insufficient attention to hydrogeologic principles that govern groundwater flow in the subsurface;
2. Model boundary conditions that are likely constraining the simulated groundwater flow trajectories to a narrow range;
3. Simulated groundwater gradients and elevation responses to changes in pumping stress that do not agree with field measurements in the Facility monitoring wells;
4. Use of critical hydraulic parameter values that differ significantly from those used for Hawaii's geology by other experienced and respected groundwater modelers with inadequate-to-no scientifically-based rationale offered to justify these large differences;
5. The simulated groundwater flow trajectories as indicated by the particle tracks are inconsistent with the observed diverse groundwater chemistry observed in the Facility monitoring wells; and
6. The groundwater flow models form a non-conservative foundation for the follow-on contaminant flow and transport models in that unrealistically high natural attenuation rates will be estimated and the risk that releases from the Facility pose to the Navy's and the Honolulu Board of Water Supply's drinking water sources will be underestimated.

Lack of model verifiability is the overarching factor that prevents DOH from accepting the GWFM. DOH has suggested field studies that if they verify model predictions can be instrumental in DOH approving the GWFM. These or other mutually agreed upon field studies are critical for resolving the uncertainties that are needed to develop an understanding of the groundwater flow dynamics beneath the USTs and in vicinity of the Facility. This understanding is essential to guide informed decision making about the degree of risk posed by the fuel storage at the Facility, and the response actions that would be required in the event of a release.

References

Department of the Navy. 2018. Well elevation Survey Report, Red Hill Bulk Fuel Storage Facility – Joint Base Pearl Harbor-Hickam, Oahu, Hawaii. Naval Facilities Engineering Command, NAVAC Hawaii. January 5, 2018

Department of the Navy. 2019a. Red Hill Bulk Fuel Storage Facility Administrative Order on Consent Tank Upgrade Alternatives and Release Detection Decision Document. Naval Facilities Engineering Command, NAVAC Hawaii. September 2019

Department of the Navy. 2019b. Conceptual Site Model, Investigation and Remediation of Releases and Groundwater Protection and Evaluation, Red Hill Bulk Fuel Storage Facility – Joint Base Pearl Harbor-Hickam, Oahu, Hawaii. Naval Facilities Engineering Command, NAVAC Hawaii. June 30, 2019

DOH. 2019. Hawaii Department of Health Evaluation of Groundwater Flow Paths in the Moanalua, Red Hill, and Halawa Regions. Prepared by Whittier, Robert B.; Thomas, Donald M.; and Beckett, G.D. July 2019

DOH. 2020. Hawaii Department of Health Technical Review – Conceptual Site Model, Investigation and Remediation of Releases and groundwater Protection and Evaluation, Red Hill Bud Fuel Storage Facility, Joint base Pearl Harbor-Hickam, Oahu, Hawaii, dated June 30 2019 Revision 01. March 30, 2020

Engott, J.A., Johnson, A.G., Bassiouni, Maoya, Izuka, S.K., and Rotzoll, Kolja, 2017, Spatially distributed groundwater recharge for 2010 land cover estimated using a water-budget model for the Island of Oʻahu, Hawaiʻi (ver. 2.0, December 2017): U.S. Geological Survey Scientific Investigations Report 2015–5010, 49 p., https://doi.org/10.3133/sir20155010.

EPA/DOH. 2018a. Agency letter to Mr. Mark Manfredi, Red Hill Regional Program Director, Naval Facilities Hawaii: Re: Comments on Ongoing Work to Satisfy the Red Hill Bulk Fuel Storage Facility (“Facility”) Administrative Order on Consent (“AOC”) Statement of Work requirements 7.1.3 (Groundwater Flow Model Report) and 7.2.3 (Contaminant Fate and Transport Report). Dated February 23, 2018

EPA/DOH. 2018b. Agency presentation to the Red Hill Groundwater Modeling Working Group Number 13, August 16, 2018

Gingerich, S.B., 2008, Ground-water availability in the Wailuku area, Maui, Hawaiʻi: U.S. Geological Survey Scientific Investigations Report 2008–5236, 95 p.

Glenn, C.R., Whittier, R.B., Dailer, M.L., Dulaiova, H., El-Kadi, A.I., Fackrell, J., Kelly, J.L., Waters, C.A., and J. Sevadjian, 2013. Lahaina Groundwater Tracer Study – Lahaina, Maui, Hawaii, Final Report, prepared for the State of Hawaii Department of Health, the U.S. Environmental Protection Agency, and the U.S. Army Engineer Research and Development Center

Honolulu Board of Water Supply. 2005. Final Report – Development of a Groundwater Management Model – Honolulu Area of the Southern Oahu Groundwater System. Prepared by Todd Engineers and ETIC Engineering. October 2005

Honolulu Board of Water Supply. 2018. Water level and deep monitoring well conductivity, temperature, and depth profile data for the wells in the Honolulu and Pearl Harbor Aquifers provided to the Hawaii Department of Health to support the Oahu Groundwater Flow Path Investigation. Data delivered December 2018

Hunt, C.D. 1996. Geohydrology of the Island of Oahu, Hawaii – U.S. Geological Survey Professional Paper 1412-B. https://pubs.er.usgs.gov/publication/pp1412B

Hunt, C.D. 2004. Ground-Water Quality and its relation to Land Use on Oahu, Hawaii, 2000-01 – Water Resources Investigations Report 03-3405. https://pubs.usgs.gov/wri/wri034305/

Lau, L.S. and Mink, J.F. 2006. Hydrology of the Hawaiian Islands. University of Hawaii Press. Honolulu Hawaii. Page 129

Lautze, N; Thomas, D; and Whittier, R. 2017. Hawaii Play Fairway compilation of Oahu geochemical data. Provided to the Navy on November 2017.

National Academy Press. 1990. Groundwater Models – Scientific and Regulatory Applications. Committee on Ground Water Modeling Assessment. Water Science and Technology Board, Commission on Physical Sciences, Mathematics, and Resources, National Research Council.

Oki, D.S., 2005, Numerical Simulation of the Effects of Low-Permeability Valley-Fill Barriers and the Redistribution of Ground-Water Withdrawals in the Pearl Harbor Area, Oahu, Hawaii: U.S. Geological Survey Scientific Investigations Report 2005-5253, 111 p.

Rotzoll, K., and El-Kadi, A.I. 2007. Numerical Ground-Water Flow Simulation for Red Hill Fuel Storage Facilities, NAVFAC Pacific, Oahu, Hawaii. Prepared for TEC Inc. by University of Hawaii, Water Resources Research Center. August 2007

United States Air Force. 2001. Final-Remedial Investigation for Waikakalaua and Kipapa Fuel Storage Annexes at Hickam Petroleum, Oils, and Lubricants (POL) Pipeline and Facilities, Oahu, Hawaii – Volume I, Contract No. F41624-95-D-8002, Delivery Order 0004 2001. Prepared by TEC, Inc.

United States Air Force. 2007. Final – Waikakalaua Fuel Storage Annex Comprehensive Remedial Investigation Report OU-1 and OU-2 Hickam Pol Facilities, Oahu, Hawaii. Prepared by TEC, Inc. July 2007

Visher, F.N. and Mink, J.F. 1964. Ground-Water Resources in Southern Oahu, Hawaii – U.S. Geologic Survey Water-Supply Paper 1778. https://pubs.usgs.gov/wsp/1778/report.pdf

Voss, C.I. 2011a. Editor’s message: Groundwater modeling fantasies-Part 1, adrift in the details. Hydrogeology Journal. 19. 1281-1284

Voss, C.I. 2011b. Editor’s message: Groundwater modeling fantasies-part 2, down to earth. Hydrogeology Journal. 19. 1455-1458

Walker, G.P.L. 1987. Chapter 41 - The Dike Complex of Koolau Volcano, Oahu: Internal Structure of a Hawai’ian Rift Zone. In Volcanism in Hawaii, Volume 2. U.S. Geological Survey Professional Paper 1350. Pages 961-993.

WellBore Navigation, Inc. 2017. Gyroscopic Directional Survey by Minimum Curvature for Valley Well Drilling. November 2017

November 10th, 2021

Gabriela Carvalho
Red Hill Project Coordinator
U.S. EPA, Region 9
Carvalho.Gabriela@epa.gov

Re: Comments on the Report "Groundwater Flow Model Report, Red Hill Bulk Fuel Storage Facility", March 25, 2020 (Revision 00) and Accompanying Draft Model Files

Dear Ms. Carvalho,

S.S. Papadopulos & Associates, Inc. (SSP&A) has completed a detailed review of Revision 00 of the report "*Groundwater Flow Model Report, Red Hill Bulk Fuel Storage Facility*" (GWFMR) prepared by AECOM Technical Services, Inc. (AECOM) on behalf of the Naval Facilities Engineering Command (NAVFAC) Defense Logistics Agency (DLA). SSP&A also reviewed draft versions of the various groundwater flow models that are described in the GWFMR and the Conceptual Site Model (CSM). This letter summarizes the review of both the GWFMR and the draft model files, presents conclusions and provides recommendations.

At times during the completion of this review, SSP&A participated in phone calls and netmeetings together with representatives of, and subject matter experts (SMEs) contracted by, the United States Environmental Protection Agency (EPA), Department of Health for the State of Hawaii (DOH) (collectively, the Agencies), and the United States Department of the Navy (Navy). During these calls, there was opportunity to ask questions and to obtain some clarifications regarding the contents of the groundwater model files, and portions of the GWFMR.

Comprehensive review of a large modeling report that describes multiple alternative models of a complex setting can result in innumerable observations and comments. This letter does not provide an exhaustive set of comments, instead provides sufficient comments in each technical area to illustrate concerns and guide appropriate actions. This letter commences with a high-level summary of the review and major conclusions; an overview of modeling objectives as outlined in the 2015 Administrative Order on Consent (AOC); and an overview of the multi-model approach. Following these overviews, more detailed comments are provided referring to contents of the GWFMR, the draft model files, or both.

Review Summary and Conclusions

The Navy has employed substantial expertise and expended substantial effort in the development of several groundwater flow models that depict, in various ways, aspects of the complex subsurface at and around Red Hill Bulk Storage Facility (RHBSF). In part because the complex hydrogeologic setting is not uniquely characterized, the Navy appropriately adopted a multi-model project approach. The Navy then endeavored to incorporate within some of the models some of the features or processes recommended for consideration and inclusion by Agency subject matter experts (SMEs) in meetings conducted during 2018 and 2019. The GWFMR documents most aspects of the multi-model development process and provides a fairly thorough description of the multiple models that were produced. Together with the extensive field characterization that took place simultaneously with model development, knowledge about subsurface conditions in the vicinity of the RHBSF has advanced considerably since the execution of the AOC.

It is important to be clear that, at this stage in the development process, no single model incorporates all features, events, and processes (FEPs) that may be important for the reliable simulation of conditions in, around, and beyond RHBSF. Although this is expected from a multi-model development process, there remain important aspects of subsurface conditions and patterns exhibited by monitoring well data in proximity to RHBSF, Red Hill Shaft (RHS), and in the direction of Halawa Shaft (HS) – referred to as the primary area of interest (AOI) – that are not accurately reproduced by any of the models in their present form. Furthermore, several lessons learned from the development of the multiple models – some of which are documented in the GWFMR, some of which have been communicated via in-person and virtual meetings – require further analysis, discussion, and integration before a smaller set of plausible or "behavioral" flow models can be developed as a reliable basis for fate-and-transport (F&T) modeling.

Of the models produced by the Navy, Model 51e may represent the single most plausible representation of the Navy's conceptual site model (CSM), for the reasons outlined below:

- The model structure represents a reasonable effort to include the main FEPs.
- The geometry of the simulated HS capture zone appears to be more in alignment with the CSM and with previous modeling such as that conducted by Rotzoll (2007).
- The model presents one of the closer reproductions of low-valued hydraulic gradients that are evident in the measured data in and around RHBSF, although the simulated gradients remain substantially higher than measured values in many cases.

Model 51e does not, however, incorporate all FEPs that are reasonably supported by SME knowledge or that are incorporated within and appear "behavioral" in other models. For example: Model 51e does not include a realistic representation of basalt heterogeneity or plausible features of the volcanic tuffs downgradient of RHBSF, which are included to some extent in other models. Where Model 51e does include parameter zones in the basalt – enabling the calibration process to estimate hydraulic conductivity values beneath saprolite – values for basalt vertical and horizontal conductivity underlying saprolite were estimated at more than order of magnitude less than surrounding basalts, which may reflect the presence of saprolite penetrating un-weathered basalt. As noted elsewhere in this letter, the approach used to develop the model layers may prevent the saprolite from impeding flow within the basalts, an important feature of the CSM. Many of the models may therefore amplify the propensity and rates of water and dissolved contaminant migration beneath saprolite toward HS, which may in turn exaggerate the risk posed by releases at RHBSF to this potential receptor via this migration pathway. The combination of empirical data and groundwater modeling also has not provided great insight into groundwater flow and dissolved constituent migration patterns local to RHBSF, such that the extent of hydraulic containment that is developed by RHS remains poorly understood. Further work is needed in this area because estimation of the extent of hydraulic containment developed by RHS is an important aspect of the assessment of risk posed to water supplies.

Consequently, further work is needed to obtain model outputs that correspondence more closely with observed conditions, particularly within the AOI. That work would be most efficiently undertaken following a period of model integration and consolidation. Thus, although the combination of field characterization, data analysis, and groundwater modeling, completed by the Navy has furthered knowledge within the AOI, the ensemble of models described in the GWFMR requires further improvement, consolidation, and review, before providing a reliable basis for F&T modeling and risk evaluation. Areas of emphasis for additional work are outlined in this letter.

Administrative Order on Consent (AOC) and Objectives of Groundwater Flow Modeling

Context for this review is provided by written agreements between the EPA, DOH, Navy and DLA (the Parties). The GWFMR was prepared in accordance with the 2015 AOC signed between the Parties. The primary objectives of the AOC and Statement of Work (SOW) therein are to "*take steps to ensure that the groundwater resource in the vicinity of the Facility is protected and to ensure that the Facility is operated and maintained in an environmentally protective manner*", including "*developing a better understanding of the hydrogeology of the area surrounding the Facility, and conducting an assessment of the risk to the groundwater resources that may be posed by the Facility*". The GWFMR is a deliverable under AOC SOW Section 7 "*Groundwater Protection and Evaluation*" which in turn supports Section 8 "*Risk/Vulnerability Assessment*" and Section 6 "*Investigation and Remediation of Releases*". The purpose of the GWFMR is described under Section 7.1 "*Groundwater Flow Model Report*" as "*to refine the existing groundwater flow model and improve the understanding of the direction and rate of groundwater flow within the aquifers around the Facility*." Groundwater flow modeling is also intended to support the development of a contaminant fate-and-transport model (CFTMR) (AOC SOW Section 7.2), and design of the groundwater monitoring network (AOC SOW Section 7.3). Specific components of the work detailed under the AOC SOW that relate to modeling are as follows:

> *(4) Navy and DLA will further develop models to better understand groundwater flow in the areas around the Facility and evaluate the fate and transport of contaminants in the subsurface around the Facility. As set forth below, based on the modeling effort, as approved by the Regulatory Agencies, Navy and DLA will develop and improve the existing groundwater monitoring network to the extent determined necessary.*
>
> *(5) Navy and DLA will develop a risk/vulnerability assessment, subject to approval by the Regulatory Agencies, in an effort to further understand the potential for and potential impacts of fuel releases from the Facility and to inform the Parties in development of subsequent BAPT decisions.*

Use of a Multi-Model Approach

Complex aquifer settings such as that encountered beneath RHBSF present many challenges to the development of groundwater models. Perhaps foremost among these at Red Hill is the role of basalt aquifer heterogeneity and compartmentalization, which presents difficulties for empirically interpreting water level and quality data laterally and vertically or developing models that correspond with those data. Field measurements alone are often insufficient to discriminate between potentially plausible alternate conceptual models (ACMs) of the subsurface, and calibration will often demonstrate unsatisfying correspondence and also not provide a single best model. In light of this, initial modeling efforts should not be anticipated to provide the "right answer" but to provide useful results that present defensible water budgets, incorporate primary FEPs, and reasonably re-produce field data such that they can be used to test hypotheses and provide a basis for F&T analysis.

In such settings, it is advisable in the early stages of model development not to attempt to produce a single model, but rather to consider a set of plausible models and then distinguish between those models that are in some sense behavioral and those that are not (Beven and Binley, 2014). This was in essence the approach used by the Navy. Although Beven and Binley (2014) distinguish models based primarily on fit (“*models that provide good fits to any observables available [behavioral models] and those that do not [non-behavioral models]*”), identifying the relative plausibility and value for decision making of different models also relies upon the knowledge and judgment of SMEs. For example, conceptual errors or simplifications in models can introduce bias that can be amplified by seeking “too good” a calibration fit (White et al., 2014). As a result, a model that provides a good fit to data but is missing one or more critical FEPs should not necessarily be considered more reliable or behavioral – and as such, weighted more heavily in subsequent applications – than a model that includes all known FEPs but provides a poorer fit. This is particularly true when the calibration objective function includes multiple components, as is the case of the Navy models. Given the current stage of model development, this review considers the representation of key FEPs in addition, and at times in preference, to calibration fit.

Primary Comments with Recommendations

Overall, through the various model incarnations, attempts have been made to incorporate the major FEPs that have been discussed by the Parties and their SMEs. This includes representation of the effect of basalt flow structure on anisotropy; the incorporation of downgradient volcanic deposits and cinder cones; and other FEPs. Although aspects of these FEPs are not known with a high degree of certainty or accuracy, reasonable efforts were nonetheless made to incorporate some of them. Exceptions to this statement are described in the comments below. The subjects of these comments are fundamental to flow model development and application, as one basis as use for F&T analysis and risk evaluation.

Representation of Subsurface Heterogeneity

There is abundant evidence for strong contrasts in the hydraulic properties of the basalt aquifer material, ranging from relatively non-conductive dense pahoehoe interiors, through to rough a’a and coarse clinkers. The basalt host rock is also intercepted by vertical and (less commonly) lateral fractures, together with lava tubes that follow the general dip and fabric of the lava flows (**Figure C-1**). The Navy groundwater models represent the subsurface using an equivalent porous media (EPM) approach. In doing so, the cumulative average effect of these heterogeneities, together with the prevailing basalt dip and strike, is represented using directional anisotropy of aquifer properties simulated in all models (with contracting values between models, and some zonation) plus, for some models, the use of the pilot point method (LaVenue and deMarsily, 2001; Doherty, 2003). The EPM assumption is very likely applicable at some scale for flow and dissolved-phase transport modeling purposes at the site; however, that scale has not been determined or demonstrated at this time. In addition, while the overall approach is fairly common practice for regional-scale flow modeling and water-resource analysis purposes, it has limitations at the scale of RHBSF for purposes of evaluating the hydraulic containment of RHS and contaminant F&T, two examples of which are provided below.

Example 1: predicting transport directions and rates. It is expected, based on the structure of the basalts, that regions of connected transmissive materials – clinker, for example – would be oriented in the direction of dip of the host lava flows. At the typical scale of clinker and a’a flows in this region, this would be expected to produce a fabric similar to that depicted in **Figure C-2**. The geometry of such a fabric can be visualized schematically and described qualitatively but cannot be represented deterministically. Work completed by the Navy consultants included a Monte-Carlo simulation of potential preferential pathways in the vicinity of RHBSF as described in the CSM Section 5.1.4:

> *"A total of 10,000 Monte Carlo simulations of random pathlines were generated …. 3,635 pathlines passed through the tank farm area. None of the pathlines through the tank farm area also passed through the Red Hill Shaft area. Therefore, the results indicated that it is unlikely that a preferential pathway exists between the tank farm area and Red Hill Shaft area in relation to historical lava flows."*

Figure D-1 from the CSM (included here as **Figure C-3**) illustrates 20 of these 10,000 Monte Carlo paths. The pattern of stochastic pathways is generally consistent with expectations based upon SME knowledge and with the concept depicted in **Figure C-2**. However, it appears from this analysis that preferential transport pathways may not pass through substantial portions of the groundwater monitoring network nor be detectable at RHS (at least, when it is not pumping). As previously noted by the EPA/DOH in comments provided on the CSM (April 22, 2019):

> *"For the Red Hill groundwater system, dissolved-phase fuel impacts are not expected to travel further than approximately 200-ft from the LNAPL source mass …. however dissolved phase impacts have been detected further than 200 feet from the tank farm, thus atypical transport conditions, such as fast-track transport features (open voids, lava tubes), may also contribute to the detections observed at Red Hill Shaft."*

The presence of unknown preferential migration pathways presents difficulties for interpreting historical groundwater sample results; for using the groundwater model without such features as the basis for F&T modeling or to support monitoring network design; and for incorporating the presence and effects of such features in the groundwater flow and F&T models.

Example 2: predicting hydraulic containment (capture zones) particularly at smaller or transient rates. The presence of an aquifer fabric like that depicted in **Figure C-2** presents difficulties in the deterministic interpretation of hydraulic containment (capture). In such a system, groundwater flow compartmentalization can be as significant laterally (such as between adjacent clinker zones) as it is vertically; and this can mean that the water recovered by pumped wells is more vertically derived than would be anticipated using a homogeneous-anisotropic assumption and approach. This possibility is supported by the easternmost extents of the boring log of the Red Hill tunnel (**Figure C-4**). Although the effects of compartmentalization may appear to homogenize at very large pumping rates, at lower pumping rates the effects of compartmentalization can be pronounced and can lead to misinterpretation of the source of water to pumped wells.

> ***Recommendation(s)***: 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 calibrate. 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.

Model Layering

The groundwater models were developed using an approach that, broadly speaking, follows the topography and more importantly follows the bounding geometry of the major formations or hydro-stratigraphic units (HSUs). When combined with the use of lateral layer "pinch-outs" and appropriate elevation adjustments, this approach leads to certain layers being in most places dedicated to specific hydro-stratigraphic units (HSUs). This approach is commonly used to provide numerical stability in simulations, particularly in settings with heterogeneous HSUs. However, the approach can have unintended consequences for the simulation of groundwater flow patterns in the presence of abrupt lateral transitions between HSUs. Two examples are provided to illustrate this.

