

A Subsurface Decision Model for Compliance with NRC's License Termination Rule Criteria

**DRAFT (Version 2)
Robert Stewart
stewartrn@ornl.gov**

**Approved by ORNL for Release for
Public Comment**

A Subsurface Decision Model for Compliance with NRC's License Termination Rule Criteria

DRAFT (Version 2)
Robert Stewart
stewartrn@ornl.gov

Approved by ORNL for Release for Public Comment

Manuscript Completed: August 27th, 2025
Date Published: TBD

Prepared by:
Dr. Robert Stewart

Oak Ridge National Laboratory
1 Bethel Valley Rd
Oak Ridge, TN 37831

Cynthia Barr, NRC Project Manager

TABLE OF CONTENTS

1			
2	LIST OF FIGURES		vii
3	LIST OF TABLES		ix
4	ABBREVIATIONS		xi
5	1 INTRODUCTION		1
6	2 BACKGROUND.....		3
7	2.1 Overview		3
8	2.2 Multi-Agency Radiation Survey and Site Investigation Manual		3
9	2.3 Triad		5
10	2.3.1 Systematic Project Planning.....		8
11	2.3.2 Dynamic Work Strategies.....		8
12	2.3.3 Real-Time Measurement Technologies		8
13	2.4 NUREG-7021 First Edition		8
14	2.5 The GEM Framework		9
15	2.6 Electric Power Research Institute Geostatistics Review		10
16	2.7 NRC Subsurface Workshops		10
17	2.7.1 Subsurface DCGLs		11
18	2.7.2 Final Status Surveys		11
19	2.8 NUREG 1757 Vol 2, Revision 2.....		12
20	2.9 S. Cohen & Associates White Paper		12
21	2.10 Subsurface and Visual Sample Plan		13
22	2.11 Interoffice Staff Memorandum		13
23	2.12 Standards		13
24	2.13 NUREG-7021 Rev 1 Approach.....		14
25	3 GEM+ Foundations		15
26	3.1 Overview		15
27	3.2 Subsurface Survey Units.....		15
28	3.3 Exposure Mechanisms		16
29	3.4 Subsurface Classification		18
30	3.5 Preliminary Site Model		18
31	3.6 Geostatistical Modeling		20
32	3.6.1 Simulation.....		20
33	3.6.2 Kriging		21
34	3.6.3 Simulation vs. Kriging.....		22
35	3.6.4 Further Resources.....		23
36	3.7 Geostatistical Site Model.....		23
37	3.8 Virtual Scanning		25
38	3.9 Multiple Radionuclides		26
39	4 GEM+ Workflow		28
40	5 Historical Site Assessment.....		29
41	5.1 Quantitative Data.....		29
42	5.2 Qualitative Data.....		30
43	5.3 Radionuclides of Concern		30
44	5.4 Locations of Concern		30
45	5.5 Impacted Media.....		31

46	5.6	Developing the PSM.....	32
47	6	Scoping Phase	35
48	6.1	Sampling	35
49	6.2	Updating the PSM	35
50	7	Characterization Phase	37
51	7.1	Early Characterization	38
52	7.2	Late Characterization	38
53	7.3	Remedial Design Support	39
54	8	Remediation	40
55	8.1	Updating Remedial Actions in the GSM	41
56	8.2	Open Pit Excavation and the DUWP-ISG-02 (USNRC 2023)	42
57	9	Final Status Survey.....	42
58	9.1	Final Status Survey Readiness Check	43
59	9.2	Hypothesis Testing	44
60	9.2.1	Overview	44
61	9.2.2	Sample Design	44
62	9.3	Virtual Scanning	46
63	9.3.1	Overview	46
64	9.3.2	Sample Design	48
65	9.3.3	Evaluation.....	48
66	9.3.4	Workflow.....	48
67	9.4	Result of Practical Equivalency Stability Assessment	48
68	9.4.1	Overview	48
69	9.4.2	Correlation Model Stability	49
70	9.4.3	Geostatistical Model FSS Predictions	49
71	9.4.4	Geostatistical Model Differencing.....	49
72	9.4.5	Decision Stability	49
73	10	AI and Large Language Models.....	51
74	11	SUMMARY	54
75	12	REFERENCES	57
76	13	Appendix: SAMPLE DESIGN STRATEGIES.....	67
77	13.1	Adaptive Fill.....	67
78	13.2	Area of Concern Boundary Design.....	67
79	13.3	Bayesian Ellipgrid.....	68
80	13.4	Check and Cover Sample Design	69
81	13.5	Final Status Survey Design	70
82	13.6	Grid Designs.....	70
83	13.7	High-Value Designs	71
84	13.8	Judgmental Design.....	72
85	13.9	MrsDM 72	
86	13.10	Random Designs.....	74
87	13.11	Threshold Radial	74
88	13.12	Structural Sampling	74
89	13.13	Variance Designs	75
90			

LIST OF FIGURES

92	Figure 2.1. A simple CSM drawn in graphical software (USEPA 2008, p. 1).....	7
93	Figure 2.2. A variety of information is provided in CSM, including quantitatively derived	
94	subsurface conditions (USEPA 2008, p. 2).	7
95	Figure 3.1. Cross sections of some example SSU configurations.	16
96	Figure 3.2. Example configurations for excavated subsurface survey units.	16
97	Figure 3.3. Exposure mechanisms anticipated in this guidance.	17
98	Figure 3.4. Example expert judgement legend.	19
99	Figure 3.5. PSM for cesium-137 contamination (Stewart 2011).	20
100	Figure 3.6. Three example subsurface simulations from the geostatistical simulation	
101	model.....	21
102	Figure 3.7. Map of expected activity levels developed by post-processing simulations.	22
103	Figure 3.8. GSM is capable of producing multiple renderings (average map, local	
104	variance, virtual scans, hypothesis readiness testing) as needed.....	24
105	Figure 3.9. 2D visualization of included and excluded (masked) cells in the foundation	
106	grid.	25
107	Figure 3.10. GEM+ scan of two local DCGLv specifications through 50 simulations.	25
108	Figure 3.11. GEM+ unity scan rule.	26
109	Figure 3.12. A single $3 \times 3 \times 1$ geometry scan point (red box).	27
110	Figure 4.1. GEM+ workflow.....	28
111	Figure 5.1. Initial PSA based on HSA.	33
112	Figure 5.2. Spatially separable concerns in a PSM.	34
113	Figure 6.1. Progression of updating an HSA to a PSA.	36
114	Figure 7.1. GEM+ in early-to-late characterization phase.....	37
115	Figure 7.2. Cesium Site MrDM from Stewart (2011).	40
116	Figure 8.1. Frames of reference for geostatistical modeling and compliance under a	
117	backfill scenario.....	41
118	Figure 8.2. GSM remedial update workflow.	41
119	Figure 9.1. An example graded core hole design for subsurface DCGLv _w compliance.....	45
120	Figure 9.2. An example of bottom-up core hole design where samples may increase and	
121	decrease per layer.....	45
122	Figure 9.3. Map of two different DCGLv _{local} scan criteria over the same place.	47
123	Figure 10.1. Screenshot of ChatGPT prompt.	51
124	Figure 10.2. ChatGPT graphic visualization of best-fit spherical model.	53
125	Figure 10.3. Approximation of ChatGPT 5.0 omnidirectional correlation model	
126	parameters in SADA 5.0.....	54