Example 1: as detailed in the GWFMR Section 4.2:

> "*Layers 2 and 3 discretize the saprolite that lies largely underneath the valleys and portions of the caprock. These model layers are absent where saprolite is absent.*" and "*Layers 4 through 9 discretize the basalt aquifer.*"

As a result, Layer 4 represents basalts adjacent to and beneath the saprolite. This is accomplished by deforming (lowering) the top and bottom elevations of the layers representing basalt in the areas where saprolite is present. At RHBSF, however, this may allow for the simulation of flow within numerically contiguous basalt units beneath saprolites that are actually discontinuous (so that flow is inhibited) in the field. The saprolites formed by the weathering of basalts in such a manner that the saprolite cuts vertically downward into the stratified and sinuous basalt flows (**Figure C-5**, left panel). As a result, lateral movement in the field within otherwise contiguous basalt flow zones can be laterally impeded. However, in the groundwater models this impedance appears to be reduced by deforming layers beneath the saprolite rather than bisecting it (**Figure C-5**, right panel). Some impedance remains, however, rather than being controlled primarily by lateral conductivity contrasts between basalt and saprolite (and secondarily the basalt vertical conductivity), the simulated flow distribution it is controlled by a combination of (local) bulk transmissivity reduction (and secondarily the basalt vertical conductivity). Support for this includes the apparent relative insensitivity of simulated pathlines in many areas to the depth and conductivity of the saprolites. Different, more plausible, results would likely be obtained by constructing the model using more uniform layers aligned with the dip and using parameter value contrasts to represent changes in material type and HSU (**Figure C-6**).

Example 2: basalt layers "pinch-out" approaching Pearl Harbor, requiring groundwater discharge to be vertically upward through the overlying anisotropic basalts: consequently, there is essentially no horizontal discharge to Pearl Harbor from the basalts. As noted above, the method used to define the model layers may diminish the effect of horizontal anisotropy and indeed lateral transitions between HSUs. Taken together, the impediment of groundwater movement and discharge towards Pearl harbor due to pinching of layers, combined with the method used to define model layers, transition between HSUs, and describe the vertical extent of saprolites (noted above), the model structure may increase the propensity for simulated groundwater to flow from RHBSF beneath the saprolites to the west-northwest (i.e., toward HS and shoreline springs located in that direction) (**Figure C-7**, **Figure C-8**). Partly as a result, there is relatively little difference in particle paths between most models parameterized with different anisotropies. This restriction of discharge toward Pearl Harbor may also contribute to the necessity (documented in the GWFMR and the subject of additional comments below) that changes in simulated groundwater pumping appear to require near-equivalent volumetric changes in recharge input to the groundwater models, which should not be necessary in a properly designed model.

> ***Recommendation(s)***: consider and evaluate alternate methods to represent lateral and vertical transitions between hydro-stratigraphic units (HSUs), and the heterogeneity within layers and within HSUs. This includes re-evaluating the method used to develop the top and bottom elevations of the model layers and their correspondence with the HSUs; and, the methods used to parameterize the HSUs within and between model layers.

Recharge

The representation and relative spatial distribution of dominantly precipitation-derived recharge at the water table is based on analyses completed previously by the United States Geological Survey (USGS). As presented in the GWFMR Section 3.6:

> "*The recharge distribution is similar for these different weather conditions, and therefore it is appropriate to uniformly scale the recharge values up or down depending on the weather. The highest recharge occurs in upland areas with lowest recharge toward the coast. This is the case within the model domain as well as to the NW and SE of the model domain. These recharge maps were developed considering several factors including land use, rainfall, irrigation, and evapotranspiration and are the most detailed representations available for areally distributed recharge across the site. Their accuracy at a local level could be questioned, but the trend is appropriate in that most recharge occurs in higher elevations, with less toward the coast*."

This approach is commonly used for parsimony in the estimation of infiltration patterns and rates at the base of the soil horizon, sometimes referred to as "deep percolation". However, the total net contribution of precipitation-derived recharge to the groundwater system remains uncertain. In addition, issues can arise if (a) the model does not incorporate appropriate routes and mechanisms to discharge the deep percolation; and (b) if complex surface conditions or strong contrasts in subsurface conditions are not accounted for when presuming that deep percolation becomes net groundwater recharge. For example, in hydrogeologic settings where there are substantial slopes, the soil infiltration capacity is limited, or where there are strong conductivity contrasts between the soil horizon and underlying aquifer, infiltration that is presumed to reach the water table and form recharge can be locally rejected, resulting in actual patterns and rates of aquifer recharge that differ substantially from apparent surface infiltration patterns and rates. Three examples are provided to illustrate these concerns and possible consequences of the recharge simulation approach currently used in the Navy models.

Example 1: As noted above, it appears it was necessary to adjust recharge during calibration to scale approximately with groundwater pumping rates that are simulated in different stress periods. This is a concern because it suggests that the sources and sinks of water to the model, as provided via the boundaries, recharge, and in transient simulations storage change, are not properly meeting pumping demands. This also appears to contradict previous correspondence which indicated that recharge would be fixed between stress periods. For example, as presented in Navy comments on the letter prepared October 19th, 2019 by SSP&A as presented in the Navy response to the EPA/DOH extension letter of January 28th, 2020:

> *"3a) The Navy team determined during preliminary calibration efforts that when recharge was used as a calibration parameter, the Parameter Estimation (PEST) software frequently assigned recharge rate multiplier values that did not match the conceptual model. The conceptual model anticipated mild differences in calibrated recharge rates to accommodate the two sets of synoptic head targets. In practice, PEST tended to make large changes to recharge rates (and therefore to the regional water budget) of 20 percent or more, which suggested that the field data did not constrain the recharge rates sufficiently to permit their use as calibration parameters. The team concluded that achieving a mild improvement in calibration to synoptic heads did not warrant the unrealistic changes to the regional water budget. The final set of models (models 51 through 59) did not use recharge rate as a calibration parameter."*

This is also evidenced in flow simulations used for particle tracking in which RHS pumping rates were increased by 2.51 million gallons per day (MGPD) from the calibrated model which leads to simulated heads across RHBSF falling to or below historical minimum levels. For simulated heads to be more comparable to historical levels, recharge would likely have to be increased by about 2.5 MGPD. A water budget analysis of the models suggests that the simulated aquifer system is heavily pumped, with about 70% of the water that enters the system via recharge and boundary inflows extracted by wells and shafts. Halawa Shaft pumping alone accounts for the withdrawal of about one quarter of system inflows (at a pumping rate of 12 MGPD) compared to total inflow of about 46.7 to 52 MGPD (excluding the 10 MGPD entering via the SE boundary in some model variants).

Example 2: Recharge patterns and relative rates are based primarily on land use and cover, rather than the underlying bedrock geologic texture. It is not clear whether and how account is made for surface run-off or the possible presence of perched or low-receiving-capacity materials that may reduce the net recharge received by the basalts. It might perhaps be expected that net recharge entering the basalts at RHBSF may be greater than net recharge entering the basalts in areas beneath the valley fill and saprolites because of the very different thicknesses and character of the intervening materials. Model 57 explored recharge uncertainty under reduced (dry though not quite drought recharge conditions); however, the inference from this simulation was undermined by the re-estimation of recharge rates, and there is no corresponding evaluation of RHS containment during wet periods.

Example 3: Apparent differences in specific capacity at RHS year-to-year suggest that the zone of containment by RHS may change substantially under wet and dry conditions (it appears the RHS specific capacity was higher in 2015 than observed in 2006 and 2017). This observation suggests that the sources of water to RHS change substantially under different conditions, and consequently that during wet periods, contamination may migrate off-site due to the combination of preferential pathways (see above) and reduced extent of containment by RHS.

> ***Recommendation(s)***: The necessity of adjusting recharge by stress period may highlight underlying issues with the representation of water sources in the model, or with the use of steady-state simulations for certain analyses, or both. Either a more comprehensive rainfall-runoff-recharge calculation approach may be needed, or some ability may be needed in the calibration to adjust recharge based upon the bedrock (receiving aquifer) geologic material type. An additional scenario should also be considered to assess containment during "wet" periods during which there is greater inflow from aerial recharge, from the dyke region, and potentially from the Moanalua valley.

Lateral Boundary Conditions

Each of the models exhibits different boundary inflows than the previously published regional groundwater model (Rotzoll, 2007: "Rotzoll's model"). While strict adherence to previous estimates is not expected or required, and differences would be anticipated from a multi-model analysis, further explanation is needed to support values that were obtained and used in the Navy models. Two examples are provided to illustrate this.

Example 1: In Rotzoll's Model, the flow into the model domain along the location of the southeastern boundary was estimated to be about 2.0 MGPD. The Navy models exhibit flows that were either essentially zero or ranged up to 9.0 to 10.7 MGPD. The higher range of values – roughly 10 MGPD – appears unlikely given that the total recharge estimated for the adjacent aquifer is only about 12 MGPD. In addition, it appears from Rotzoll's Model that the influx estimated from the upslope dyke area is about 50% greater (roughly 30 MGPD) than is assumed in the Navy models, and that the inflow is conceptually interpreted as occurring higher up in the geologic sequence than in the Navy models. Although the Navy and Regulator SMEs have discussed the possibility of substantial groundwater flow from the Moanalua Valley direction through the AOI, further justification is needed to lend support to this possibility.

Example 2: Previous modeling together with groundwater elevation mapping and water budget analyses suggest there may be a non-trivial inflow of water from the northwest area from the Schofield Plateau (i.e., into the northwest boundary of the Navy model domain). Any such inflow would ultimately contribute to discharges that occur at springs along the shores of Pearl Harbor, which in many or all of the Navy models appear to be fed almost entirely by a combination of flow originating at the upslope dyke area, together with recharge accrued between this upslope area and the springs, and – in models with a high southeastern inflow – via the direction of Moanalua Valley. If a significant proportion of the discharge occurring at harbor area springs arises from the northwest boundary, this may substantially alter flow patterns within the AOI.

> ***Recommendation(s)***: An attempt should be made to obtain volumetric budgets useful for developing boundary inflow estimates from Rotzoll's Model, other suitable regional-scale models, flow-nets, or via SME concurrence. This effort should include discussions with Agency SMEs regarding the potential for inflows at the southeast and northwest boundaries.

Calibration Data Concerns and Inconsistencies between Observed Data and Simulated Results

In several places in the GWFMR and other deliverables and Navy presentations, concern has been expressed about the quality of the data available for flow model calibration. While the density in space and time of groundwater monitoring data at RHBSF is – in relative terms and given the large scale of the site – less than at many other underground storage facilities, concerns about data *quality* expressed in the GWFMR and Navy presentations appear exaggerated. Two examples are provided to highlight this.

Example 1: The GMFMR (Section 3.1) states:

> *"Also, the apparent gradients at the shallow Facility basalt wells are not consistent (can be uphill or downhill) when Red Hill Shaft is pumping. When Red Hill Shaft is not pumping, the apparent gradients in shallow Facility basalt wells all point uphill toward RHMW04 on Figure 3.1-5a. On Figure 3.1-5b, these apparent gradients all point away from RHMW01 in all directions as though*

> *that was an area of high recharge. Therefore, the Facility well water level differences should not be overinterpreted, due to the very small difference values that are within the error limits of water level measurements at any one well."*

The excerpt as written conflates two related but separate issues: the presence of flat and difficult-to-discern gradients (due to small-valued measured water level differences) versus interpretive value. The presence of consistently small water level differences, leading to difficult-to-discern hydraulic gradients and flow directions, is information. Measured head differences may be small, but they are not necessarily erroneous. Groundwater flow models that exhibit large head differences in areas where the data indicate small head differences are therefore inconsistent with the data, and contradict the evidence presented by the data that head differences and thus gradients are flat and uncertain. This concept is illustrated schematically in **Figure C-9** which depicts two probability density functions (PDFs). In this figure, the PDF on the left is narrow and has a well-defined peak; conceptually, this peak represents the model output which suggests consistent and large-valued head differences. The PDF on the right is wide and has a poorly defined peak; conceptually, this peak represents the measured water level data which suggest low-valued and at times indeterminate head differences. Even though the PDF on the right (representing the data) is wide, reflecting uncertainty and the fact that the resulting hydraulic gradients and flow directions are not known with confidence based on measured head differences, the PDFs differ distinctly and the model PDF does not correspond to the data PDF.

Example 2: The interpretation that small-valued head differences (and resulting flat gradients) have limited interpretive value leads the Navy to develop a method for weighting certain observation data in the model calibration that deemphasizes them. As a result, the calibration emphasizes matching regional head differences and gradients at the cost of reasonably reproducing head differences and gradients at RHBSF. This is not immediately evident in the figures presented in the GWFMR (for example **Figure C-10**) but can be visualized using plots that focus on RHBSF. For example, during stress period three, five of the eight facility wells compared to RHMW-01 exhibit simulated head differences that are in the opposite direction in every Navy model (**Figure C-11**): the observed head difference between RHMW-01 and RHMW-04 while the RHS is not pumping was -0.19 feet, yet the range of simulated differences was between 0.14 and 0.29 feet. This difference is not an issue of data quality; rather, as illustrated by the schematic PDFs shown in **Figure C-9**, it represents a systematic difference between the models and the field data.

> ***Recommendation(s)***: In the absence of demonstrated errors, it is important to accept the data and not conflate small differences in value with error and lack of interpretive value. Although the small differences between observed and simulated results are within the limits of measurement error (and might be given smaller weight during calibration), when differences exceed reasonable limits, they should be penalized accordingly. As a partial mitigation of this issue, the Navy should incorporate data obtained February through March 2019 – which covered a planned shutdown-and-rebound test at RHBSF – into the calibration. Doing so is likely to provide substantial value for understanding groundwater flow and hydraulic containment dynamics close to RHBSF.

Pathline Figures for Depicting and Comparing Capture Zones

Particle tracking figures presented in the GWFMR to delineate the extent of hydraulic containment (capture) developed by RHS, and depict potential sources of water to HS are challenging to compare and contrast and make it difficult to discern and compare the ultimate extent of capture at different rates.

Recommendation(s): The Capture Frequency Map (CFM) approach is recommended for depicting hydraulic containment (capture) and the sources of water to supply wells. This is illustrated in **Figure C-12**. To create this figure, forward particle tracking models developed by the Navy that simulate RHS as pumping were used together with a regular grid of particles (i.e., 99 columns by 73 rows with equal spacing of 100 ft) to delineate capture. After the initial simulations were completed using RHS pumping rates from the Navy models, the pumping rate was reduced from 4.65 MGPD to 3 MGPD and then to 1 MGPD to assess the extent of capture corresponding to the reduced pumping. A CFM was created for each of three pumping rates by summing the particles captured at RHS at each particle starting location and dividing by the total number of models (i.e., 12). The image more completely depicts and contrasts the extent of capture under these varying conditions and rates, using the Navy models. (*Note: these figures are provided only to illustrate the application of the CFM approach using the Navy models. The extents of capture depicted in these figures do not represent an opinion of the Agencies or their SMEs as to the probable extents of capture of RHS.*)

Relative Weighting of Alternate Models

The highest-weighted models presented in the GWFMR Table 5-7 *Summary of Multimodel Applicability for Risk-Based Decision Making* are those that simulate heterogeneous basalts using the pilot point method. In concept, the representation of heterogeneous basalts using pilot points is more consistent with the CSM than a presumption of homogeneity; however, the relative-weighting or ranking rests too heavily on the improved calibration fit obtained using pilot points. Such an improvement in fit statistics would be expected by the introduction of a much larger number of parameters to the calibration process, but should not necessarily be interpreted as improved reliability, because of the greatly increased parameterization of those models.

Recommendation(s): if using calibration statistics for multi-model ranking, more comprehensive methods should be considered that incorporate the concept of calibration or residual standard error, degrees of freedom, and parameterization, such as those described by Poeter and Anderson (2005), among others. Inclusion and representation of FEPs should, however, also be included in the weighting, ranking, and consideration of the plausibility of models.

Additional Comments

The following is a brief list of additional noteworthy comments:

- Model bottom elevation: the base elevation of freshwater in the model is based upon a simulated saltwater lens that appears to suggest that Pearl Harbor is freshwater rather than saltwater.
- Transfer Function Noise (TFN) analysis: the use of the TFN technique with analytical expressions to derive clean response functions at monitoring wells to include in the calibration of complex numerical models is appropriate. However, two concerns arise with the results and presentation:
  - The use of analytic expressions that assume anisotropic homogeneous conditions to derive response functions should not be conflated with the interpretation that the subsurface can be reasonably represented as an anisotropic homogeneous one for purposes of flow and F&T modeling.
  - The TFN provides a "clean" drawdown-recovery response, which is important to the estimation of bulk aquifer properties. However, estimation of the extent of hydraulic

containment (capture) developed by pumped wells, shafts, or tunnels also requires that the groundwater model reasonable reproduce measured groundwater elevations and head differences (gradients).

- Verification modeling: the verification results provide little inference regarding the absolute o relative performance of the various models. Reasons for this include: (1) Simplification of pumping cycles in the verification simulations obscures differences and similarities between simulated and observed responses. Simulated stress periods do not represent the pumping / non-pumping cycles (short-duration cycles, and one longer-duration cycle that results in extended recovery, are not modeled). (2) Graphical depiction of the simulation results which appears to incorporate vertical "offsets" between the simulation outputs and measured data. As noted above, correspondence with groundwater elevations and head differences is also necessary.
- Graphics comparing model output with measured data: many of the graphs in the GWFMR are prepared at a scale that makes it difficult to evaluate model performance. For example, GWFMR Figure 3.3-1 illustrates results for all wells on a single figure using the same color for each group of wells. This method of presentation obscures potentially meaningful differences between well groups and between individual wells within each group, so that very little meaningful inference can be made from the plot regarding the performance of the models.
- Observation data used for calibration: GMFMR Section 4.5 states: *"The water level differences were initially provided unit weighting for calibration because they are indicative of gradients that govern flow magnitude and direction, which are a primary objective for the model."* It appears from the model files that regional head differences were given unit weighting, while weights at RHBSF were often an order of magnitude less than unit weighting. This may be a contributing factor to the poor fit to head differences at the facility. There may also be inconsistencies in depicted observation locations: for example, GWFMR Figure 3.1-1b illustrates well locations used for calibration, yet there is no reference to most of these wells in the calibration files.
- Parameter values estimated during calibration: in several models, hydraulic conductivity values for one or more HSUs were estimated at their upper or lower limits and in some cases estimated values appear greater than maximum values presented in Table 4-1 of the GWFMR. In addition, parameter ranges listed in GWFMR Table 4-1 do not appear to match those listed in the calibration input files, nor to match GWFMR Table 5-2. There also appear to be parameters listed in GWFMR Table 5-2 that are not listed in GWFMR Table 4-1 nor appear in the model files. These apparent discrepancies may, however, arise from version differences between the model files reviewed and those depicted in the GWFMR.

Conclusions

The main conclusions from this review were provided at the commencement of this comment letter. I hope you find the foregoing review helpful. If you have any questions or comments regarding the contents of this review letter, please do not hesitate to contact me.

With regards,

Matthew J. Tonkin, PhD, President
S.S. Papadopulos & Associates, Inc. (SSP&A)
Williamsfield, Illinois

References

Beven, K. 2019. Towards a Methodology for Testing Models as Hypotheses in the Inexact Sciences: Proceedings of the Royal Society A 475: 19.

Beven, K. and Freer, 2000. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology.

Beven, K., and A. Binley. 2014. GLUE: 20 Years On: Hydrological Processes 28: 5897-5918.

Beven, K., and J. Freer. 2001. Equifinality, Data Assimilation, and Uncertainty Estimation in Mechanistic Modelling of Complex Environmental Systems Using the GLUE Methodology: Journal of Hydrology 249: 11-29.

Department of the Navy (DON). 2019. Conceptual Site Model, Investigation and Remediation of Releases and Groundwater Protection and Evaluation, Red Hill Bulk Fuel Storage Facility, Joint Base Pearl Harbor-Hickam, O‘ahu, Hawai‘i; June 30, 2019, Revision 01. Prepared by AECOM Technical Services, Inc., Honolulu, HI. Prepared for Defense Logistics Agency Energy, Fort Belvoir, VA, under Naval Facilities Engineering Command, Hawaii, JBPHH HI.

Department of the Navy (DON). 2020a. Groundwater Flow Model Report, Red Hill Bulk Fuel Storage Facility, Joint Base Pearl Harbor-Hickam, O‘ahu, Hawai‘i; March 25, 2020, Revision 00. Prepared by AECOM Technical Services, Inc., Honolulu, HI. Prepared for Defense Logistics Agency Energy, Fort Belvoir, VA, under Naval Facilities Engineering Command, Hawaii, JBPHH HI

Department of the Navy (DON). 2020b. Investigation and Remediation of Releases Report, Red Hill Bulk Fuel Storage Facility, Joint Base Pearl Harbor-Hickam, O‘ahu, Hawai‘i; March 25, 2020, Revision 00. Prepared by AECOM Technical Services, Inc., Honolulu, HI. Prepared for Defense Logistics Agency Energy, Fort Belvoir, VA, under Naval Facilities Engineering Command, Hawaii, JBPHH 30 HI.

Grueber, C.E., S. Nakagawa, R.J. Laws, and I.G. Jamieson. 2011. Multimodel Inference in Ecology and Evolution: Challenges and Solutions: Journal of Evolutionary Biology 24: 699-711.

LaVenue, M., and de Marsily, G., 2001, Three-dimensional interference test interpretation in a fractured aquifer using the pilot-point inverse method: Water Resources Research, v. 37, no. 11, p. 2659–2675, doi:10.1029/2000WR000289.

Poeter, E., and D. Anderson. 2005. Multimodel Ranking and Inference in Ground Water Modeling: Ground Water 43, no. 4: 597-605. Doherty, J., 2003, Groundwater model calibration using pilot-points and regularization: Ground Water, v. 41, no. 2, p. 170–177, doi:10.1111/j.1745-6584.2003.tb02580.x.