127	Figure 13.1. Adaptive fill sample design.	67
128	Figure 13.2. AOC boundary design showing a ring of new samples covering the	
129	boundary between contaminated (gray) and non-contaminated zones.....	68
130	Figure 13.3. Bayesian Ellipgrid samples in tighter patterns over areas where the	
131	contamination is more likely to be found.	69
132	Figure 13.4. Check and cover sample design sample configurations under different	
133	confidence settings (from Stewart et al. 2011).	70
134	Figure 13.5. Gridded designs.....	71
135	Figure 13.6. High-value design places samples in areas with the highest modeled	
136	values.	72
137	Figure 13.7. MrsDM evaluation of proposed locations.....	73
138	Figure 13.8. A 2D random sample design.	74
139	Figure 13.9. Three example geometric patterns for bounding an isolated hot-spot	
140	measurement.	74
141	Figure 13.10. Variance-based sample design shows new samples (gray triangles) in	
142	areas of highest model variance (e.g., kriging variance or simulation point	
143	variance).....	75
144		

145

146
147
148
149
150
151
152

LIST OF TABLES

Table 7.1. Early-stage design options..... 38

Table 7.2. Later stage design options 39

Table 9.1. Virtual scan results based on geostatistical simulation..... 47

Table 9.2. Example tabular virtual scan results. **Error! Bookmark not defined.**

Table 10.1. ChatGPT model selection and parameterization. 52

Table 10.2. ChatGPT lag distance and semivariance backing information. 52

153

ABBREVIATIONS

154	AOC	area of concern
155	CCM	contaminant concern map
156	CSM	conceptual site model
157	DCGL	derived concentration guideline level
158	DCGLv	derived concentration guideline level for subsurface (volume)
159	DQO	data quality objective
160	EMC	elevated measurement comparison
161	EPA	US Environmental Protection Agency
162	FSS	final status survey
163	GEM	Geospatial Extension to MARSSIM
164	GSM	geostatistical site model
165	GUI	graphical user interface
166	HSA	historical site assessment
167	MARSSIM	Multi-Agency Radiological Site Survey Investigation Manual
168	MrDM	Multi-scale Remedial Design Model
169	MrsDM	Multi-scale Remedial Sample Design Model
170	NRC	US Nuclear Regulatory Commission
171	ROC	radionuclide of concern
172	ROPE	results of practical equivalence
173	PRV	post-remedial replacement value
174	PSM	preliminary site model
175	SADA	Spatial Analysis and Decision Assistance
176	SC&A	S. Cohen & Associates
177	SIS	sequential indicator simulation
178	w	wide-area
179		

180

181 *Many complex decommissioning sites are expected to contain areas of residual radioactivity in*
182 *subsurface soils and building structures. These complex sites will need to be*
183 *investigated/surveyed and a determination made as to the need for subsurface remediation and*
184 *a final decision made on license termination. — Tom Aird, US Nuclear Regulatory Commission.*
185

186

1 INTRODUCTION

188 In January 2012, the US Nuclear Regulatory Commission (NRC) released NUREG-7021,
189 articulating focused efforts to address radiological subsurface compliance. The work integrated
190 state-of-the-art (as of 2009) geostatistical modeling and decision support frameworks with the
191 organizing principles that had previously shaped the Multi-Agency Radiation Survey and Site
192 Investigation Manual (MARSSIM) guidance for radiological surface compliance in the vadose
193 zone. Since 2009, significant advances have been made in decision support, geospatial
194 modeling, and computing. This revision to NUREG-7021 updates the previous document by
195 further strengthening and tightening the subsurface approach under a workflow called GEM+.

196 For more than two decades, MARSSIM (USNRC 2000) has been a widely used and rigorously
197 tested statistical approach used to support compliance decision-making for radiologically
198 contaminated surfaces (surface soil or building surfaces). The details are complicated, but
199 MARSSIM has two primary goals. First, the average concentration of any single radionuclide
200 may not exceed the *wide-area* derived concentration guideline level (DCGL_w). Second, within
201 this area, no smaller area may have concentrations that exceed the derived concentration
202 guideline level elevated measurement comparison (DCGL_{EMC}) value. Additionally, total activity
203 over multiple radionuclides may not exceed a cumulatively derived limit. MARSSIM works by
204 dividing a site into unimpacted and impacted areas. Unimpacted areas require no further survey
205 activities. Impacted areas are further divided into smaller exposure units called survey units,
206 classifying those survey units according to how likely they are to be contaminated, determining
207 the size of the survey units based on risk, and providing guidance on the number of samples
208 and percent scanning needed based on the risk and other considerations. Of course, this
209 description is a cursory summary of a carefully crafted and highly detailed compliance path
210 found in the guidance. MARSSIM's success can be attributed to the use of tractable approaches
211 facilitated by the ease of access to contaminated media, standard statistical methods, and
212 inexpensive scanning technologies that check between sample locations.