White, J. T., Doherty, J., & Hughes, J. 2014. Quantifying the predictive consequences of model error with linear subspace analysis. Water Resources Research, 50(2), 1152-1173. https://doi.org/10.1002/2013WR014767

Figures

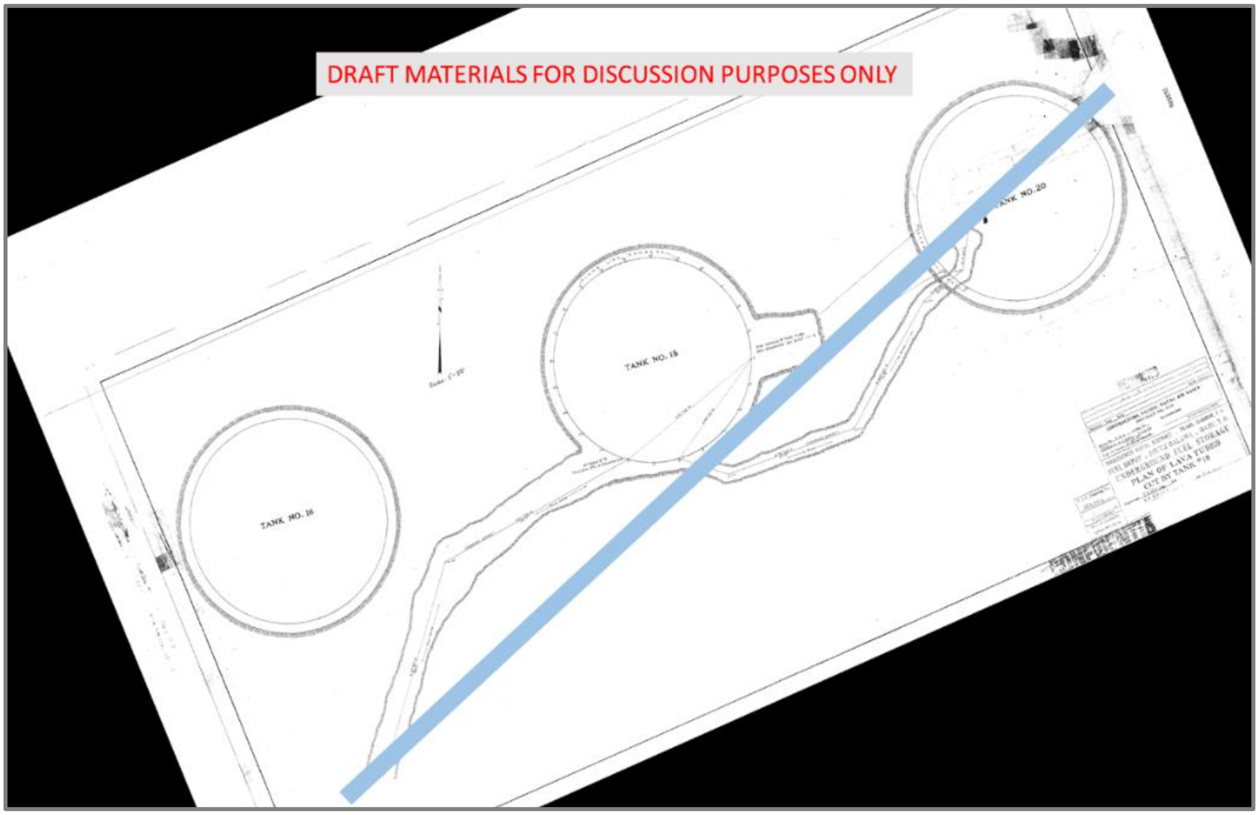


Figure C-1: Illustration of Mapped Lava Tube from CSM Report

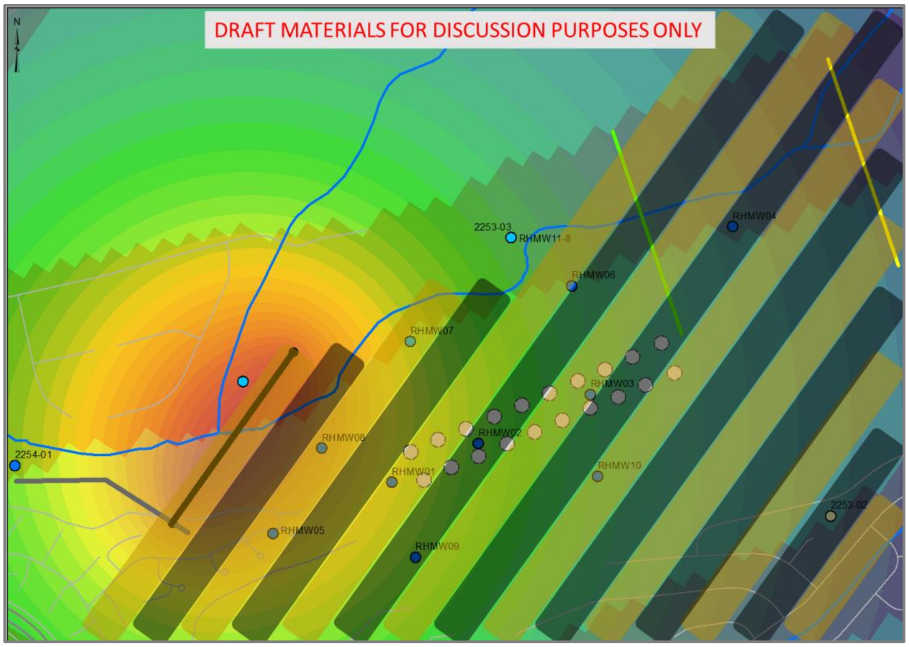


Figure C-2: Gross Schematic of Bedrock Fabric and Orientation of any Preferential Pathways

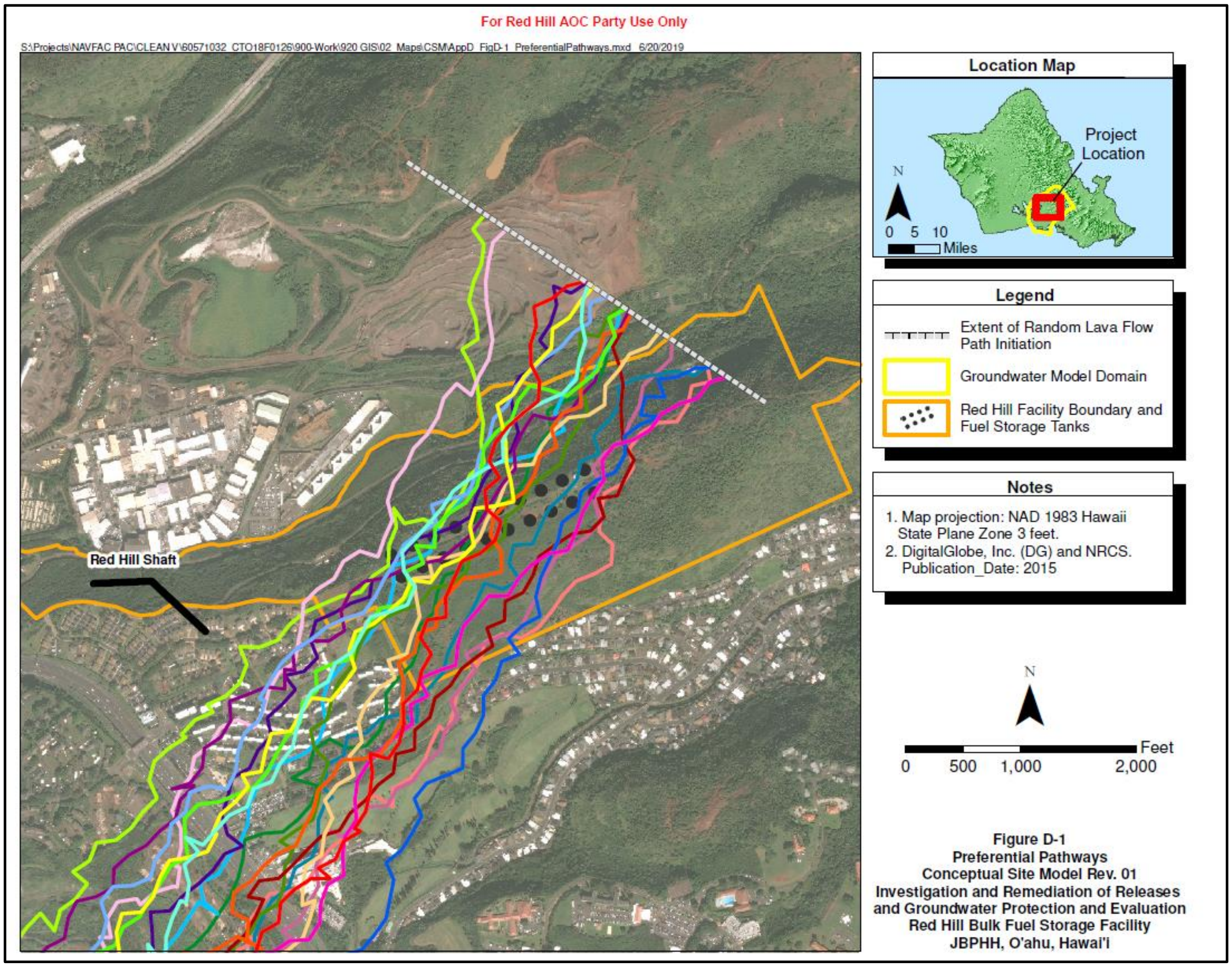


Figure C-3: Results of Structure-Imitating Stochastic Lava Simulations

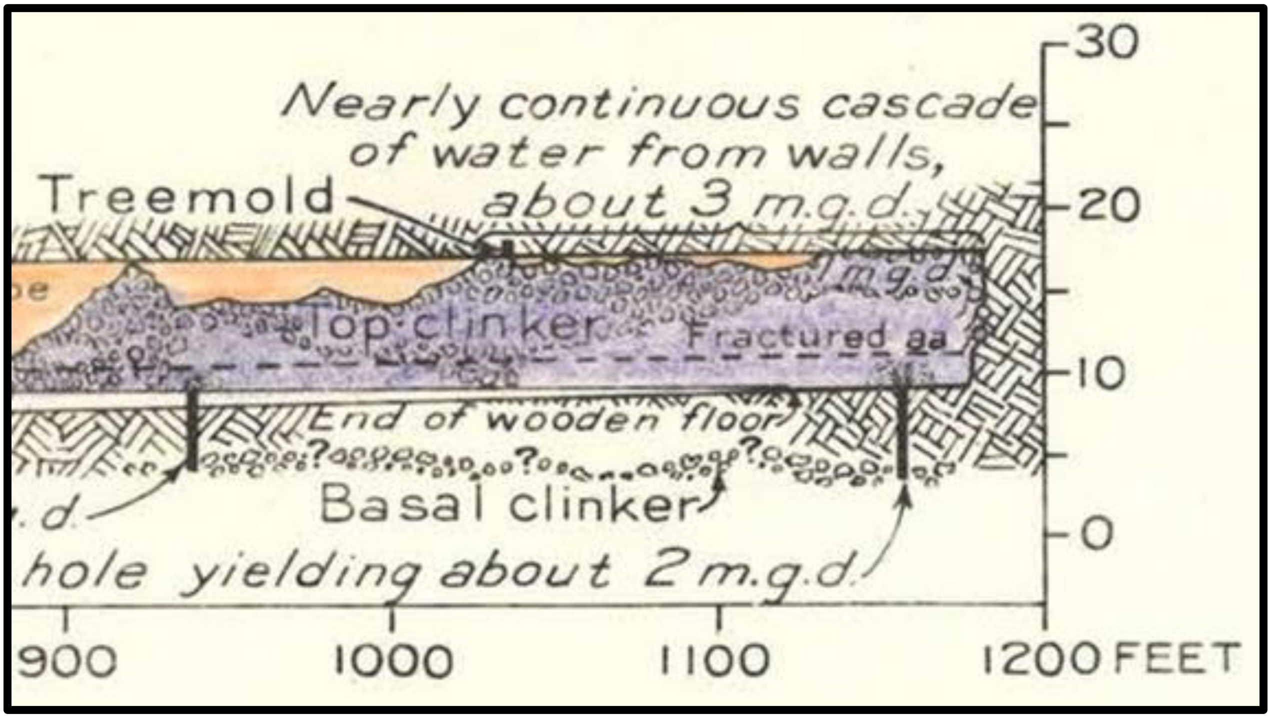


Figure C-4: Extract from Boring log for Red Hill Tunnel

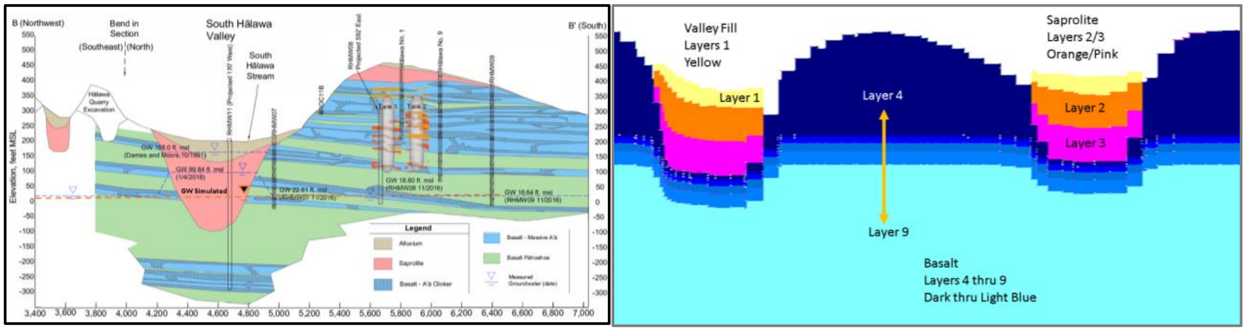


Figure C-5: Conceptual Model for Saprolite-Basalt (Left) and Simulated Representation (Right)

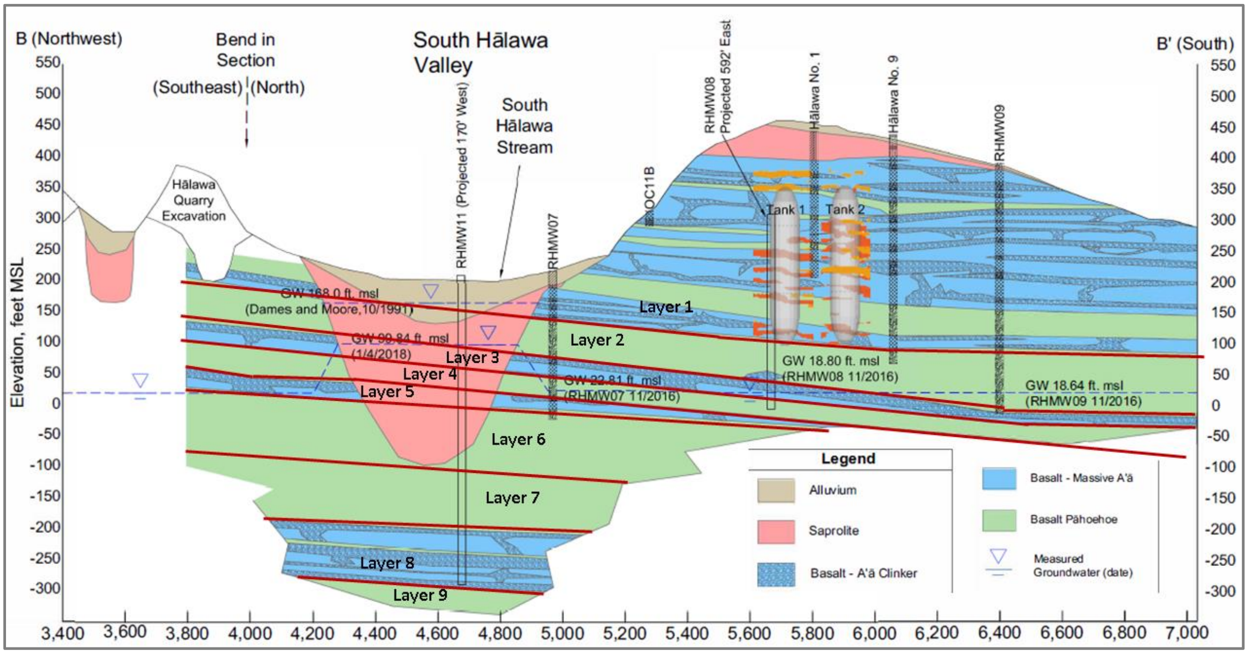


Figure C-6: Alternative Representation of Saprolite-Basalt Relationships

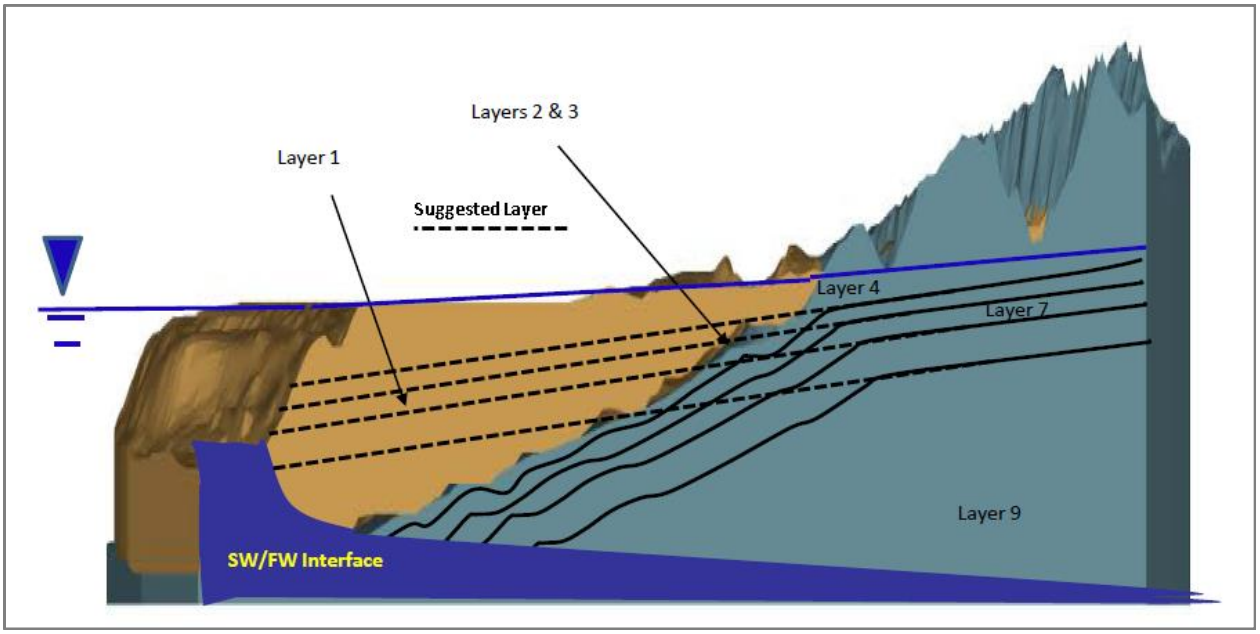


Figure C-7: Alternative Representation of Basalt-Caprock

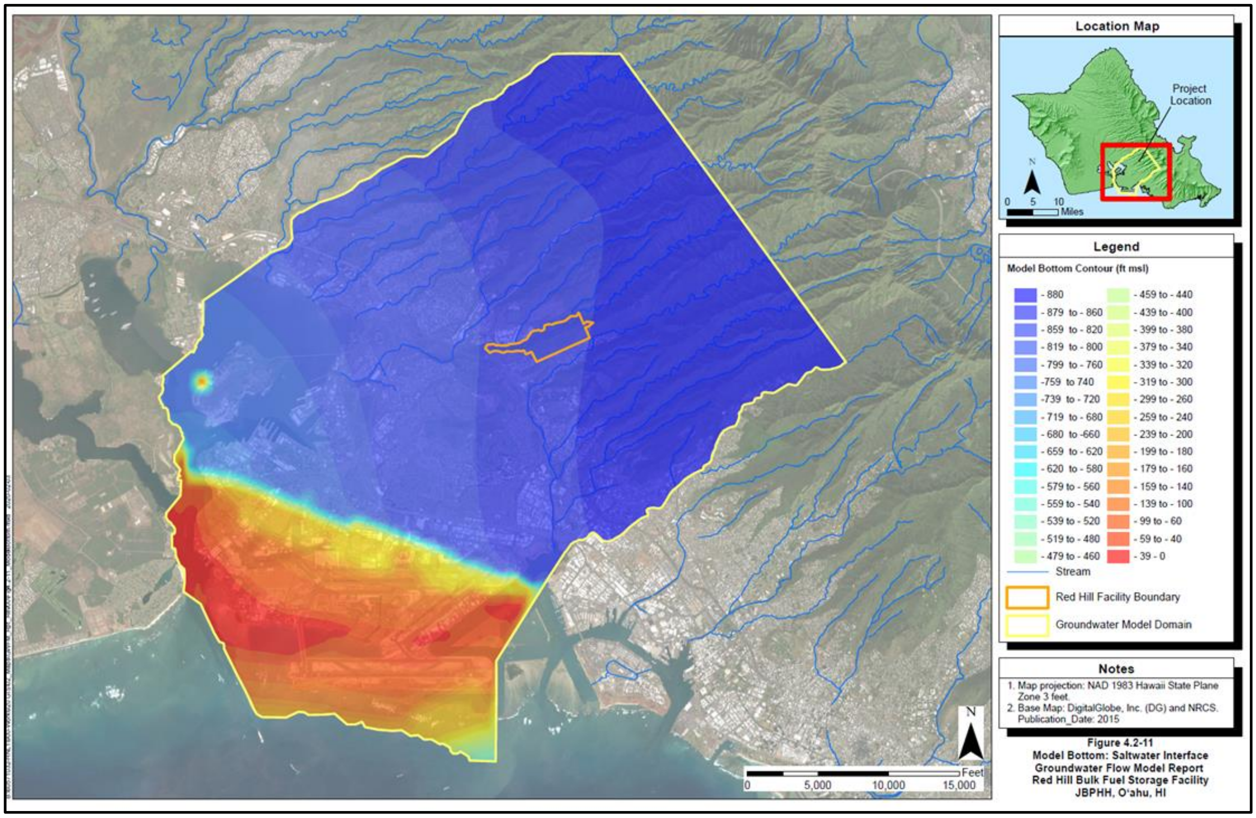


Figure C-8: Representation of Freshwater-Saltwater Interface

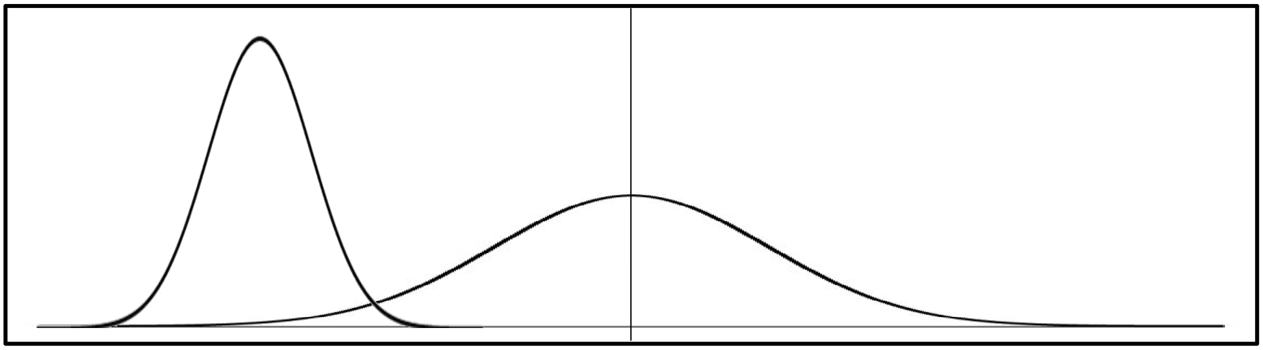


Figure C-9: Schematic Illustration Comparing Two Different Probability Density Functions (PDFs)

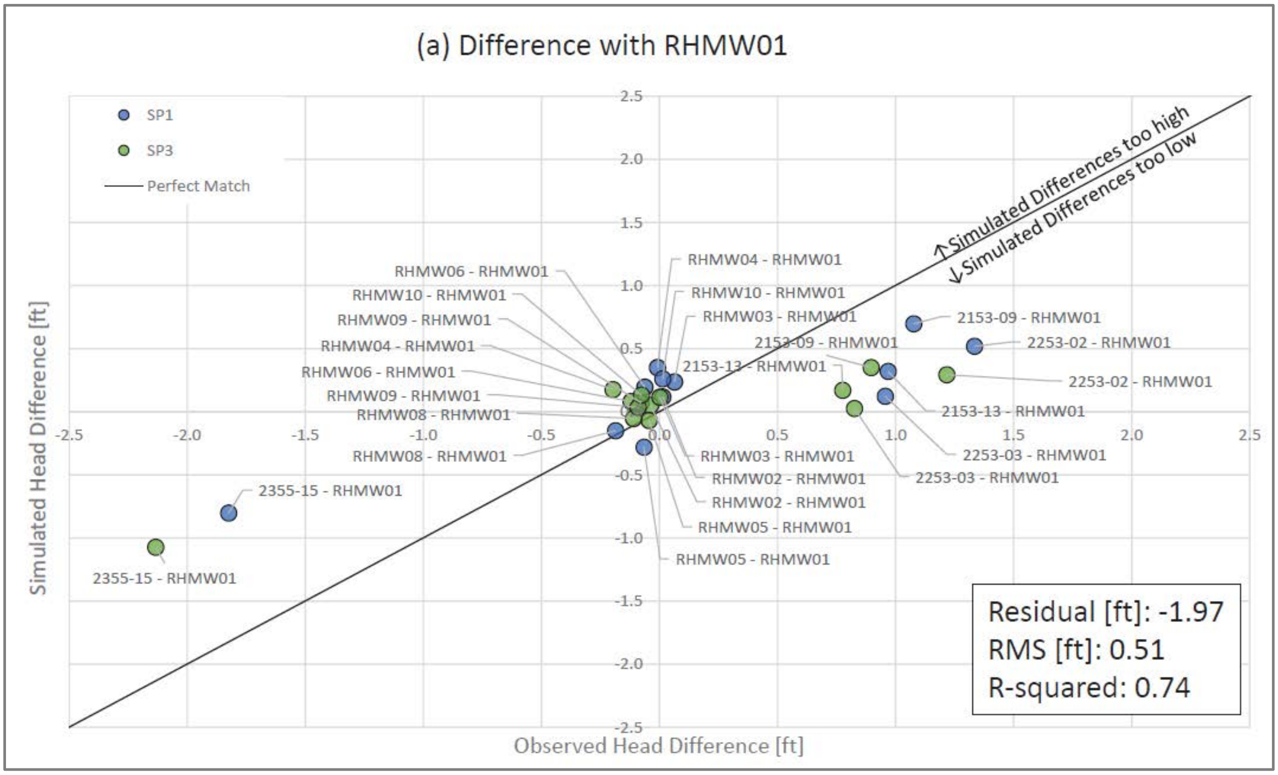


Figure C-10: Example Graphical Depiction of Simulated and Observed Head Differences

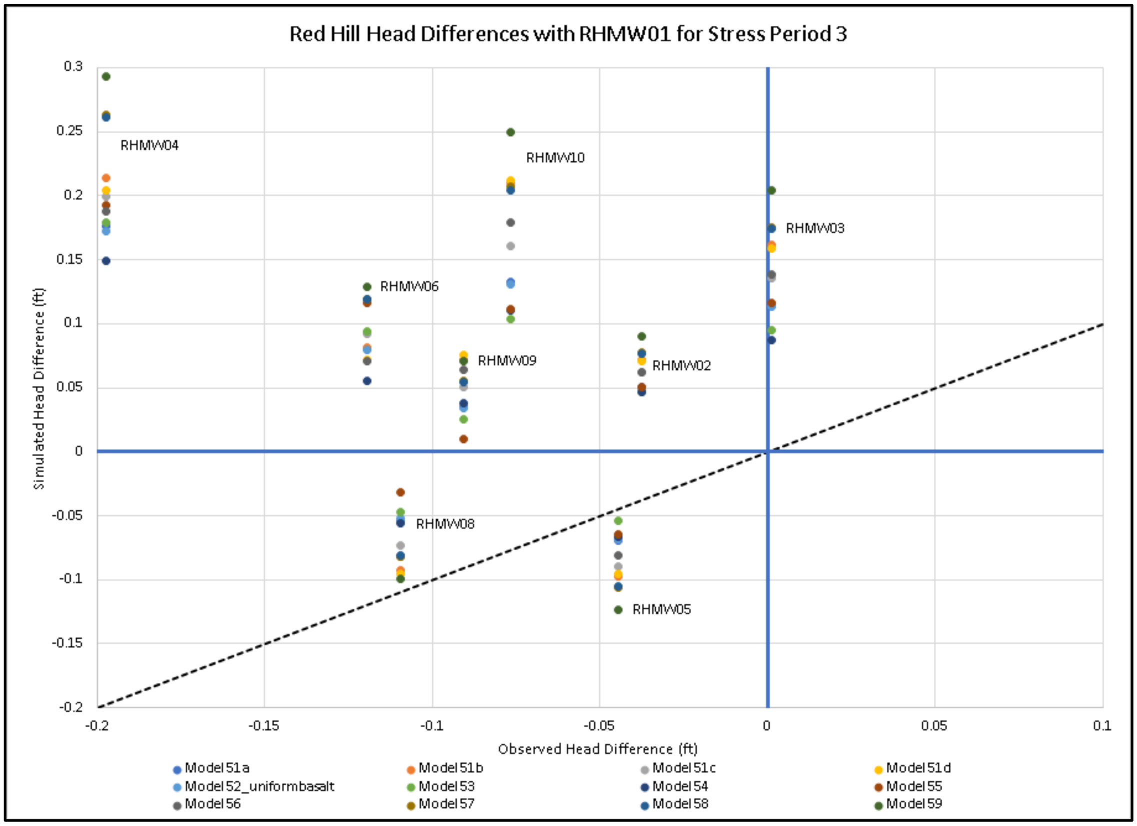


Figure C-11: Alternative Graphical Depiction of Simulated and Observed Head Differences

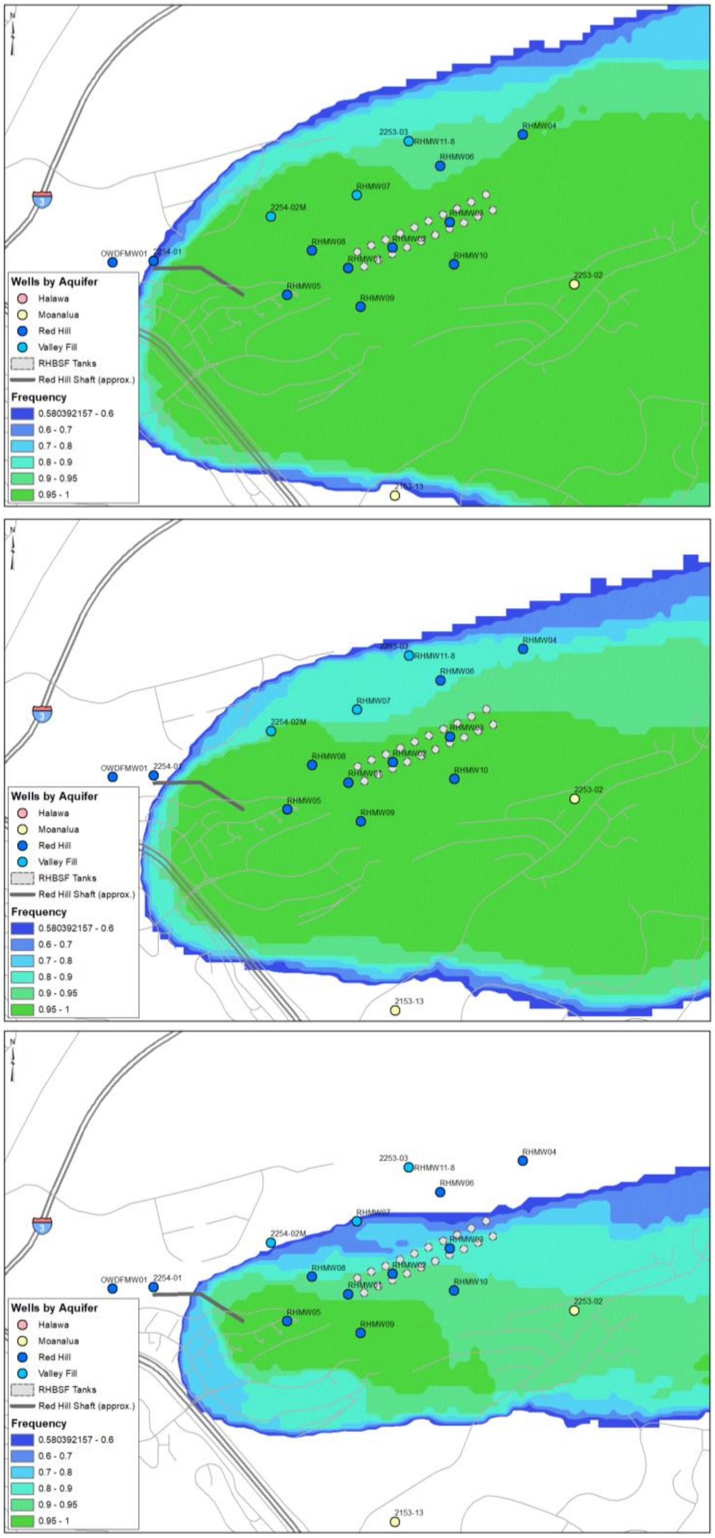


Figure C-12: Example Capture Frequency Maps at Three Alternate Pumping Rates for Red Hill Shaft

Tables

Table C-1: List of Parameter Literature Values and Estimated Values

(Bold indicates estimated value at/above max/min)

Parameter	Minimum Literature Value	Maximum Literature Value	Minimum Estimated Value	Maximum Estimated Value	Average Estimated Value	Percentage at or below Minimum	Percentage at or above Maximum
Caprock Kh (marine)	500	2500	**2500**	**5000**	**4808**	0.0%	100.0%
Caprock Kv (marine)	0.001	15	0.18	11.87	9.3	0.0%	0.0%
Caprock Kh (alluvial)	0.1	20	**0.1**	**20**	18.5	7.7%	92.3%
Caprock Kv (alluvial)	0.001	2	0.09	**20**	**13.8**	0.0%	69.2%
Valley fills, Kh	2	200	**1**	**200**	16.3	92.3%	7.7%
Valley fills, Kv	0.01	10	0.001	3.37	0.26	92.3%	0.0%
Saprolite under valley fill, Kh	0.1	10	4.81	**10**	5.7	0.0%	7.7%
Saprolite under valley fill, Kv	0.001	0.1	0.002	**0.8**	0.072	0.0%	7.7%
Saprolite under caprock, Kh	0.1	10	0.8	**10**	4.7	0.0%	7.7%
Saprolite under caprock, Kv	0.001	0.1	0.002	**0.8**	0.084	0.0%	7.7%
Tuff overlying marine, Kh	0.01	200	**200**	**500**	**477**	0.0%	100.0%
Tuff overlying marine, Kv	0.01	15	**0.01**	3.17	0.30	69.2%	0.0%
Tuff overlying alluvial, Kh	0.01	200	10	20	10.8	0.0%	0.0%
Tuff overlying alluvial, Kv	0.01	15	**0.001**	0.18	0.031	69.2%	0.0%
Tuff cone, Kh	0.01	50	**0.001**	0.089	**0.008**	92.3%	0.0%
Tuff cone, Kv	0.001	5	**0.001**	0.008	0.002	92.3%	0.0%
Basalt, Kh	500	20000	1814	8280	3657	0.0%	0.0%
Basalt, Kv	2	200	44.54	**200**	136	0.0%	23.1%

Table C-2: List of Estimated Vertical Anisotropies

Vertical Anisotropy	Maximum	Minimum	Average
Caprock (marine)	13889	421	1531
Caprock (alluvial)	222	1	49
Valley fills	1000	59	928
Saprolite under valley fill	2500	6	737
Saprolite under caprock	2500	6	596
Tuff overlying marine	50000	158	35429
Tuff overlying alluvial	10000	100	7002
Tuff cone	11	1	2
Basalt	151	9	48

DOH Review: Navy Groundwater Flow Models & Related Issues with the Navy CSM for the Red Hill Facility

By:

The Department of Health Hawaii (DOH)
Technical subject matter experts
Robert Whittier, Don Thomas, G.D. Beckett
& Anay Shende

In coordination with EPA, Region 9

October 19, 2021

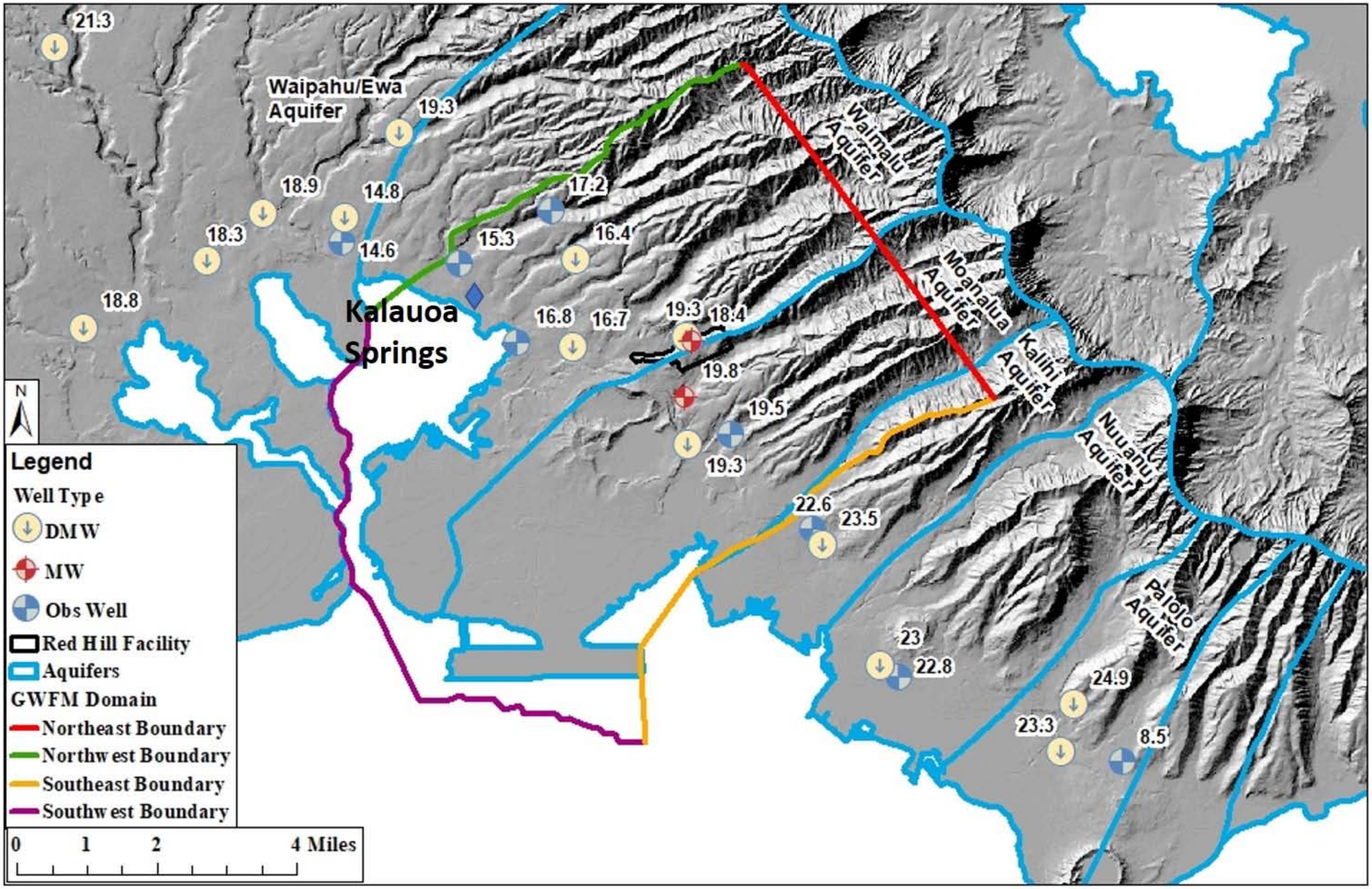
GWFM - Drinking Water Risk Concerns

1. GWFMs boundary conditions have uncertainty
   a) Chosen BCs are reasonable for primary models
   b) Data indicate other boundary conditions are probable
2. Validation doesn't ensure the model adequately replicates groundwater flow trajectories
   a) Currently used comparative data – g.w. gauging
   b) Verification simulations appear not to match elevations
   c) Alternative verification data sets
3. Model conclusions and data contrasts are problematic
   a) Critical question: Do the model results support the conclusions in the IRR Report?
   b) And future CF&T (Part II discussion)

Critical Drinking Water Risk Evaluation Questions

- Does pumping the Red Hill Shaft mobilize groundwater from beneath all tanks toward the Red Hill Shaft?
- Is there an unobstructed hydraulic pathway from beneath the tanks to the Halawa Shaft?
- Over-arching question:
  - Is the model informative for answering either or both of those questions?
  - Can the models adequately inform CF&T (Part II)?