213 In the narrowest sense, the MARSSIM technical compliance machinery is not easily moveable
214 to the subsurface because of issues such as the cost of sampling, inability to perform
215 exhaustive scanning, complications associated with volumetric exposure units, increased media
216 complexity associated with soil properties, and general inaccessibility. However, advances in
217 geospatial modeling and decision support since MARSSIM allow translation of its guiding
218 principles to subsurface problems. In geospatial language, these principles recognize that
219 (1) exposure can occur over multiple spatial scales (wide area vs. local elevation), (2) multiple
220 clean-up levels are possible according to the spatial scale and subsurface depth selected, and
221 (3) compliance should be based on the totality of evidence and linked to interpretable statistical
222 endpoints. However, executing on these principles for the subsurface requires advanced spatial
223 methods and a graded approach that balances the intent for high statistical certainty against the
224 subsurface sampling densities required to meet that certainty in three dimensions (subsurface)
225 instead of two (surface).

226 The first edition NUREG-7021 was early in articulating the application of advanced geospatial
227 and decision methods in a MARSSIM-inspired subsurface workflow for subsurface compliance
228 and provided some clarity in the details. It maximized the informatic value of what was known,
229 leveraged the machinery of spatial autocorrelation, replaced crisp binary exhaustive screens
230 with probabilistic modeled values, and conducted hypothesis testing under the assumption of
231 independent and identically distributed data with modeled confidence. The heart of the idea was
232 to use the contaminant concern map (CCM) as the common operating picture to quantitatively

233 compile the best state of knowledge possible. The CCM drives sampling and remedial choices
234 that could lead to a defensible and transparent compliance decision. Using the phases of
235 investigation as the backbone, the approach annotated each with a geospatial strategy for
236 advancing the CCM, offering quantitative continuity across the stages and accumulating a body
237 of evidence for or against compliance. This work was an important first step but left open
238 several key issues. What is the best way to process multiple subsurface exposure scenarios
239 and multiple derived concentration guideline levels for subsurface (DCGLvs)? What is the best
240 way to optimally remediate and sample, considering different exposures and their corresponding
241 exposure unit sizes/geometries? How should a formal final status survey (FSS) be developed
242 and interpreted?

243 This report advances NUREG-7021 by leveraging major principles of the Geospatial
244 Extension to MARSSIM (GEM) framework (Stewart 2011), advancing them further with the
245 current state of the art and evolving a more comprehensive subsurface approach called GEM+.
246 The GEM+ workflow has the following important improvements over the original NUREG-7021
247 guidance.

- 248 • A formal FSS hybrid design based on statistically selected soil samples and
249 geostatistical *virtual scans* of the subsurface.
- 250 • Concurrent consideration of multiple DCGLvs, spatial scales, and depths
- 251 • Tighter articulation of spatial modeling across the phases of environmental investigation
- 252 • Complimentary support to the open pit excavation scenario under DUWP-ISG-02.
- 253 • Tools for designing multiscale remedial geometries and corresponding sample designs

254 This revision to NUREG-7021 does not propose to address all questions or mitigate all known
255 challenges. Rather, it aims to provide a clearer picture from a geospatial decision support
256 standpoint while leaving significant latitude for adaptation and modification. Any
257 recommendations here do not override site-specific conditions and the needs of the data quality
258 objective (DQO) process.

259 GEM is designed for complex subsurface scenarios. Small scale scenarios where the amount or
260 spread of contamination can conservatively be addressed with a conservative total removal
261 strategy may not find this guidance helpful. Examples would be a small volume of isolated and
262 elevated activity that is economically better to simply extract the entire volume, including a
263 considerable buffer of clean soil around it, than engage in a full characterization and remedial
264 design approach. Complex scenarios involve substantially distributed or complex configurations
265 of contamination in the subsurface, soil that has elevated levels but licensees wish to leave in
266 situ, issues of groundwater, and so forth. In this case GEM+ is appropriate.

267

2 BACKGROUND

268 2.1 Overview

269 Evolution of subsurface decommissioning guidance beginning with foundational DQO guidance
270 to the present-day update to NUREG-7021 has been a decades-long effort and originates from
271 a much wider set of regulatory policies and guidance. In the widest context, federal guidance
272 documents provide and interpret environmental regulation for federal agencies and the public
273 (USOMB 2007). These documents often translate policy within a scientific context, promoting
274 responsible and consistent methods for responding to environmental pollution. The DQO
275 process is a cornerstone of regulatory guidance for investigating contaminated lands. The
276 process provides guidance on setting project objectives, specifying decision errors, and
277 identifying information needs, including type, quantity, and quality of data (USEPA 2006a). First
278 appearing in the 1980s (USEPA 1980), the DQO process has motivated several follow-up
279 guidance documents (e.g., USEPA 1987a, 1987b, 1994b, 1989a, 1989b, 1992a, 1992b, 1994a,
280 1997, 2000a, 2001c, 2002a, 2002b, 2002c, 2003b, 2006b, 2006c, and 2006d; USNRC 2000)
281 and has shaped the landscape of environmental investigations for the last 40 years. Here we
282 provide an overview of recent relevant works specific to the update of NUREG-7021. The
283 account starts with the MARSSIM guidance and briefly traces key documents from that point
284 forward in chronological order.

285 2.2 Multi-Agency Radiation Survey and Site Investigation Manual

286 MARSSIM provides guidance for determining whether a site is in compliance with a radiation
287 dose or risk-based value (USNRC 2000). Specifically, the guidance is focused on contamination
288 at the surface, either in the surface soil layer or on building surfaces. Compliance assessment is
289 driven by three key concerns: (1) determining whether the site-wide average exceeds a $DCGL_w$,
290 (2) determining whether any localized elevated areas exceed an elevated area $DCGL_{EMC}$, and
291 (3) whether total activity across multiple radionuclides exceeds a cumulative activity. The first
292 question is addressed using hypothesis tests and sampling strategies to determine the number
293 and location of samples that show that the mean (or median) concentrations in the survey unit
294 are less than the DCGLs. Scan instrumentation is used to detect local elevations between
295 samples that could also contribute to dose. A couple of approaches such as the Unity Rule and
296 Gross Activity DCGL handle the case of multiple radionuclides (see for example, p. 4-7, 4-8 in
297 MARSSIM). Although the details can be tedious, these approaches are tractable, facilitated by
298 the ease of access to contaminated media, standard statistical methods, and inexpensive
299 scanning technologies that check for residual activity between sample locations.