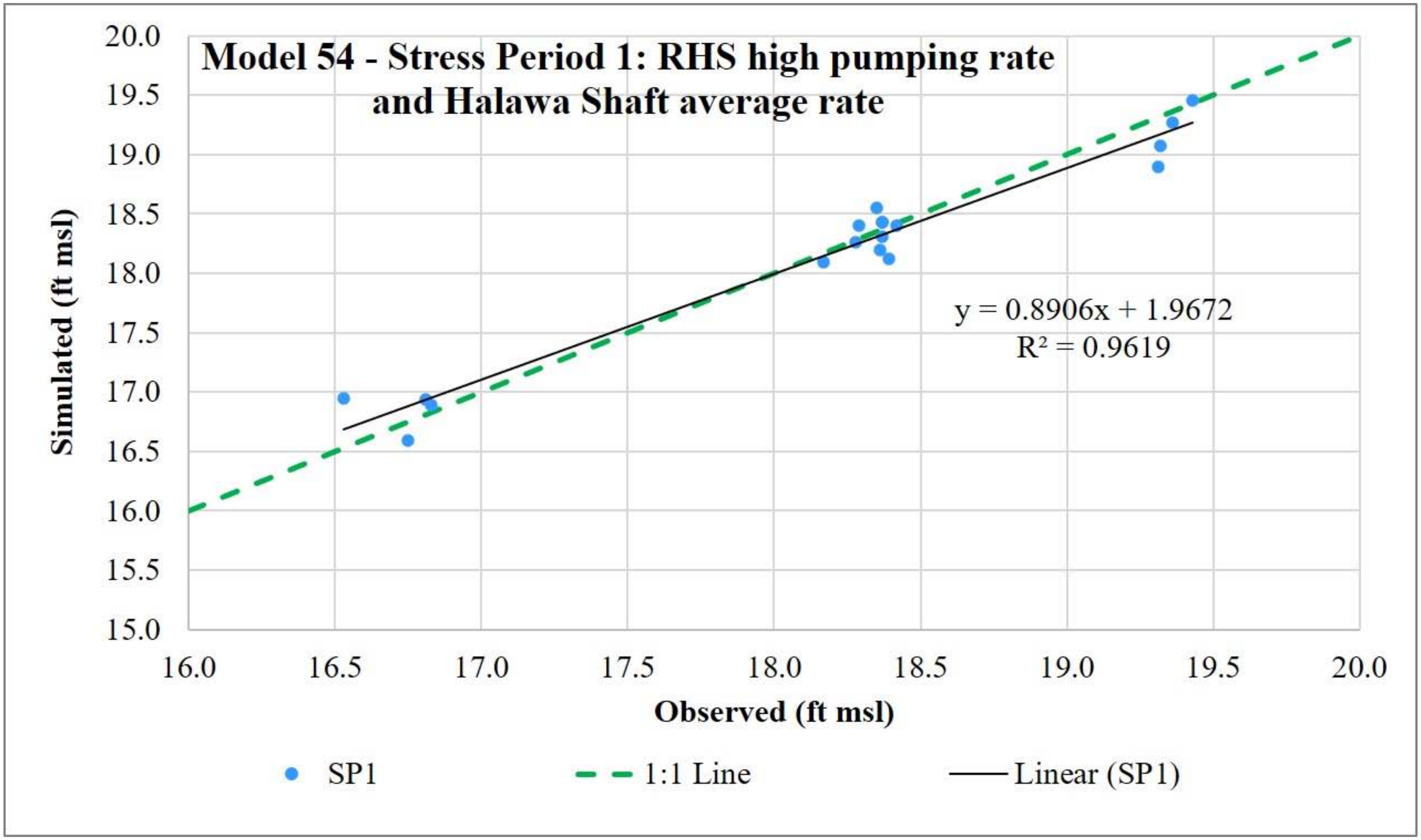
GWFM Boundary Conditions



Model Validation – Compare to Site Data

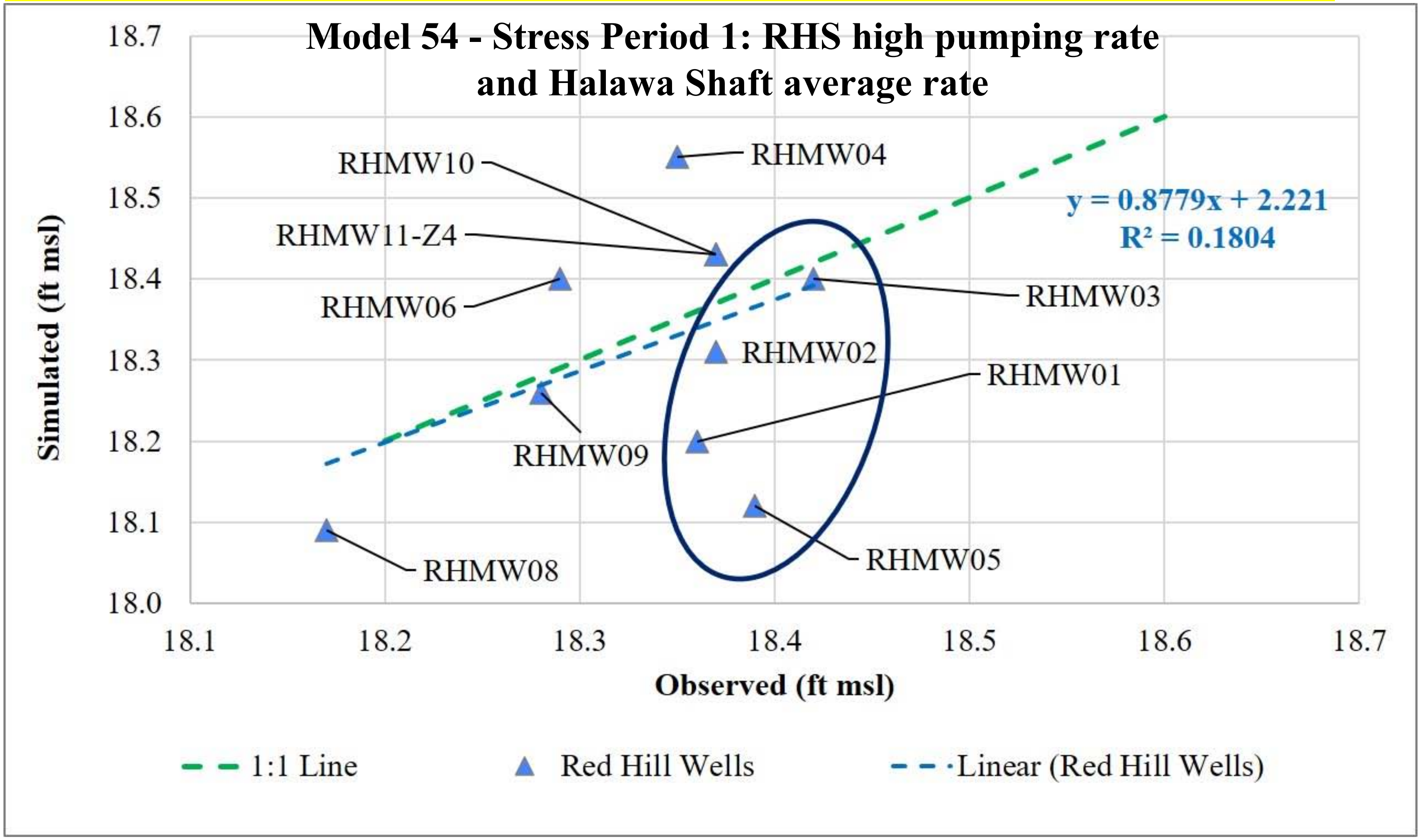
1. Methods & data currently used
    a) Groundwater elevations
    b) Transient responses
    c) Others
2. Concern with current comparative data
    a) Mis-match between modeled and measured gradients
    b) Groundwater elevations have low accuracy (Part II review)
3. Alternative groundwater behavior data
    a) Chloride and other natural tracers

Regional Water Levels

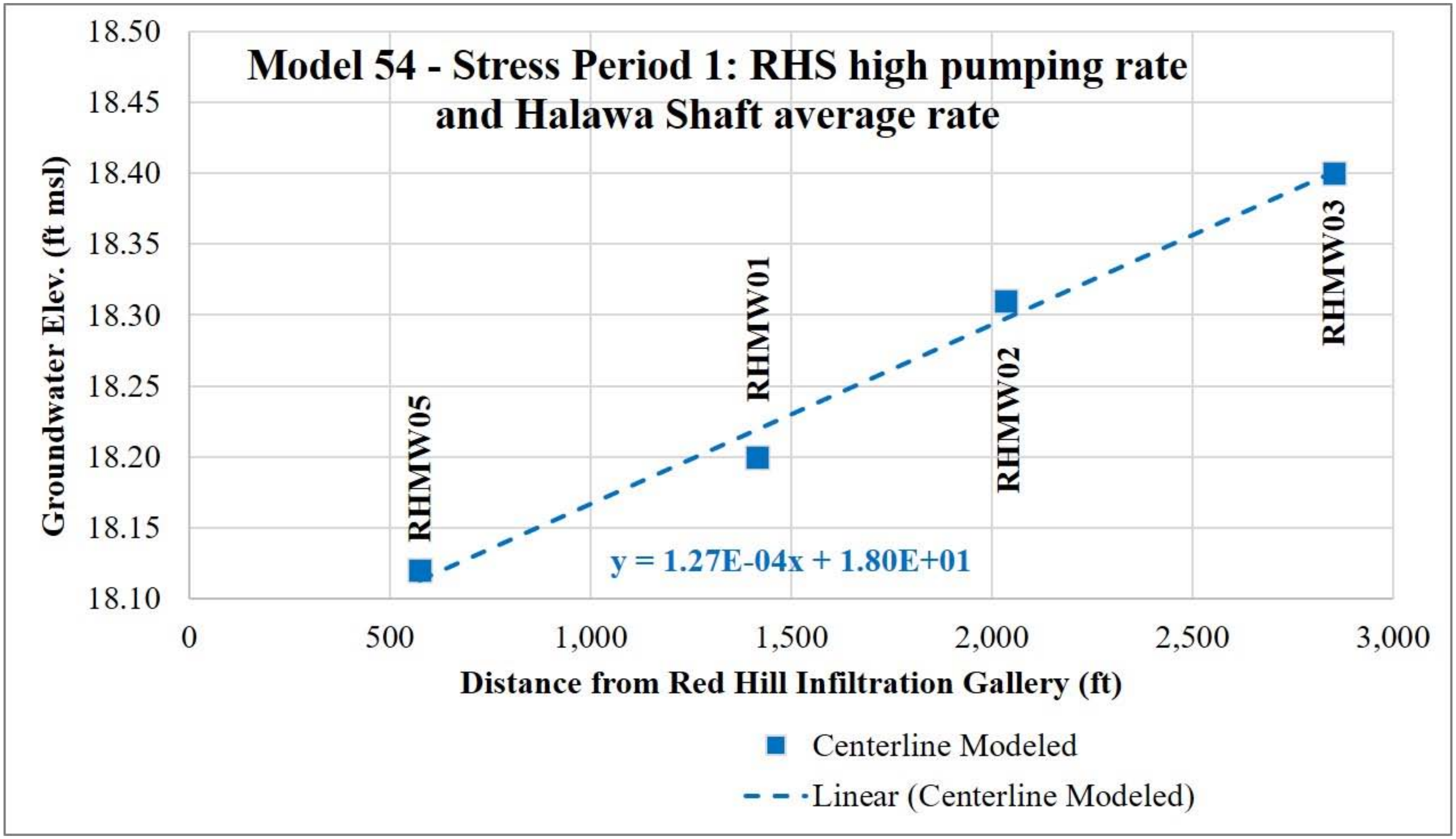


Local Water Levels

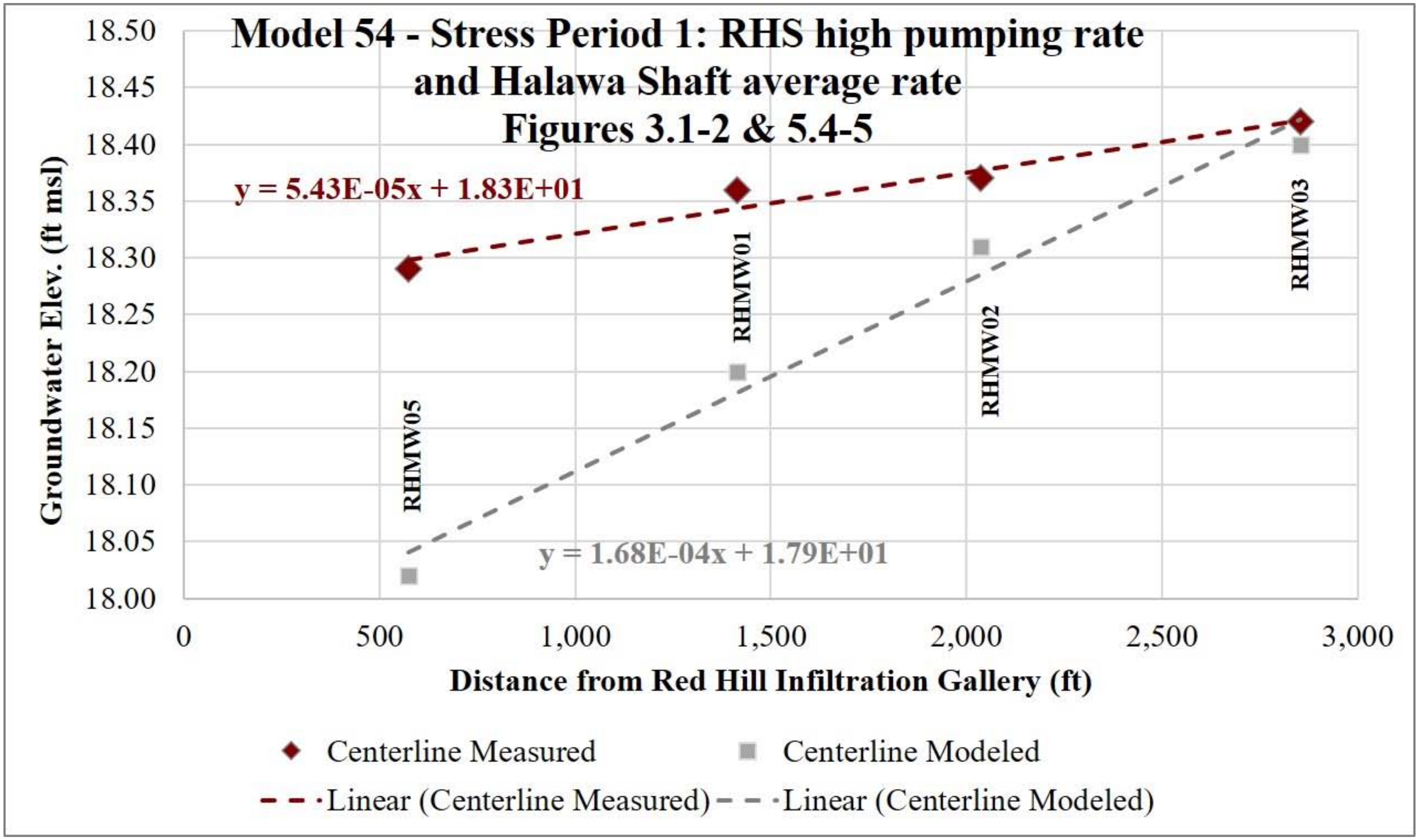
(Navy GWFMs do not match local data)



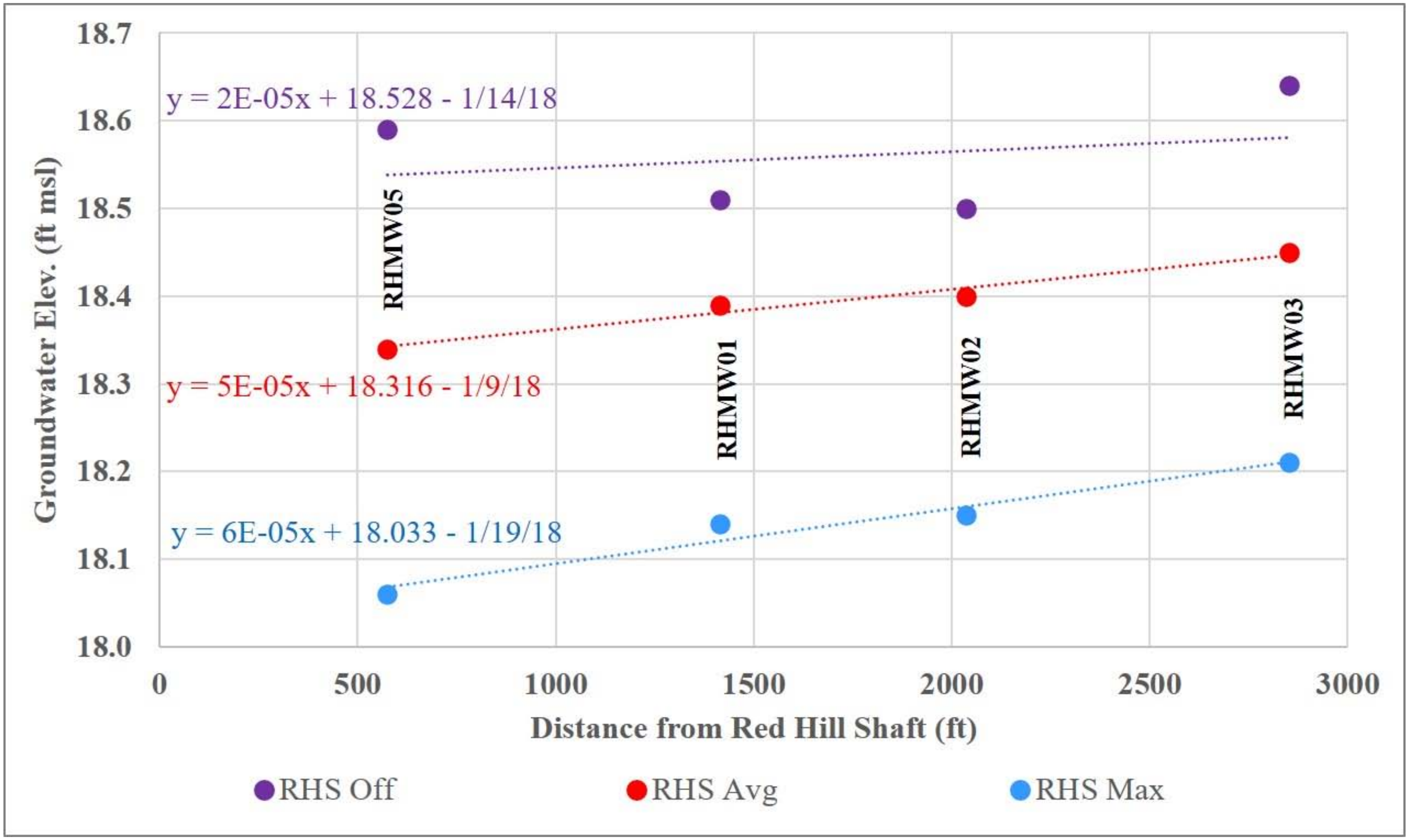
Gradient beneath and downslope of the tanks
(output from Navy GWFMs)



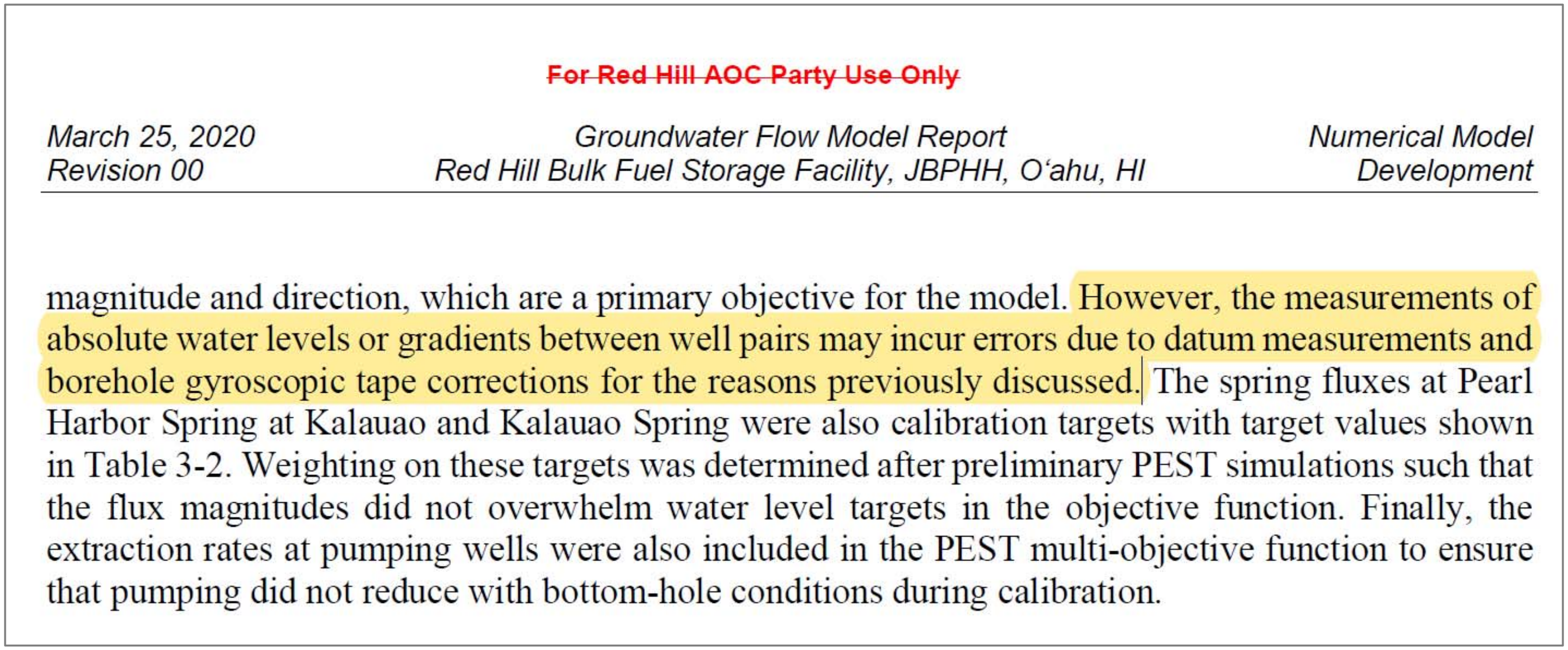
Measured vs. Modeled RH Ridge Gradients
(Gradient beneath and downslope of the tanks)



Red Hill Ridge gradient - under three different pumping conditions



Reliability of GW Elevation Data



~~For Red Hill AOC Party Use Only~~

March 25, 2020
Revision 00

Groundwater Flow Model Report
Red Hill Bulk Fuel Storage Facility, JBPHH, O'ahu, HI

Numerical Model
Development

magnitude and direction, which are a primary objective for the model. However, the measurements of absolute water levels or gradients between well pairs may incur errors due to datum measurements and borehole gyroscopic tape corrections for the reasons previously discussed. The spring fluxes at Pearl Harbor Spring at Kalauao and Kalauao Spring were also calibration targets with target values shown in Table 3-2. Weighting on these targets was determined after preliminary PEST simulations such that the flux magnitudes did not overwhelm water level targets in the objective function. Finally, the extraction rates at pumping wells were also included in the PEST multi-objective function to ensure that pumping did not reduce with bottom-hole conditions during calibration.