300 In the MARSSIM framework, impacted portions of a site are divided into homogenous
301 geographical areas called *survey units*. A survey unit is a geographical area with a specific size
302 and shape that serves as the basis of the investigation and over which compliance with a
303 release decision will be made. Three survey unit classifications are possible for impacted areas
304 of the site:

305 Class 1 areas: Areas that have, or had prior to remediation, a potential for radioactive
306 contamination (based on site operating history) or known contamination (based on previous
307 radiological surveys).

308 Class 2 areas: These areas have, or had prior to remediation, a potential for radioactive
309 contamination or known contamination but are not expected to exceed the DCGL.

310 Class 3 areas: Any impacted areas that are not expected to contain any residual radioactivity or
311 are expected to contain levels of residual radioactivity at a small fraction of the DCGL.

312 The selected path of investigation within MARSSIM depends on this classification. In general,
313 MARSSIM is motivated by two criteria. The first criterion is referred to as the Derived
314 Concentration Guideline for site-wide activity ($DCGL_w$). The $DCGL_w$ is the release criterion by
315 which estimates of site-wide activity are compared. The second criterion is referred to as the
316 Derived Concentration Guideline for elevated measurement comparisons ($DCGL_{EMC}$). This value
317 pertains to the upper limit permitted by localized areas of elevated activity. An analyst must
318 check to see the sum of fractions of all of the elevated areas and the wide area meet the dose
319 limit. Licensees may also need to show that multiple contaminated media meet the dose.
320 MARSSIM is therefore concerned with the survey-wide activity level and localized activities
321 within the unit. If a survey unit fails these checks, then additional steps may be taken before
322 release.

323 For the survey unit-wide comparison, simple statistical tests are used that assume
324 independence in the observations. Typically, nonparametric tests, such as the sign test or
325 Wilcoxon rank-sum (WRS) test, are used to test the hypothesis that the survey unit-wide
326 average is less than the $DCGL_w$. The local activity levels are evaluated by comparing scanning
327 results and/or sample measurements directly against the $DCGL_{EMC}$. Any result that exceeds the
328 $DCGL_{EMC}$ requires additional investigation or remediation before the survey unit can be
329 released.

330 Knowledge of the forthcoming assessment permits the prior estimation of the number of
331 samples required under the assumption of spatial independence. The number of samples will be
332 statistically sufficient to achieve certain Type I and Type II error rates given the $DCGL_w$, certain
333 assumptions about sample variance, and a parameter called the lower bound of the gray region.
334 Estimates for the number of samples may be affected by background samples as well if the
335 radionuclide is naturally occurring. Either way, the samples must also have a sufficient density
336 in the survey unit to ensure that any potential elevated areas will be discovered with high
337 probability (Class I and II units). A measurement is elevated if it exceeds the $DCGL_{EMC}$. This
338 comparison may be conducted not only against direct measures but also against radiological
339 scan data.

340 MARSSIM therefore has the following important properties (succinctly described):

- 341 • A historical site assessment (HSA) is conducted. This determines the likelihood of
342 contamination often in a very qualitative way.
- 343 • If warranted, a scoping survey is performed to assess the severity and magnitude of the
344 assessment and possible remedial action.
- 345 • If warranted, a characterization effort is conducted to better delineate the extent of
346 contamination.
- 347 • If warranted, a remediation effort can be conducted to remove or mitigate activity levels
348 below each DCGL.
- 349 • The FSS is conducted.
- 350 • Survey units must pass a survey-wide activity limit ($DCGL_w$) and a local activity limit
351 ($DCGL_{EMC}$), both of which depend on an exposure scenario.
- 352 • Survey units pass the survey-wide limit by a formal statistical hypothesis test that
353 assumes the observations are independent.
- 354 • Survey units pass the $DCGL_{EMC}$ by direct measurement comparison and by
355 comprehensive scan data with assumed source geometry.

- 356 • Because of these well-formed decision outcomes, the number of samples required can
357 be estimated before the investigation.
- 358 • The assessment is conducted, and the site passes or fails.

359 The important facts about surface compliance can be summarized as follows:

- 360 • Assume no knowledge is available regarding contaminant location (other than the
361 designation of Class I, II, or III).
- 362 • The surface is easy to access.
- 363 • Samples are relatively inexpensive and easy to acquire.
- 364 • Scanning can provide a comprehensive “reality check.”

365 Subsurface contamination presents circumstances that do not warrant a direct application of
366 MARSSIM. First, different methods for sub-surface exposure can result in multiple DCGL_w
367 values including groundwater contamination scenarios, plant uptake routes, and scenarios
368 where the subsurface is disturbed and brought to the surface (e.g., excavation or well
369 penetration). DCGL_w evaluations may require consideration of multiple and depth specific
370 values. See NUREG-1757, Volume 2, Rev. 2, Appendix J for exposure scenarios that should be
371 considered.

372 Unfortunately, thoroughly checking for elevated areas is impossible using scan technologies.
373 Between sample locations, neither direct data measurements nor comprehensive scan
374 measurements can exhaustively measure activity, yet the demand for reasonable certainty
375 remains high. Scans down boreholes can be conducted, but, because of the physics of
376 radiation, a large amount of activity can be detected only at a very limited distance (few feet).
377 The number of boreholes would then need to be increased geometrically to meet MARSSIM
378 grade requirements. Still, the need for higher data collection persists, and this increased
379 demand will require additional access costs and intrusive measures (such as core holes).
380 Continued adherence to strictly high-quality lab results will likely be cost prohibitive as well.

381 **2.3 Triad**

382 In response to these circumstances, the US Environmental Protection Agency (EPA) has
383 articulated the *Triad* model (Crumbling 2001a). Triad represents a concerted effort by experts
384 from the public and private sectors to create a modern approach that lays the groundwork for a
385 second-generation DQO process (Crumbling 2002). Triad methodology spans the project life
386 cycle, providing continuity among management practices, scientific methods, and technological
387 advances that emphasize the quality of the decision. At the center of Triad is the *conceptual site*
388 *model* (CSM). A CSM is a representation of site knowledge that evolves throughout the
389 investigation. CSMs communicate knowledge about a variety of issues, including geology,
390 exposure pathways, spatial distribution of contamination, and transport mechanisms (Crumbling
391 2001a; USEPA 1992b). Under Triad, the CSM drives data collection by identifying knowledge
392 gaps. The CSM is reciprocally informed and evolved by the outcome of those data (Crumbling
393 2001a). Triad recognizes the value of accurate laboratory analysis but also calls for the
394 inclusion of screening and field detection methods that are typically faster and less expensive
395 (Crumbling 2004). The combination of speed and reduced costs can yield greater sampling
396 density and better support for CSM evolution.

397 Triad recognizes that two conflicting goals in an environmental investigation can often deter,
398 distract, and even cripple the decision process. On the one hand, environmental decision-
399 making should be based on sound science. On the other hand, projects are usually expected to

400 control costs and be as inexpensive as possible. Sound science implies that a sufficient amount
401 of quality data can be collected to fully understand the underlying processes. Given the complex
402 conditions that plague many sites, collecting these data usually requires numerous samples.
403 Depending on the measurement requirements and access costs, the price can be substantial.

404 This issue is further exacerbated by the common interpretation of the phrase “data quality.”
405 Historically, the focus has been on analytic quality, which in practice emphasizes the highest
406 possible accuracy for each measurement. This focus has been driven to some degree by
407 regulatory pressure on sample accuracy, evidenced by the rejection of screening and field
408 detection methods in many final decisions. Unfortunately, higher analytic accuracy requires
409 higher cost. As a result, project managers may necessarily limit the number of samples
410 collected (Crumbling 2002).

411 Triad approaches this problem by expanding the concept of data quality from an analytic
412 viewpoint to a decision-support viewpoint. Furthermore, emphasis is placed on the use of
413 alternative and real-time measurements along with alternative lines of evidence to inform
414 understanding and clarify uncertainty. This focus brings to the front the idea of sufficient data
415 accuracy and the value of perfect information (Back et al. 2007; Dakins et al. 1996; Lyon et al.
416 1994; Kaplan 1993). The value of information and its accuracy must be questioned in light of
417 how well it supports the decision process rather than purely how accurate the value is. As a
418 trivial example, suppose a decision limit of 100 pCi/g is established. Method A, an expensive
419 sampling technology, can detect trace radioactivity as low as 0.1 pCi/g and measure it to
420 several significant digits. However, Method A does no better in supporting the decision than the
421 less expensive Method B, which can detect activities as low as 20 pCi/g to an accuracy of
422 ± 1 pCi/g. Both measurements are well below the criteria of 100 pCi/g. Therefore, overly
423 accurate sampling wastes valuable resources.

424 The problem is a double-edged sword because expensive sampling choices ultimately reduce
425 the total number of samples afforded. In light of decision quality, data quality shifts from a
426 question of accuracy to a question of sufficiency. With the gravitational center shifted from a
427 solely data view to a decision quality view, the value of understanding data quality is increased
428 rather than diminished. In fact, data quality is assigned a larger purpose, namely how well it
429 informs the conceptual site model. When more data are available, greater detail is possible in
430 the CSM. Therefore, within the context of this subsurface decision framework, further emphasis
431 is placed on sample location because environmental processes are always a spatial problem.

432 In a perfect world, “decision quality” would be equivalent to “decision correctness.” However,
433 decision correctness is often unknown (usually even unknowable) at the time a decision must
434 be made. In many cases, correctness may never be known owing to the situational complexity
435 and conditions that evolve over time. The term “decision quality” therefore means that decisions
436 are defensible against reasonable scientific or legal challenges (Crumbling 2002) given the best
437 available information and knowledge afforded by financial and professional resources at the
438 time.

439 The foundation of Triad is the CSM. A CSM is a representation of site knowledge that evolves
440 over the course of investigation. CSMs can take on a variety of forms. Some CSMs are simple
441 graphical depictions, as in Figure 2.1, or complex and quantitatively derived models, as in
442 Figure 2.2. Both figures are taken from USEPA (2008). Under Triad, the CSM drives data
443 collection by identifying knowledge gaps. The CSM is reciprocally informed and evolved by the
444 outcome of those data (Crumbling 2001a). The CSM ultimately informs the decision-making

process, and a focus on increasing the content and information value of the CSM should direct activities throughout the investigation life cycle.

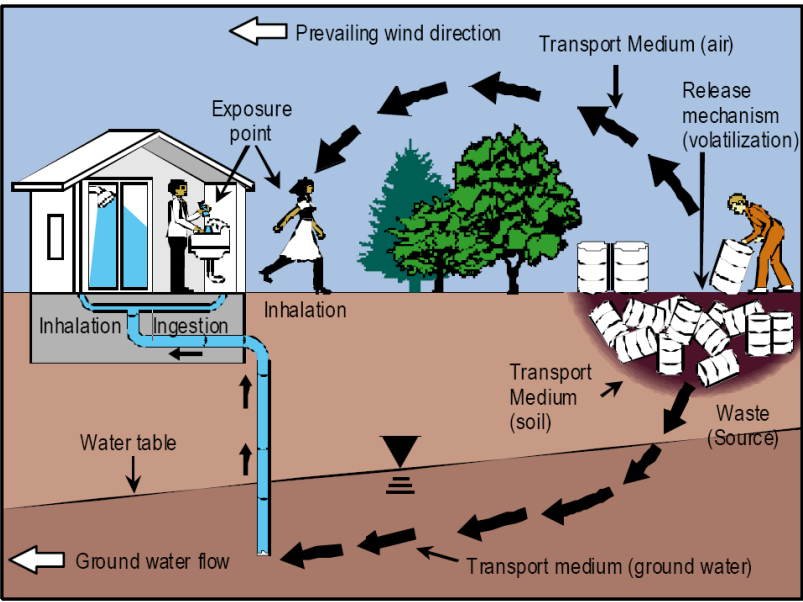


Figure 2.1. A simple CSM drawn in graphical software (USEPA 2008, p. 1).