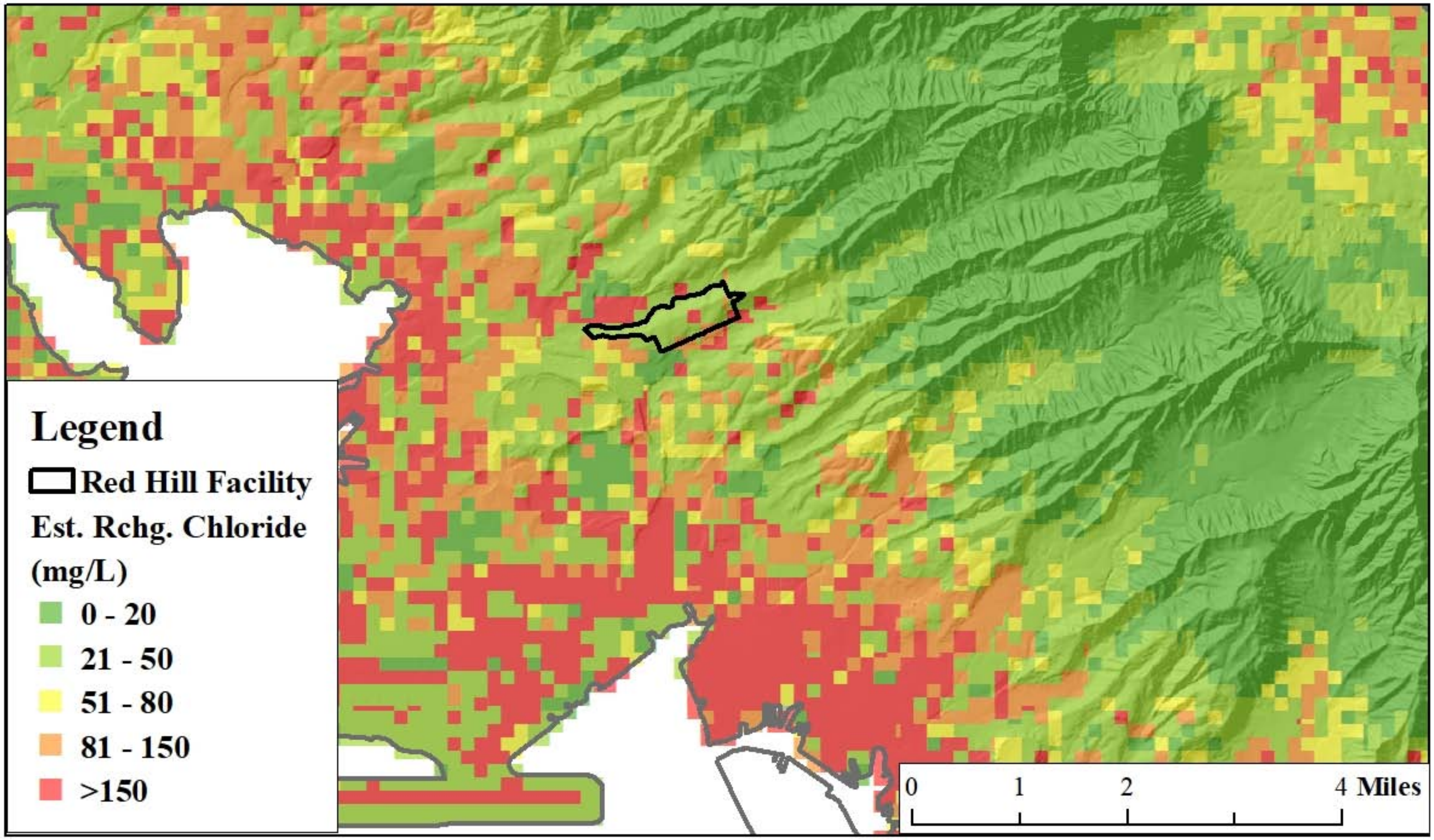
9. Groundwater Data

Imagery ©2018 DigitalGlobe, Google, Map data ©2018 Google

Chemistry shows indication of a poorly mixed system

- Chloride conc. vary from ~40->1000 mg/L
- Southeast very different from northwest
- Northwest chlorides still highly variable
- A large flux of groundwater down the Red Hill ridge should show better mixing

Estimated Chloride Conc. in Recharge

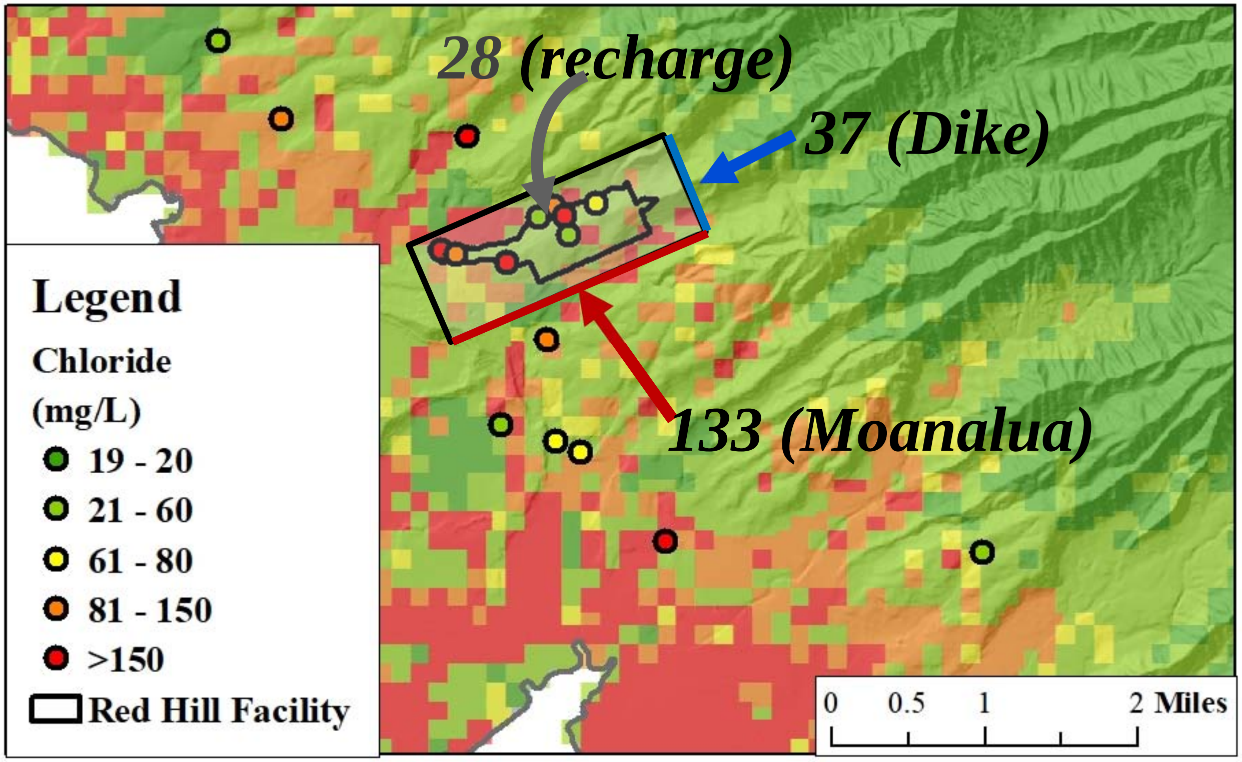


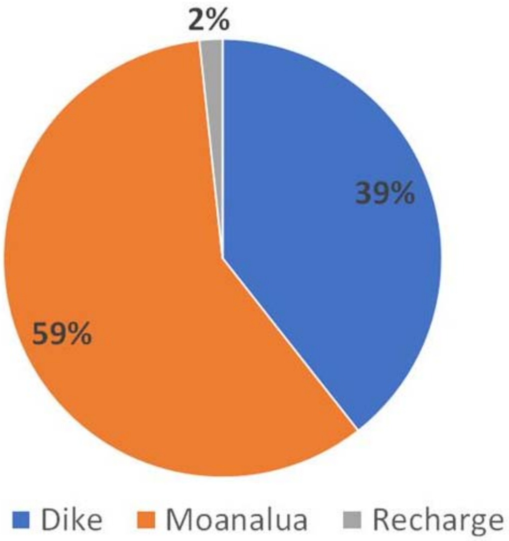
- Chloride in recharge estimated using the chloride mass balance approach
- Chloride concentration at the Facility <50 mg/L
  - Except for one pixel

Using Geochemistry to Refine Models

(without needing explicit CF&T simulations)

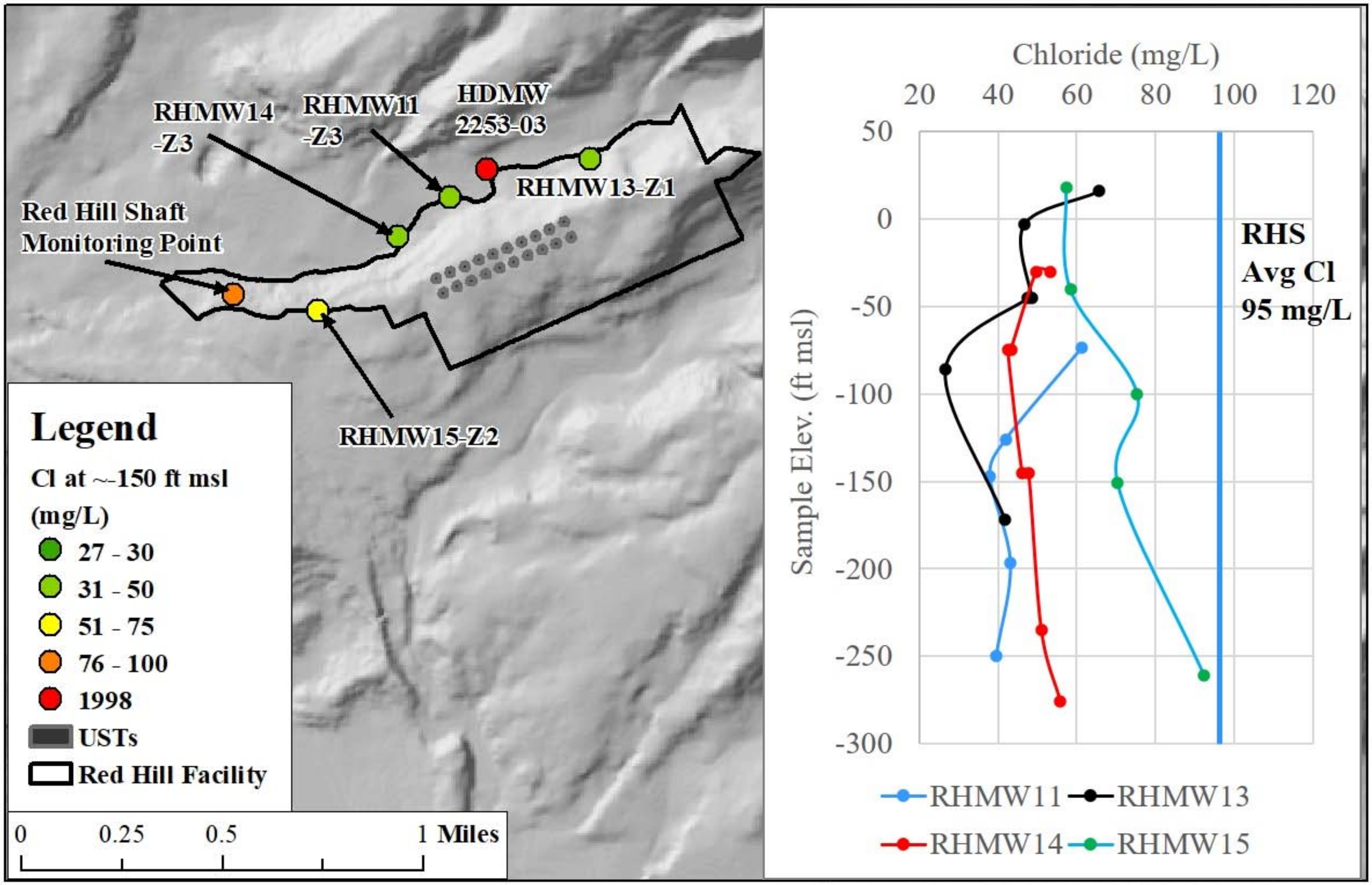
- Mixing Equation
  - $C_{mix}=(C_1*Q_1+C_2*Q_2+C_3*Q_3)/(Q_1+Q_2+Q_3)$
  - 93 mg/L = 2%*28 + 38%*37 + 59%*133
- Red Hill Shaft average chloride conc. ~ 95 mg/L
  - Chloride concentration is weighted Cl sum from the source areas



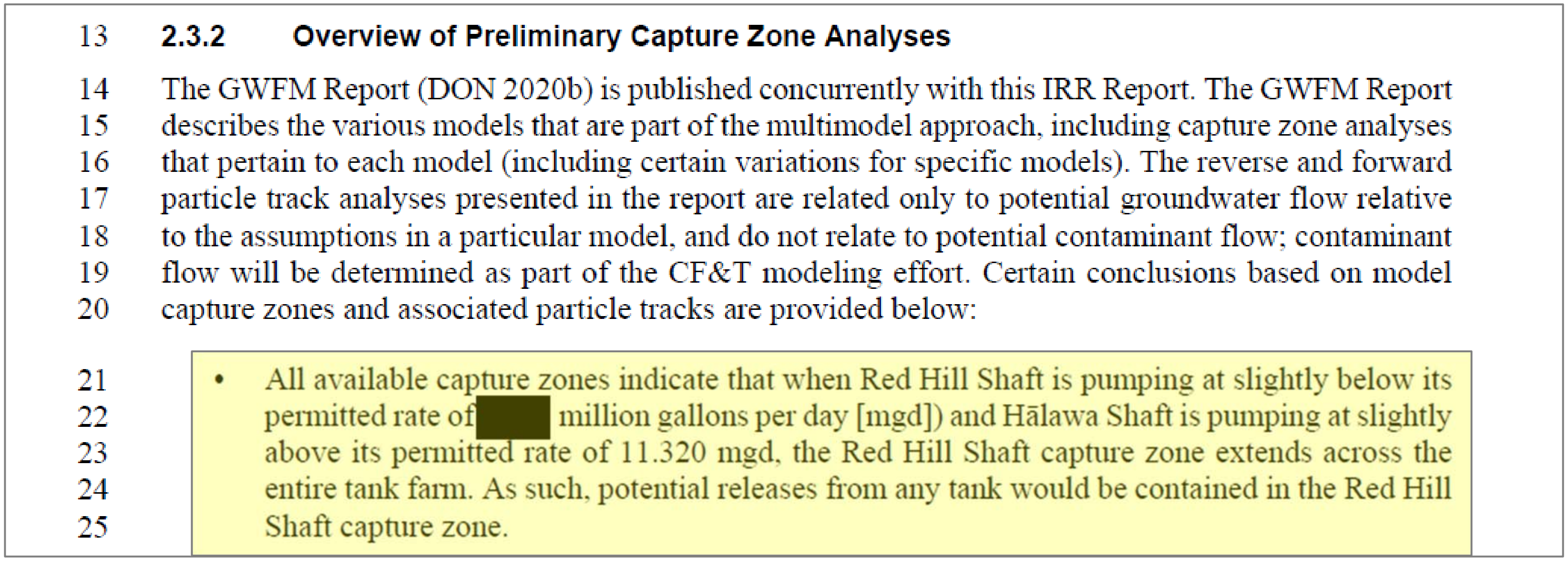


Numbers denote assumed chloride concentration

It is unlikely that chlorides originating in the Halawa region elevated the chloride concentration in the RHS



Application of model conclusions



2.3.2 Overview of Preliminary Capture Zone Analyses

The GWFM Report (DON 2020b) is published concurrently with this IRR Report. The GWFM Report describes the various models that are part of the multimodel approach, including capture zone analyses that pertain to each model (including certain variations for specific models). The reverse and forward particle track analyses presented in the report are related only to potential groundwater flow relative to the assumptions in a particular model, and do not relate to potential contaminant flow; contaminant flow will be determined as part of the CF&T modeling effort. Certain conclusions based on model capture zones and associated particle tracks are provided below:

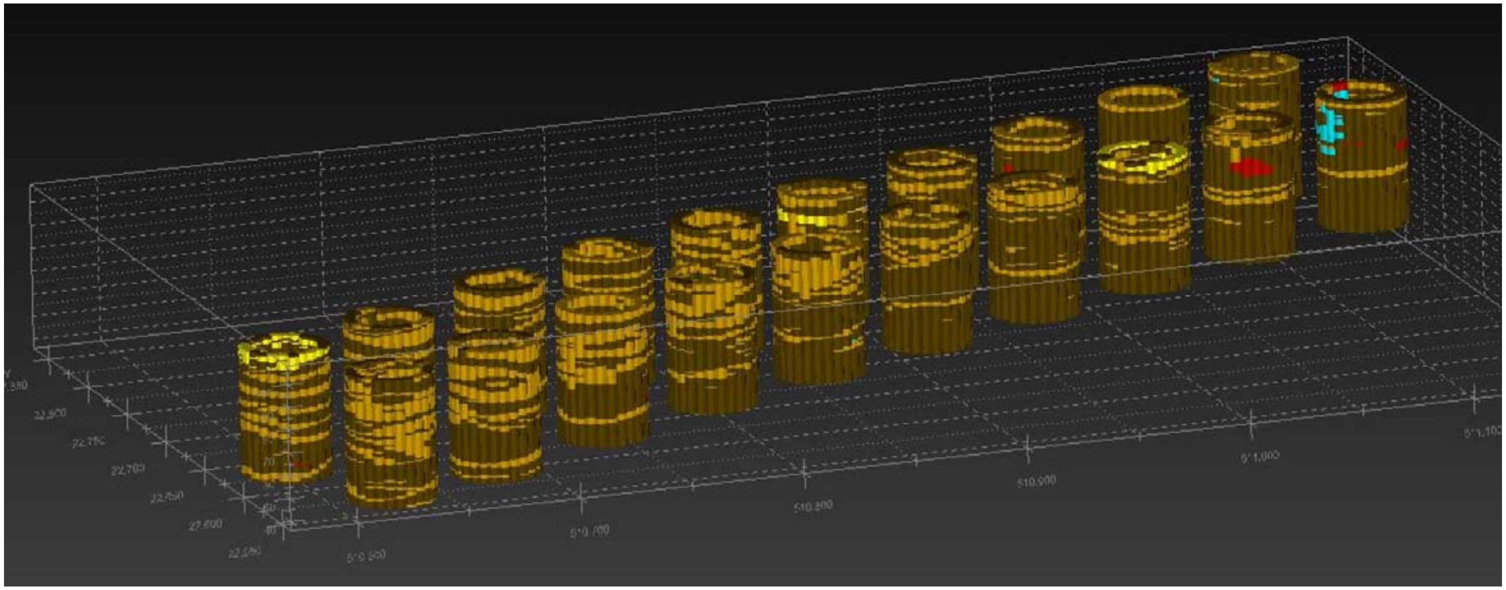
- All available capture zones indicate that when Red Hill Shaft is pumping at slightly below its permitted rate of [redacted] million gallons per day [mgd]) and Hālawa Shaft is pumping at slightly above its permitted rate of 11.320 mgd, the Red Hill Shaft capture zone extends across the entire tank farm. As such, potential releases from any tank would be contained in the Red Hill Shaft capture zone.

- Investigation and Remediation of Releases Report; Page 2-18
- Issues previously discussed cast doubt on the assumption the Red Hill Shaft will contain the offsite migration of any contaminant plume
- The model results are currently not informative for developing release response plans
  - Questions regarding the ability of the RHS to capture a contaminant plume and the risk the Halawa Shaft remain unanswered

Further GWFM & CSM Review Items

DOH Technical Team:

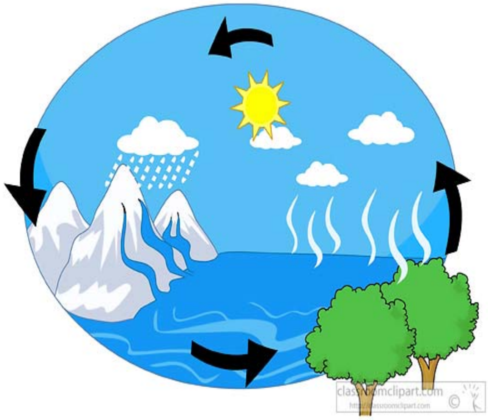
Dr. Thomas & Rowland, UH
Robert Whittier, DOH/SWPP
G.D. Beckett, C.Hg.
Anay Shende, DOH
Dr. Matt Tonkin, EPA (review)



Key Groundwater Model Objective

- The purpose of this deliverable is to refine the existing groundwater flow model and improve the understanding of the direction and rate of groundwater flow within the aquifers around the Facility (AOC, 2015)
  - *To do this, the underlying hydrogeologic conditions must be refined and better understood in light of new data not available to prior modeling*

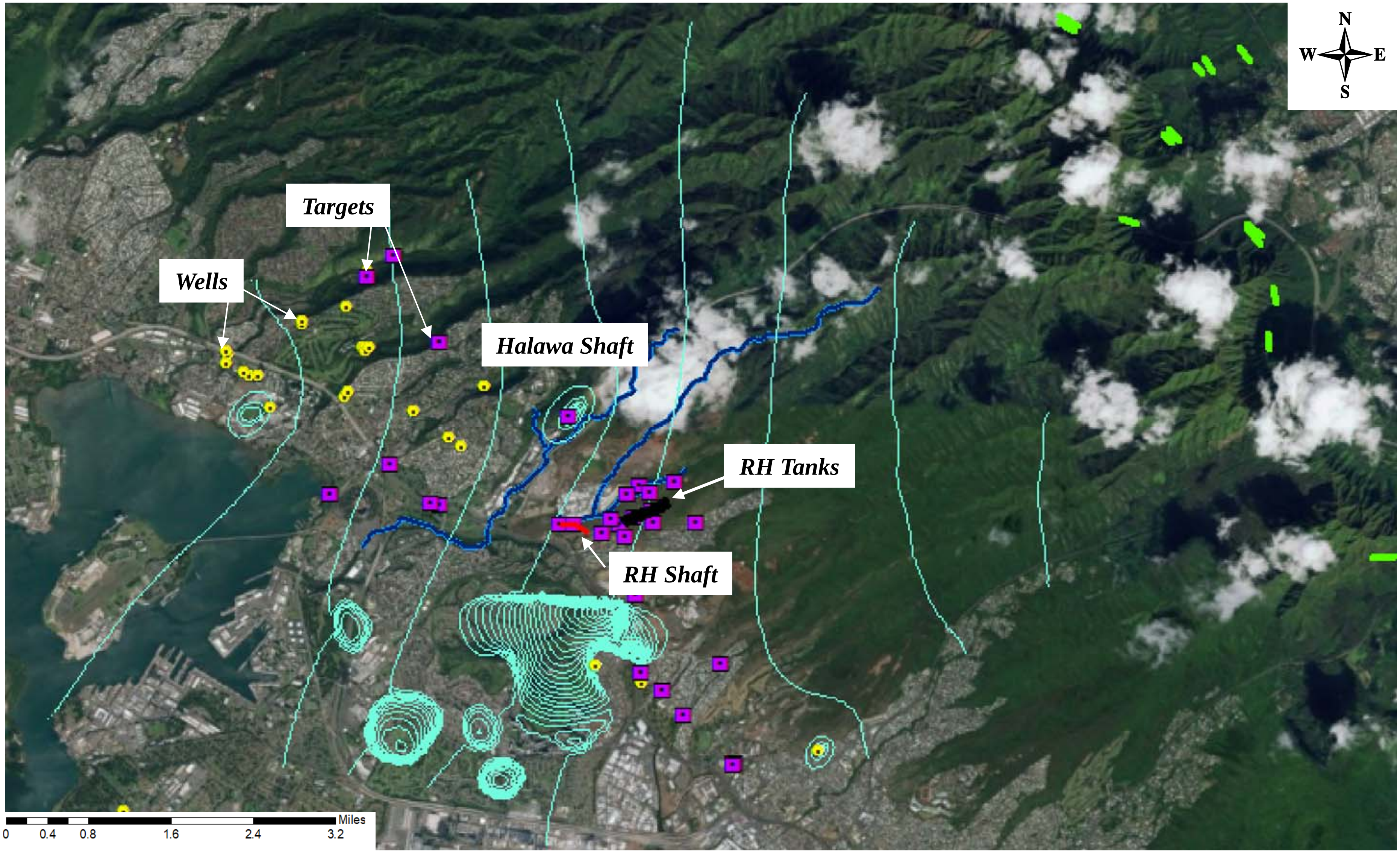
The Navy Has Delivered Multiple Models



- Key review questions:
  - Do the models represent local heads?
  - Gradients?
  - Transient aspects?
    - Pumping from Red Hill & Halawa shafts
    - Monitoring well response “groupings”
  - Do transient simulations better past models?
  - Are models consistent with geochemistry?
    - And with dissolved-phase patterns?
  - Are models parameters appropriate?
- Will the model(s) inform risk estimates?
  - Most uncertain aspect is NAPL
    - Where is it presently & in what state?
    - How far/fast could releases travel?
  - What are the key processes?
  - Are those adequately described & demonstrated?

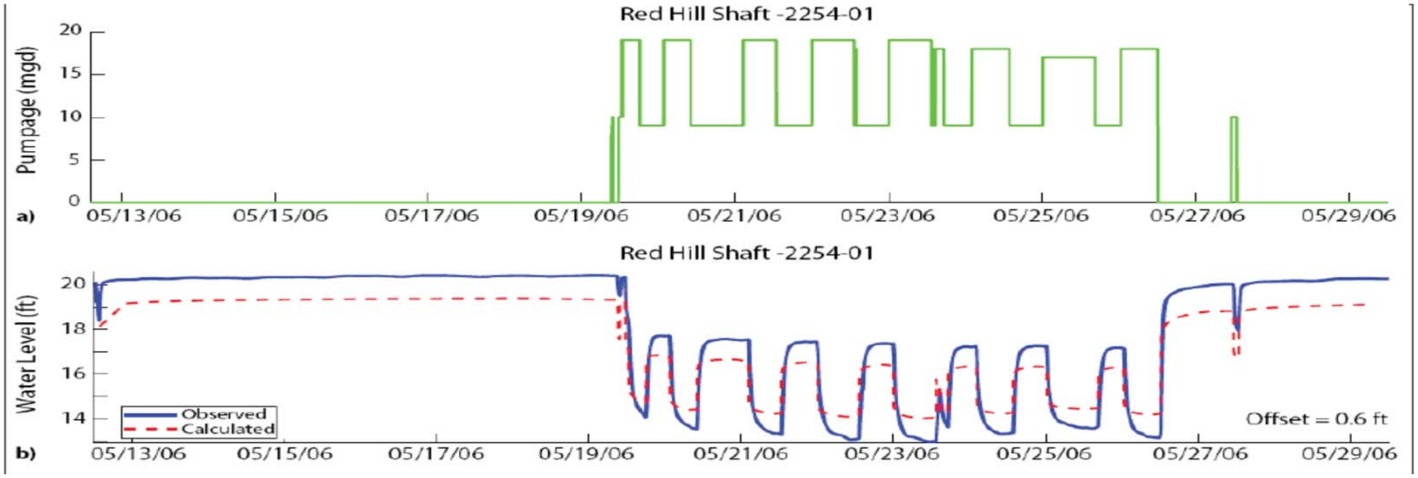
General Area/Model Map

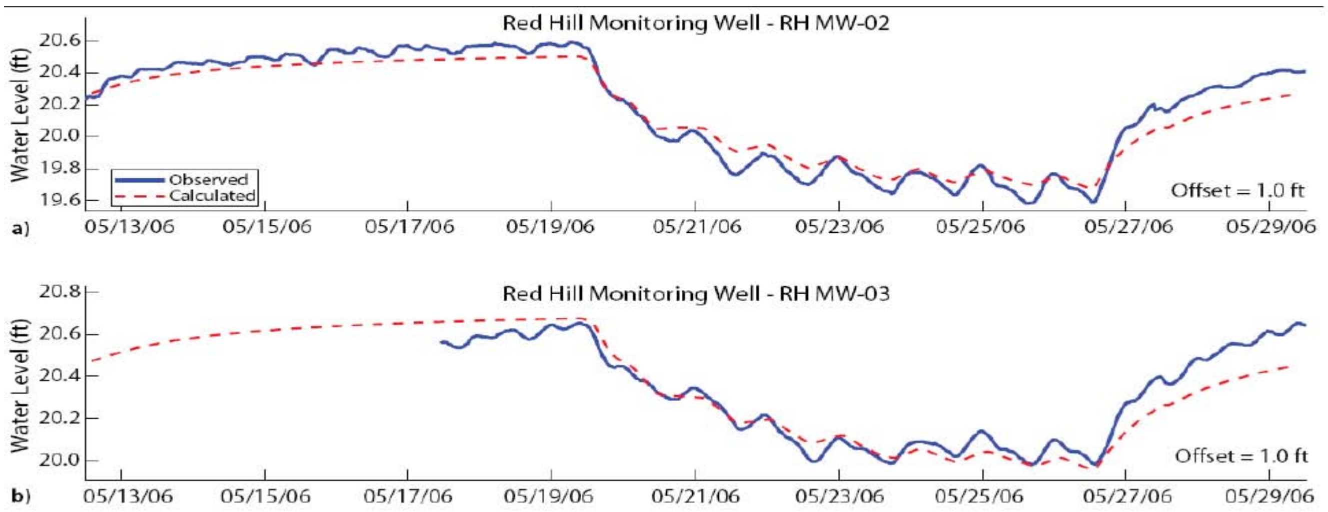
(Halawa Shaft On, RH Shaft Off)



The Primary Issue with the Prior Model

(calibrated to drawdown, but not to heads; complexity)

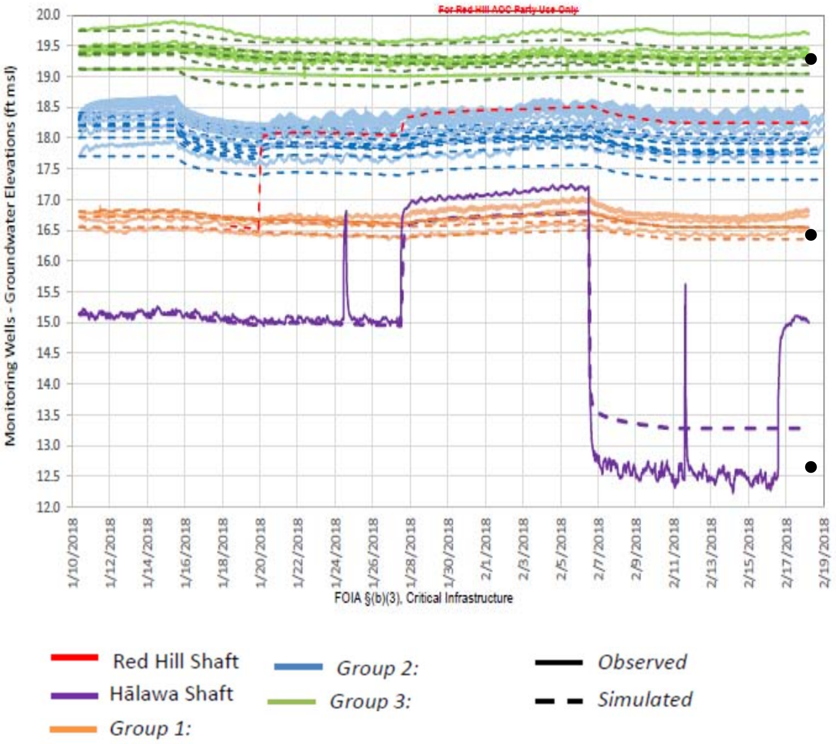




Kolja Rotzoll and Aly I. El-Kadi, 2007

Objectives of Verification Models

(GWFMs apparent mismatch to g.w. elevations)



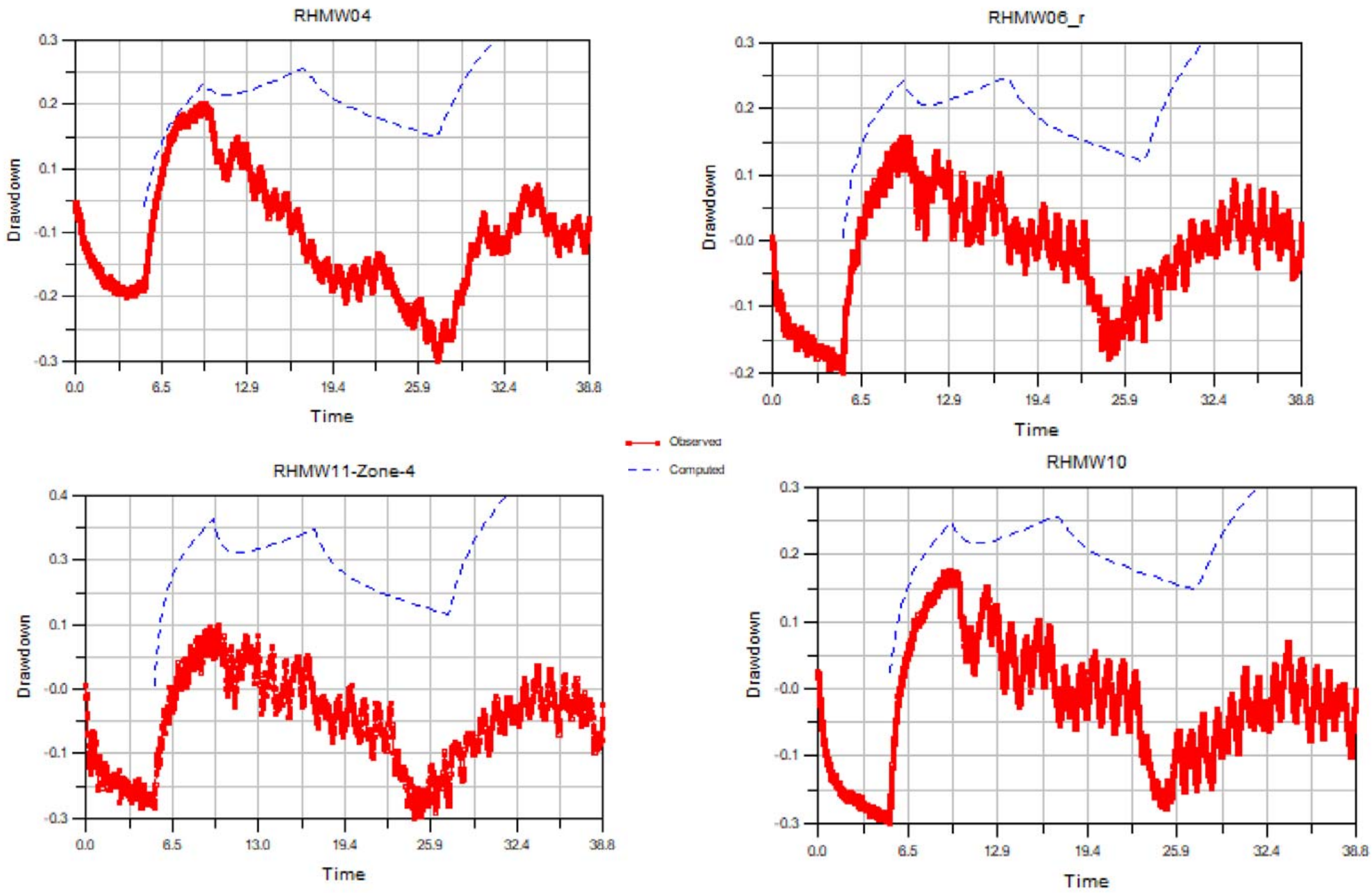
Abbreviated legend

Figure 5.1.1-7, Redacted GWFM Rept, Mar 2020

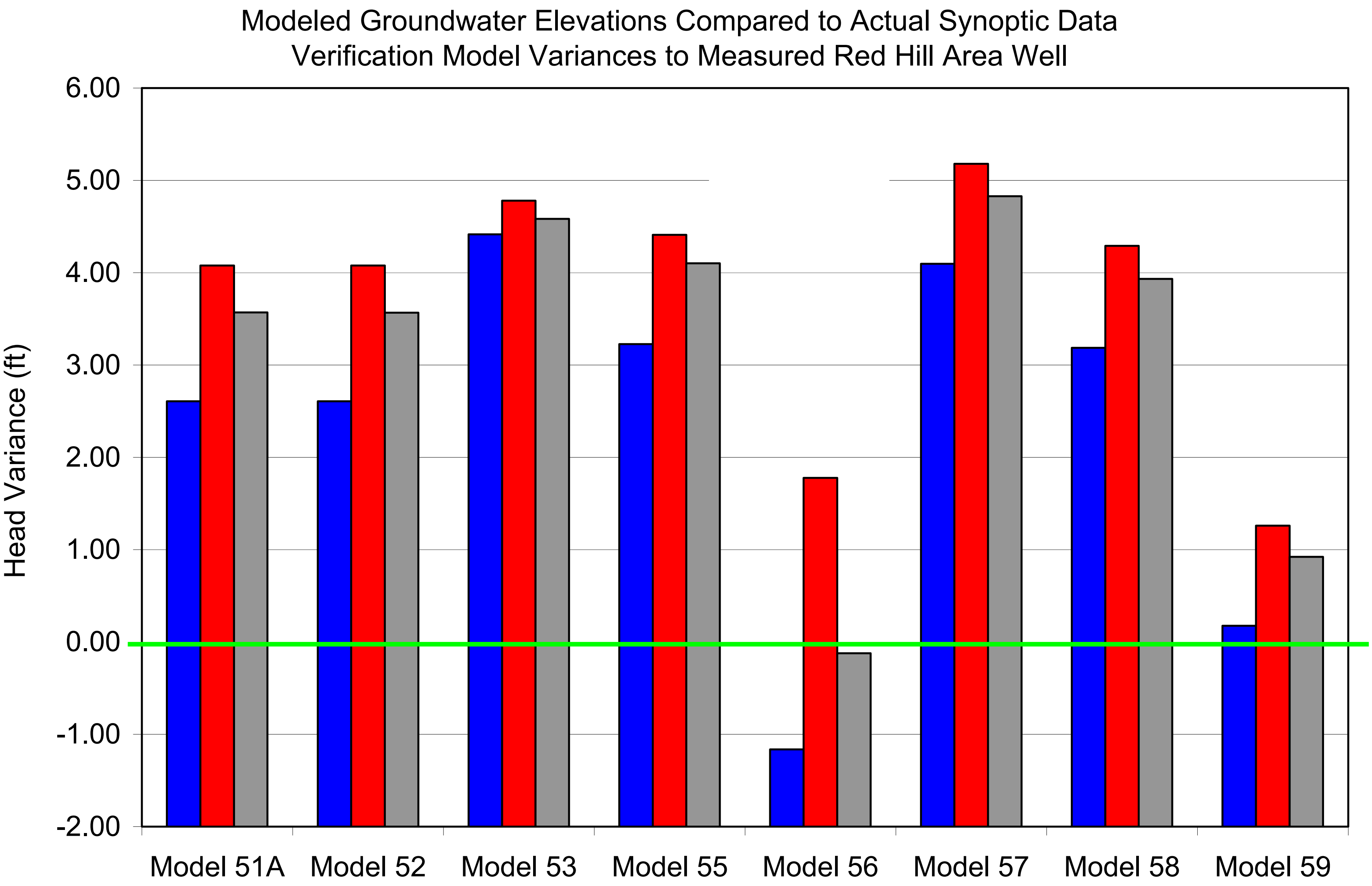
- Verification means just that
  - A "blind" test of the GWFMs predictions
  - How well do they agree with elevation data?
- How is this typically implemented?
  - Calibrate main models
  - Run against site data from another time
  - See how well each model reflects the data
- Purpose
  - Identify deficiencies in main models
  - Identify which are "best fits"
  - Consider transient implications
  - Consider compartmental responses (& others)
- Issue, we cannot replicate the reported results
  - Plots do not agree with modeled output
  - May be a superposition (drawdown upon measured)
- The g.w. elevation offset was prior model issue
  - Recall primary AOC objective

Example Hydrographs; M51a Verification

(charts are direct model output – GWV)

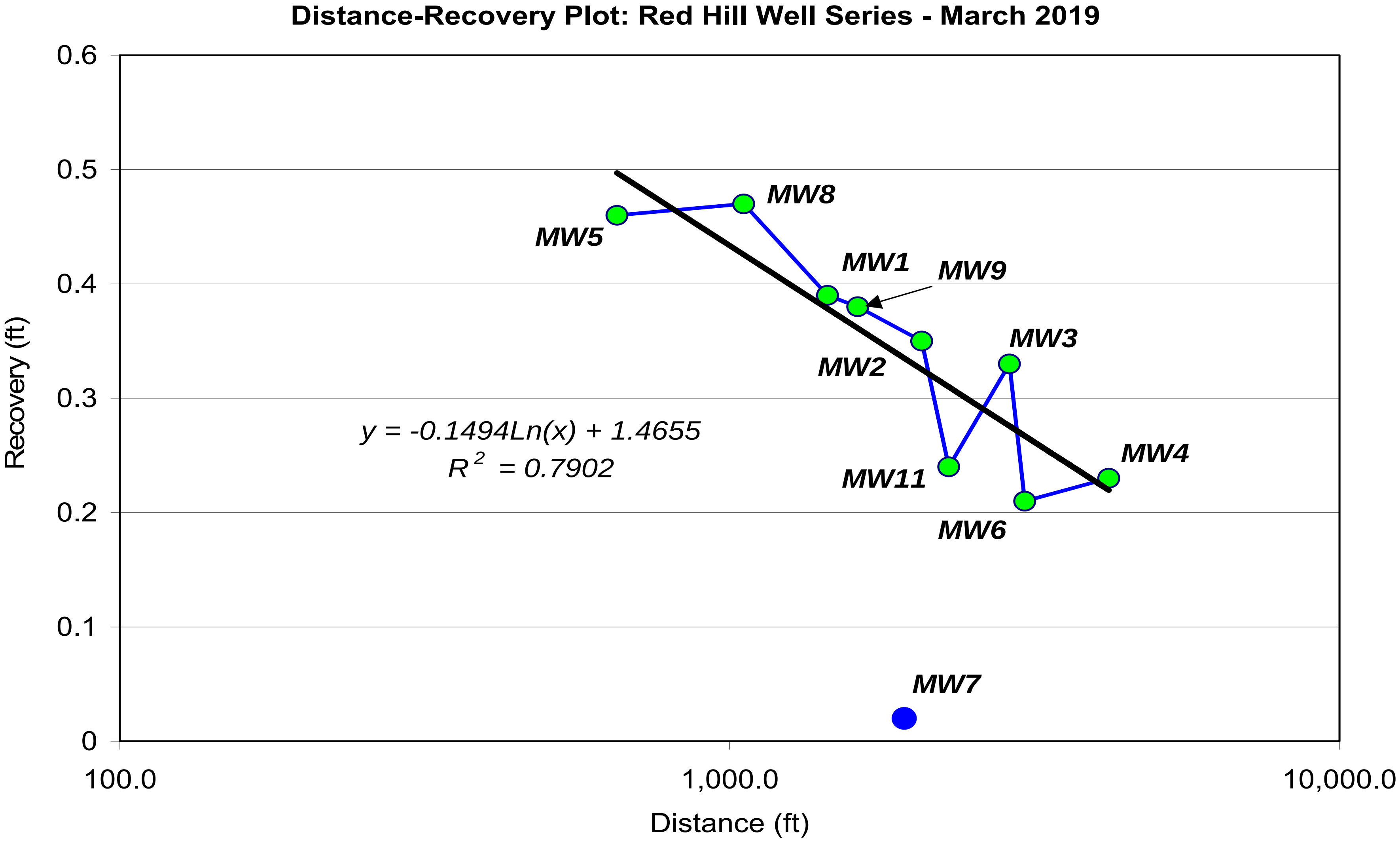


GW Elevation Variance – Transient Models



Non-Uniform Distance Drawdown Behavior

(indicates complexities not captured by models)

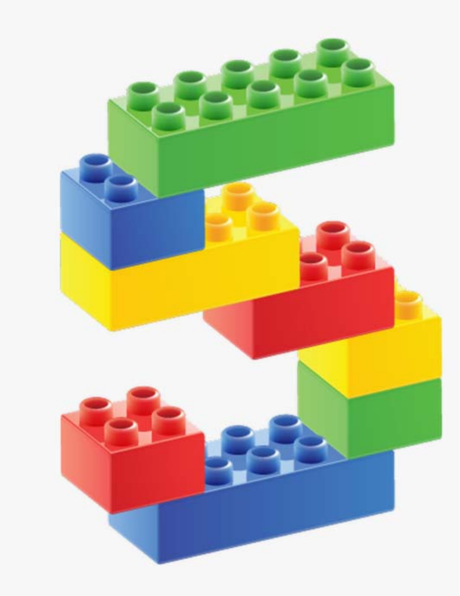


Prior Key Parameters v. Navy Models

(ranges are inconsistent & w/o explanations)

	Oki, 2005			**Navy GWFM - avgs**		
Hydrostratigraphic Unit	Kv	Kt	Kl	Kv	Kt	Kl
Volcanic-rock aquifer	7.5	1,500	4,500	65	1,000	2,999
Caprock, upper-limestone unit	25	2,500	2,500	0.01	500	500
Caprock, low-permeability unit						
Above Waianae Volcanics	0.3	0.3	0.3	0.01	1	1
Above Koolau Basalt, west of Waiawa Stream	0.01	0.01	0.01	0.01	1	1
Above Koolau Basalt, east of Waiawa Stream	0.6	0.6	0.6	0.01	1	1
Valley-fill barriers	0.058	0.058	0.058	0.01	1	1

Key Model Review Observations



- GWFMs do not match heads, diminishing reliability
  - Particularly in transient verification runs
  - Similar issue as in prior modeling (2007)
- GWFMs use atypical parameters for Hawaii aquifer
  - If retained, in-depth justification needed
- GWFMs do not use CSM geologic details – SSPA work
  - Impact of heterogeneity needs detailed evaluation
- GWFMs do not comport with natural g.w. tracers
  - Complex distributions may imply multiple source waters
- GWFMs capture zones not supported by field data at pumping rates similar to those modeled
  - Approaches used may overestimate capture potential
  - Gradient issues & complexity not covered
- The current GWFMs are not reliable for decisions
  - For CF&T, risk analyses and mitigation decisions
- Modifications will be needed (SSPA work follows)

Ongoing Issues with the Navy CSM

The CSM being the fundamental basis for the Navy GWFMs,
future CF&T/Risk Evaluations and the overall key conditions at the
Red Hill Bulk Fuel Storage Facility