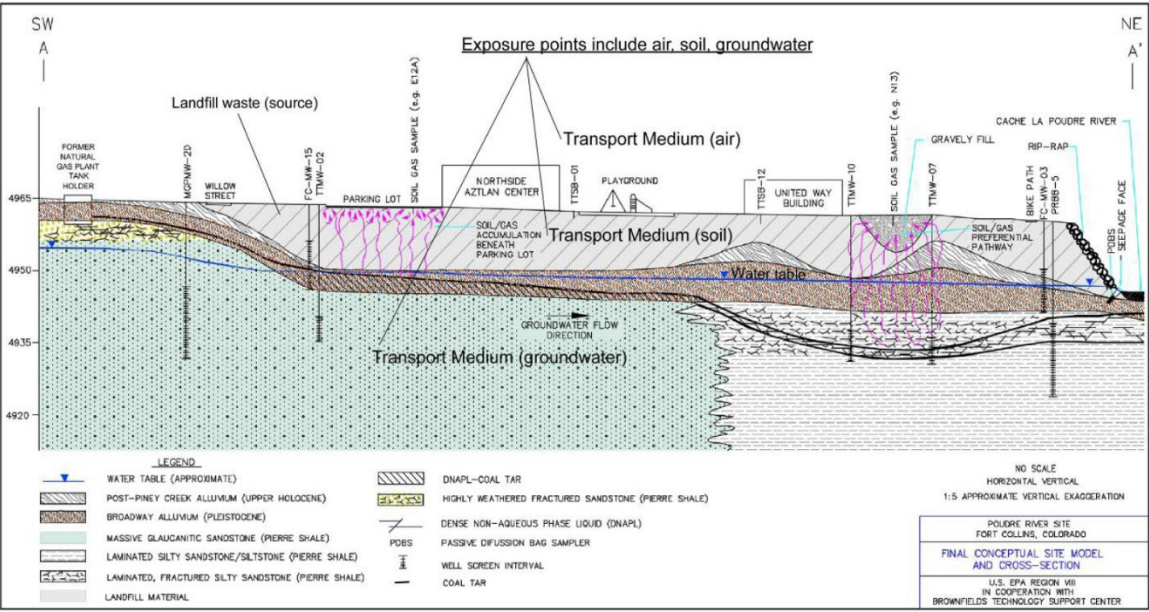


Figure 2.2. A variety of information is provided in CSM, including quantitatively derived subsurface conditions (USEPA 2008, p. 2).

The Triad approach comprises three core elements (Crumbling 2008): systematic project planning, dynamic work strategies, and real-time measurement technologies.

2.3.1 Systematic Project Planning

Systematic planning includes the following:

- Building an atmosphere of trust; transparent, open communication; and cooperation between parties working toward a protective, yet cost-effective resolution of the “problem.”
- Gaining consensus on the desired outcome for the project.
- Developing a CSM from existing information.

The CSM plays a central role in the Triad approach. The CSM represents the current state of knowledge, including processes, history, exposure, and sampling results. The idea is to continually mature and evolve the CSM as more information and understanding unfold. The CSM is the foundation of confident project decision-making and should be taken seriously. Many site managers do maintain an accurate and operational CSM that can inform the regulatory endgame, but, unfortunately, regulatory practice rarely develops this concept to its logical conclusion (Crumbling, 2008). This document describes one implementation of the CSM, called the geostatistical site model (GSM), which does bring this concept to a logical conclusion in the compliance endgame.

2.3.2 Dynamic Work Strategies

This managerial element encourages the use of dynamic, flexible planning processes that permit the use and evolution of the CSM in determining where resources should be spent. This element is referred to by Crumbling (2008) as process quality control and can yield tremendous project savings and improved decision outcomes.

2.3.3 Real-Time Measurement Technologies

This element embodies several themes. First is the use of cheaper, faster, sufficiently accurate data that can be collected in greater abundance and analyzed in a timely fashion to benefit CSM evolution. A greater abundance of data can lead to a more detailed and defensible CSM. Examples include in situ methods, geophysics, and rapid turnarounds from traditional labs. Second is the use of computer tools to capture, store, display, manipulate, and model measurements to support evolution of the CSM. This component is an important focus of this document, which brings the full power of geographic information systems (GISs) and spatial modeling systems to the center of the discussion, explicitly for the purpose of updating the CSM.

2.4 NUREG-7021 First Edition

The intent of this first edition was fully articulate how MARSSIM principles may be supported in the vadose zone. Several major challenges face this goal including accessibility of the subsurface, lack of comprehensive scans, and an increase in media complexity that calls for equally advanced scientific approaches. This report applied the principles of Triad to extend MARSSIM approaches into the subsurface and brought to bear a substantial and continually advancing set of tools from spatial analysis, modeling, and the GIS community. Specifically, the first edition offered recommendations on the flow and arrangement of components that form, update, and evolve an explicitly constructed, spatial variation of the conceptual site model, called the contamination concern map. This map is a specific implementation that focuses on the likelihood of exceeding a decision criterion at a local scale

and directly addresses uncertainty in volume extent and location. The map matures over each major phase of the investigation and serves as a basis for compliance reasoning by using a set of decision frameworks. The outcome can inform investigators and regulators alike of a reasonable course of action in the final assessment.

2.5 The GEM Framework

Stewart¹ (2011) picks up where the NUREG-7021 left off and moved the needle farther in his dissertation work titled “A Geospatial Based Decision Framework for Extending MARSSIM Regulatory Principles into the Subsurface.” That work explores simultaneously defining, sampling, remediating, and accomplishing compliance across multiple spatial scales and mechanisms of exposure). Subsurface direct exposure mechanisms operate on the idea that soils below ground may someday be excavated and brought to the surface. This is an organizing principle in Lively (2012) where computed DCGL values are based on surface DCGLs modified by volume and mixing factors that reflect physically lifting, mixing, and spreading soils above ground. Dig scenarios include well digging, trench building, or full excavation to build new facilities with basements or deep foundations. Note that Lively (2012) does not consider groundwater pathways which can drive DCGLs. GEM allowed for specification and compliance assessment simultaneously across multiple DCGL values. These values can reflect lift and spread scenarios that vary by volume and depth. These tasks can be completed by enhancing the Triad CSM first with a qualitative/quantitative Contaminant Concern Map and then into a fully quantitative CCM that would reach for the statistical clarity of standard hypothesis testing at the surface while admitting the severe complexity of the subsurface.

The GEM framework includes the following key goals:

- Simultaneously manage multiple DCGL values without undue burden on the investigation.
- Alleviate the risk of encountering the modifiable area unit problem by defining a fine spatial scale for analysis that can be aggregated to any exposure scale/scenario of interest.
- A site fails compliance if the probability of exceeding a DCGL for any exposure unit of any size and shape, situated anywhere within the survey area (including the survey unit itself), exceeds an established probability limit. This is a model-based expansion of the traditional DCGLw test in MARSSIM.

Out of this dissertation work (Stewart 2011) came new valuable pieces of the subsurface puzzle, including the Regulatory Limit Rule which sets a compliance rule across multiple spatial scales and serves here as the GEM+ *virtual scan* step. The Multi-Scale Remedial Design (MrDM) and Multi-Scale Remedial Sample Design (MrsDM) algorithms process fine-scale geostatistical simulations derived from hard and soft data into optimal soil removal plans and sampling intended to maximize the next iteration of remedial planning. That work also formally published the Check and Cover sampling strategy, which had not yet appeared because of the substantially delayed publication of NUREG-7021. The significance of this work is that it laid a solid foundation from HSA to compliance check that addressed the problem of multiscale compliance and provided rigorous methods for optimally reaching that multi-objective goal.

¹ The same author as the preceding NUREG-7021 manuscript.

2.6 Electric Power Research Institute Geostatistics Review

In 2016, the Electric Power Research Institute (EPRI) published *Guidance for Using Geostatistics in Developing a Site Final Status Survey Program for Plant Decommissioning* (EPRI 2016). This foundational update provides a comprehensive view of the state of the art in geostatistical methods, software, and industrial applications. The work emphasizes the benefits of geostatistical modeling to the decommissioning process and offers specific guidance for using geostatistical analysis during nuclear power plant decommissioning. It begins with an accessible overview of the principles of geostatistics followed by industrial case studies, including decommissioning sites. Also included in the work is a comprehensive survey of geostatistical modeling software packages and in-depth exploration of their capabilities and limitations. The report offers solid guidance for conducting a geostatistical investigation that is conveniently articulated in the language of decommissioning and the phases of investigation. This work's guidance is well aligned with NUREG-7021 perspectives and provides an invaluable resource for investigators engaging in this present update.

2.7 NRC Subsurface Workshops

To advance the progress of NUREG-7021 further, NRC held two public workshops in July 2021 and May of 2022. These workshops convened leading experts in the field who brought forward and discussed a number of issues regarding subsurface contamination and summaries of these can be found in USNRC (2021) and USNRC (2022a). A number of presentations were given covering CSMs, DCGLs, EMCs, and Sample Designs. we summarize the outcomes here from a geospatial vantage.

Phases of Investigation

At the first workshop, S. Cohen & Associates (SC&A) emphasized the importance of CSMs in driving sampling and modeling decisions throughout the assessment life cycle. In particular, the evolution of a CCM (Stewart 2011) posed as numerical, geospatial counterpart to a CSM that evolves empirically by using sampling and geostatistics as the site investigation moves from HSA to final test for compliance (USNRC 2021). In the Radiation Safety and Controls presentation these concepts were amplified further by a common data environment that includes more information such as imagery, 3D rendering, and blueprints. conveyed implemented through the use of Geospatial Information Systems (GIS) (USNRC 2021). In the second workshop (USNRC 2022b), SC&A further proposed that geostatistics (e.g., kriging and prior information (Bayes) be used to make better use of data. After a release criteria is decided, an initial survey design is created by applying Bayesian Ellipgrid to an HSA-based map of exceedance probabilities. Secondary sampling designs based on geospatial methods such as Markov-Bayes can further refine data in subsequent steps until the FSS is conducted. In fact, SC&A assessed that Bayesian Ellipgrid and Markov-Bayes are the two most promising approaches to sampling design in the subsurface. In the second workshop, PNNL attendees provided an overview of geophysical surveys and the wide range of information they provide that can be very useful in forming CSM, CCMs, and drive sampling strategies. These kinds of data can also be ingested by geostatistical models as auxiliary data to improve estimation and sample design strategies. The take home message here is that continually updated geospatial maps of data and modeling using the best information available is key to a successful and graded compliance approach to complex subsurface contamination.

2.7.1 Subsurface DCGLs

Given the number of possible exposure scenarios and applicability by depth, the use of multiple subsurface (DCGL_v) values was proposed. Example scenarios include construction of buildings, lift and spread excavation, intruder, and groundwater. DCGL_v values will depend on whether contamination is found in the unsaturated or saturated zones. Argonne provided insight that DCGL_v values computed in RESRAD can vary as a function of thickness, area, and cover. Existing guidance on computing subsurface DCGLs was noted in USNRC (2022b) and USNRC NUREG/CR-7268. NRC provided a presentation on the challenges of dealing with multiple DCGLs when developing sample designs and analysis. In the second workshop, NRC's overview continued to emphasize the need for DCGL_vs and DCGL_{vEMC}s for subsurface residual activities (USNRC 2022b). Geospatial modeling is relevant here as it can estimate both the vertical and horizontal spread of contamination indicating which subsurface zones or scenarios are important to consider. Furthermore, the dissertation (Stewart 2011) provided algorithms for simultaneously considering multiple DCGLs in characterization and remedial design.

Presentations also touched on questions of DCGL_v spatial scale. On the one hand SC&A posed a DCGL_v analogous to a DCGL_w and what test would be analogous to the MARSSIM type hypothesis testing. Additionally, questions around whether a DCGL_{vEMC} should be computed were raised and under what scenarios that exposures might occur. NRC slides noted that intrusion scenarios (as in well placement) were a case where subsurface DCGL_{vEMC} would apply. For open pits, it was noted that you could scan for elevated areas as indicated by MARSSIM). But this leaves open the question of contamination left in place. In this present report, we assess that while it is possible to develop frequentist statistical approaches with fixed scale subsurface support (wide vs local) more ground can be gained by repositioning the test as a model evaluation where geostatistical simulations estimate probability of exceeding any DCGL_v of interest regardless of the spatial support of that compliance value. In the case of local elevations, it is generally true that geostatistical models tend to produce estimates that are at or below the highest value encountered through sampling. However, both EPRI and Gogalak emphasized that EMC detections could be geostatistically addressed by evaluating higher percentiles of the local distribution against a DCGL_{vEMC}.