The Hawaii Hard Rock Release Experience

Source: Don Thomas, 2021

- Fuel releases often move quickly
    - Typically in complex pathways
        - Primary & secondary transport
    - Often difficult to characterize
- Fast-track/other geologic features exist
    - Lava tubes, voids, fractures, clinkers
    - Confining beds & non-volcanics
    - Preferred & random orientation scales
    - Often sparse distribution, large effect
- Weathering of rock is complex
    - Bulk rock properties may not apply
- For Red Hill
    - How is the architecture arranged?
    - How will fuel behave within that?
    - Effects on capture/remediation?
    - All relates to g.w. protection goals
        - And sole source aquifer preservation

Overview – Unresolved CSM Issues

Source: Dr. Scott Rowland, 2021

- Red Hill is under-characterized
  - Compared to similar sites
  - Results in high uncertainty in the CSM
- Complex geology is noted in CSM
  - But, simplified in GWFMs
  - Insufficient basis for appropriate CF&T
  - G.W. & CF&T behavior appears more complex
- Data indicate TPH beyond RH Ridge
  - CSM interprets these as artifacts (generally)
- CSM interprets LNAPL migration to SW
  - But available data indicate otherwise
- CSM indicates fuel retained ~ 30-ft depth
  - Not supported by available data
- Fuel retention characteristics are unknown
  - Fuel/NAPL parameters inapplicable
  - Geometry unconstrained by data
  - Dynamics are critical to g.w. protection
- Many other issues remain
- In total, CSM is not reliable for g.w. protection

Example Navy CSM Cross-Section

(schematic rendering, but details are not in GWFMs)

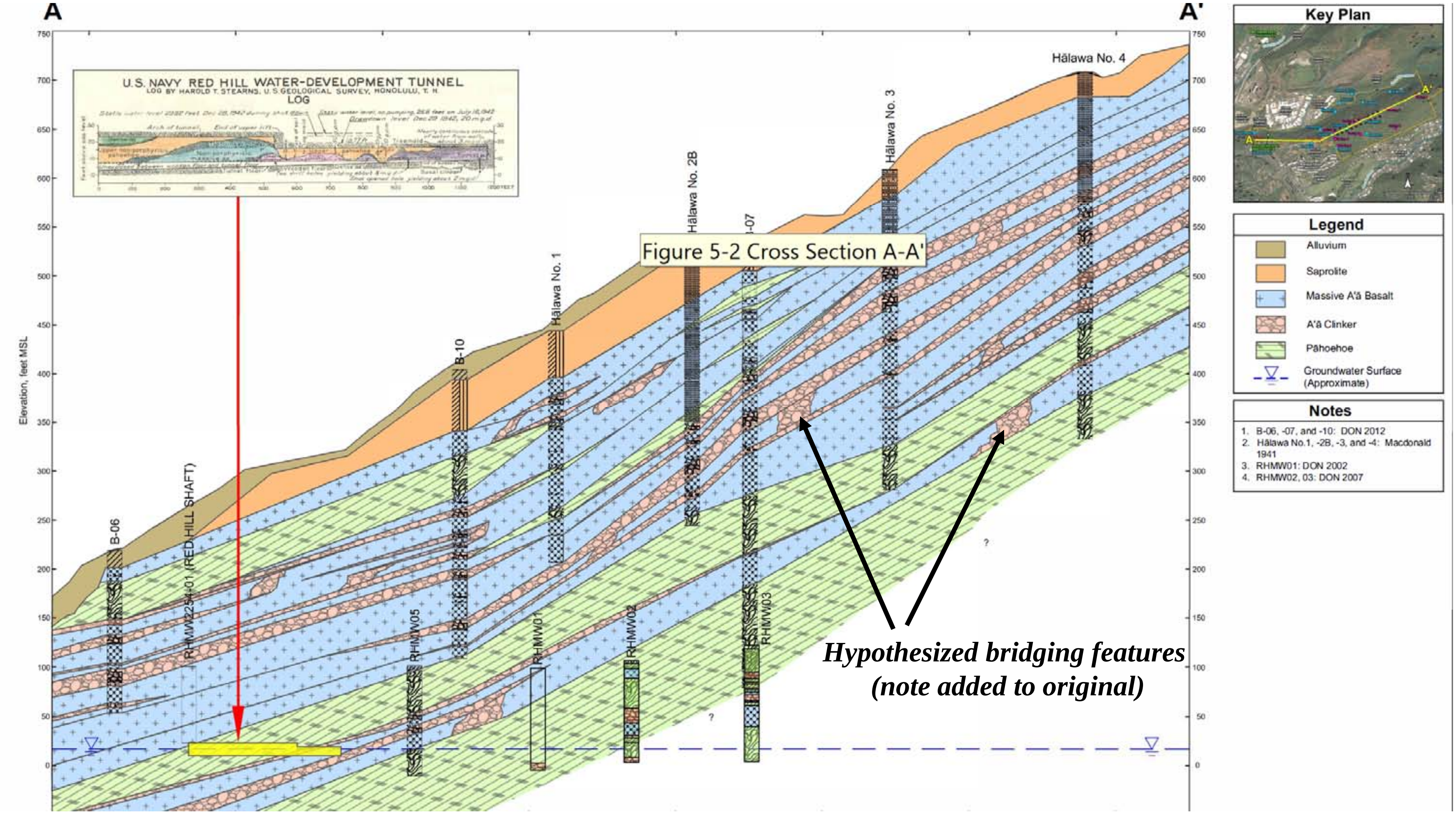


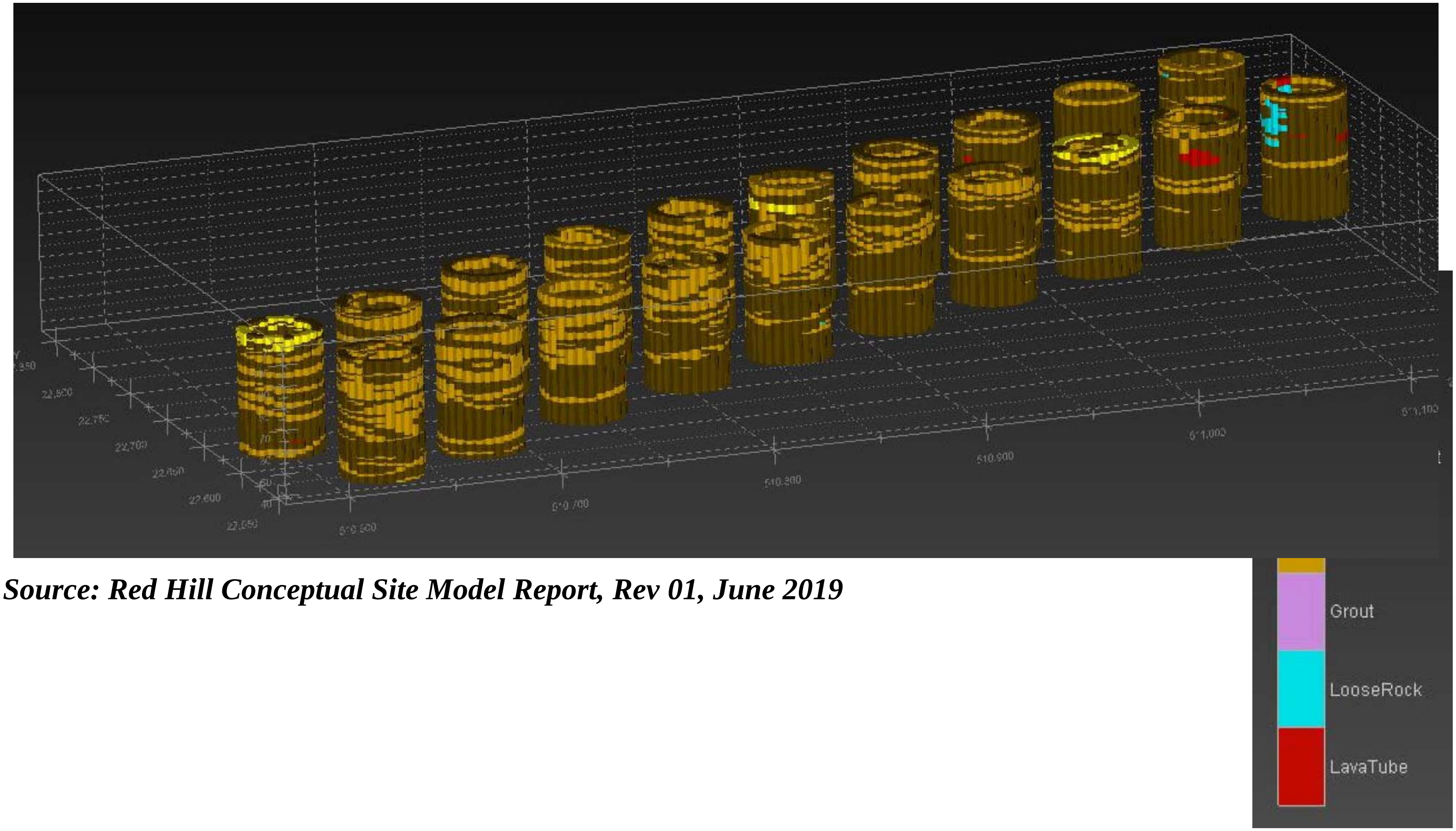
Figure 5-2 Cross Section A-A'

Hypothesized bridging features (note added to original)

Source: Red Hill Conceptual Site Model Report, Rev 01, June 2019

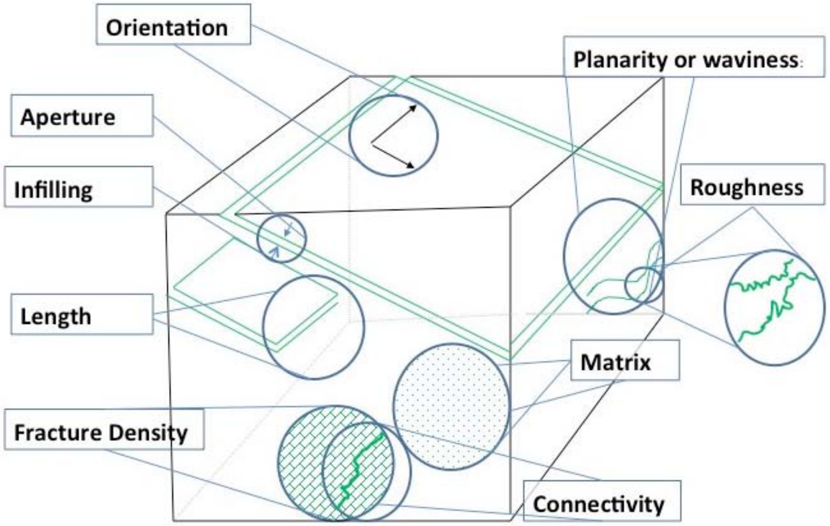
Navy 3D Lithologic Model – Barrel Logs

(same issue, Dr. Tonkin will address)



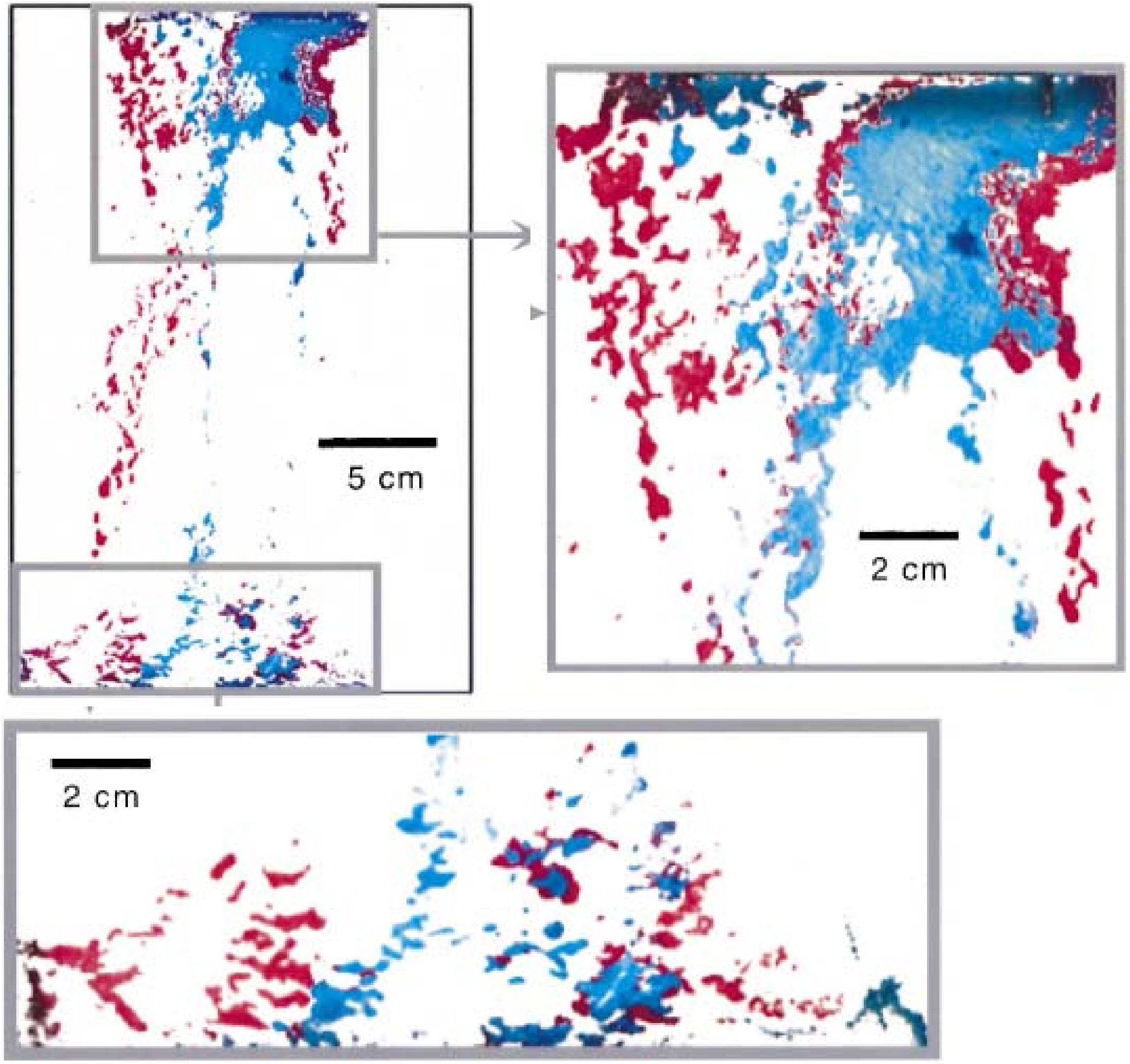
Source: Red Hill Conceptual Site Model Report, Rev 01, June 2019

Outcrop Interpretation – Dr. Scott Rowland (UH)



Source: ITRC, 2017

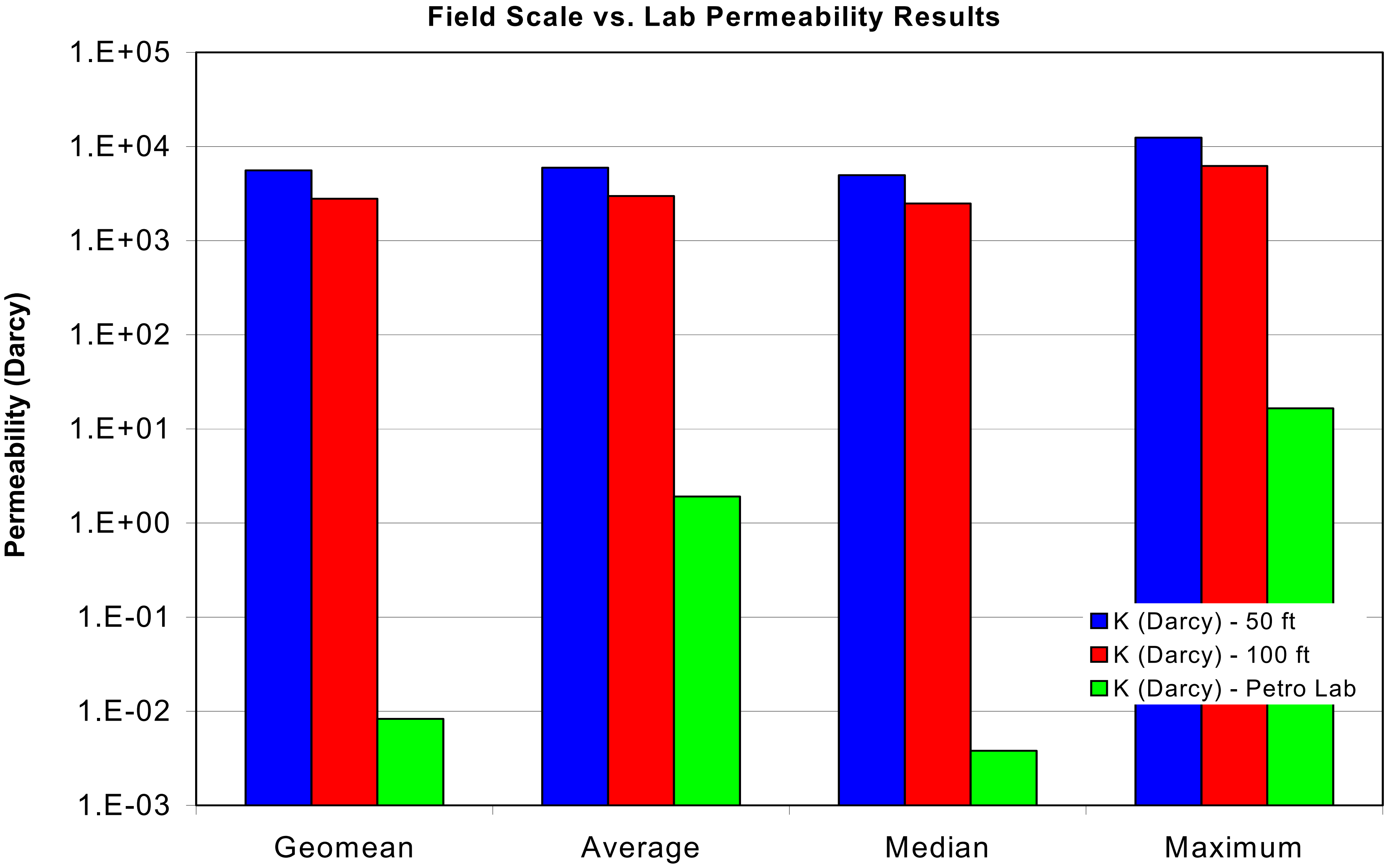
Complex NAPL Distribution in a Fracture



Geller et al., 2000

Lab vs. Field Scale – Permeability

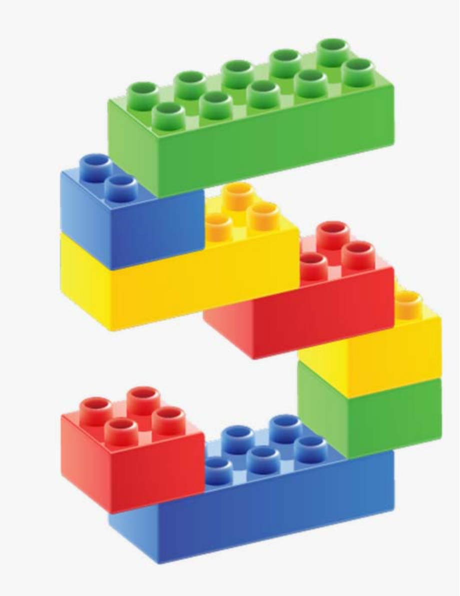
(Site lab data are not comparable to field scale)



Data source: Conceptual Site Model, June 2019, Rev 01

Summary of CSM Review (to date)

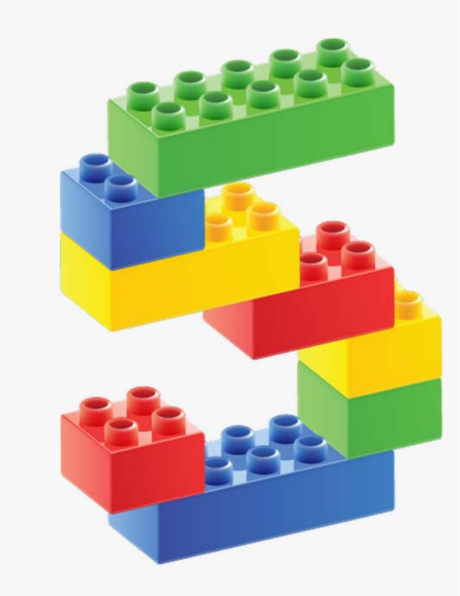
(broad issues, other details pertain)



- Issues have been ongoing, unresolved by new data
  - Or interpretations unconstrained by available data
- The site is not well characterized – (safety concerns)
  - Fate of 2014 & 2021 releases are undelineated
  - Data suggest fuel has reached the water table under RH
- Geologic complexity noted in CSM
  - But not explored at the needed level of detail
  - No assessment of EPM scale or applicability
- Groundwater flow paths and behavior is uncertain
- Distal detections are considered generally valid
  - Reported by certified labs & independently validated
  - There is TPH-range mass in GCs
  - Detections are consistent with other data/patterns
- NSZD rates are likely overestimated & uncertain
  - RHMW03 & RHMW01, net thermal profiles, no NAPL
  - Plume size and character likely larger than estimated
- The whole of the RH Tank Farm has likely had releases
  - CSM does not account for long & variable fuel history
  - And those implications for CF&T/risk/mitigation

Implications of CSM Concerns

(relative to groundwater protection matters & TUA)



- G.W. capture of releases is not demonstrated
  - By field data or adequately by GWFMs
- NSZD may not be reliable as a cleanup method
  - RHMW03 interpreted impacts remain > 20 yrs
- G.W. protection depends on several factors
  - How fuel migrates under release conditions
  - Speed and effectiveness of release detection & actions
  - Cannot be addressed by GWFMs alone
- Capture may not be an appropriate G.W. remedy
  - Fuel migration & remedy must be aligned
  - Capture is not a cleanup method – relies on uncertain NSZD
  - ***However, g.w. treatment may protect water services***
- Red Hill Shaft is indicated to be at risk from releases
  - Proximity & low-level TPH detections (including July 2021)
  - Dilution & NSZD make this both surprising & concerning
- Risk evaluations must be connected to a conservative CSM
  - Presently, there is insufficient conservatism in the CSM
  - Along with high uncertainty that is not addressed