2.7.2 Final Status Surveys

Many open questions remain in how to address the FSS and decision to release the site. Questions around whether Scenario A (fails release unless proven otherwise) or Scenario B (passes release unless proven otherwise) are applicable. Questions about whether a subsurface analogy to the power of the hypothesis test vs sample size should be developed and whether this be accomplished as a geostatistical post processing step. These lead to final status questions about sample density, depth, and distribution to achieve a level of confidence and limit decision errors. One attendee assessed that MARSSIM tests (Wilcoxon Rank, Sign) would still apply for subsurface class 3 where little, or no, potential for radionuclide concentration(s) or radioactivity above background exists. These questions continue to be addressed in this present report which proposes a graded workflow for demonstrating compliance using geostatistical uncertainty as an analogous model to the hypothesis test well understood under MARSSIM. However, as EPRI correctly points out in the workshop, extending to the subsurface increases the number of dimensions resulting in sparser data sets. This is compounded by inaccessibility, heterogeneity, and complex processes in the subsurface. While geostatistical modeling can help alleviate some of these challenges, it is reasonable to admit that a framework is required that makes compliance possible in the face of substantial spatial uncertainty.

Geospatial Methods

In the first workshop, presentations favorably reported on present state of the art and example applications of geospatial methods. EPRI presentation reported relevant ISO standards (covered below) and a positive report on geostatistical methods used for decommissioning nuclear plants, laboratories and research facilities in France, Spain, and Belgium, among other countries, leading to tangible cost savings. EPRI provided a roadmap for applying geostatistics that could support modeled compliance estimation at multiple spatial scales, a principle that this present report emphasizes as well. PNNL provided an overview of Visual Sample Plan and SC&A provide an overview of the differences, pros, and cons of SADA vs VSP. In the second workshop, PNNL provided a deeper examination of how VSP could be used in a subsurface decommissioning project. In the second workshop, NRC noted a need still exists for evaluation of geospatial data analysis and modeling tools including GIS and geostatistics. Readers are pointed to workshop materials at <https://www.nrc.gov/docs/ML2211/ML22117A070.html>.

2.8 NUREG 1757 Vol 2, Revision 2

In July 2022, NRC published a second revision of NUREG-1757 Volume 2, which added substantial language regarding surveys and dose modeling considerations related to subsurface residual radioactivity. A summary of NUREG/CR-7021 and use of geographical information system and geospatial tools was included. Among other topics, surveys of open excavations and reuse of materials were discussed in Appendix G.3. The technical basis for the approaches discussed in Appendix G.3 were expanded upon in DUWP-ISG-02 (USNRC 2023), which presents detailed guidance extending MARSSIM to open surfaces in the subsurface including basement substructures (see discussion below). An update to Appendix J of NUREG-1757, Volume 2, was also made in Rev. 2 which provided additional guidance on consideration of exposure scenarios for buried or subsurface residual radioactivity. This guidance update was based on lessons learned from review of license termination plans (LTPs) for reactor licensees. Integration of dose modeling with FSSs, including use of multiple DCGLs for various environmental media including surface and subsurface soils was also a topic in Appendix G.

2.9 S. Cohen & Associates White Paper

In September 2022, SC&A wrote a white paper on that assessed NUREG-7021 and other resources in light of progress in this subject area. The paper reiterates the continued challenges in subsurface surveys, including lack of clear mechanisms, inaccessibility of the subsurface, lack of comprehensive scans, and increased media complexity. It cautions of likely increases in the number of complex decommissioning sites that are expected to become active soon, many with likely subsurface residual radioactivity. The authors call for guidance that emphasizes sound geospatial and rigorous decision-making methods that recognize the inherent limitations associated with subsurface investigations. Additionally, the paper summarizes industry-accepted practices, references for NRC-proposed activities, and examines several subsurface scenarios, including sampling and survey strategies. A key observation is that the Bayesian Ellipgrid and the Markov–Bayes approaches appear to be two highly promising approaches for survey design. At nearly 400 pages, it covers a wide range of subsurface scenarios, including vadose zone contamination. In the main, the propositions regarding subsurface soils are well aligned with the goals of this report. Consequently, the paper has been beneficial in updating this present report.

2.10 Subsurface and Visual Sample Plan

In November 2022, Pacific Northwest National Laboratory produced a review document (Huckett et al. 2022) looking at possible geostatistical approaches for subsurface scenarios, situating their Visual Sample Plan (VSP) package and other laboratory capabilities among these possibilities and pointing to recent developments in the literature. The report reinforces many of the challenges detailed in other reports and proposes a way forward as well as open research questions that must be answered. For example, they propose decomposing the 3D problem into a series of separately modeled disjoint 2D problems that may be approached with MARSSIM statistical machinery as usual. They recognize the potential for complex volumes for which such layering is not easily performed. The article provides instructional background on geostatistical principles, points to recent publications on the subject, and prioritizes how VSP should be updated to accommodate 3D modeling. The report concludes by emphasizing the need for further research to adequately support FSSs and points to the importance of model-based inference rather than design-based inference as a viable path forward. The present report is a significant examination of model-based inference as a way to balance the need for decommissioning confidence against the high cost of gaining this confidence in traditional or design-based ways.

2.11 Interoffice Staff Memorandum

In October 2023, NRC published “Division of Decommissioning, Uranium Recovery, and Waste Programs [DUWP] Interim Staff Guidance [ISG] DUWP-ISG-02” (USNRC 2023) for public comment. This document provides interim guidance for survey of open surfaces in the subsurface including open excavations, basement substructures planned for backfill, and the backfill materials themselves. The guidance extends MARSSIM principals to surfaces located in the subsurface including survey-unit classification and size, number of samples, MARSSIM and alternative statistical tests and sampling methods. The memorandum discusses exposure scenarios pertinent for subsurface, importance of source geometry modeling in characterization, and relevance of multiple DCGLs that vary with potential exposure scenario and soil depth. The guidance also discusses consideration of dose from existing groundwater contamination as well as use of commonly used decommissioning dose modeling codes such as RESRAD-ONSITE for developing DCGLs for subsurface residual radioactivity. Methods to assess the importance of key parameters to the dose assessment as well as methods to obtain additional technical support for these parameters are provided. Finally, a number of lessons learned related to subsurface investigations are also provided. This present guidance compliments this memorandum and a section is dedicated to how GEM+ can interact with the recommendations therein.

2.12 Standards

ASTM provides several standards in geostatistics that are also useful to consider. ASTM D5922-18 is a standard guide for conducting a geostatistical site investigation including analysis, interpretation, and modeling of spatial variation (ASTM 2018a). Document D5923-18 provides standards for choosing a kriging method in site investigations (ASTM 2018b). Document D5924-18 provides guidance for the selection of simulation approaches in geostatistical site investigations (ASTM 2018c), and D5549-19. Finally, D5922-18 provides a standard guide for content of geostatistical site investigation report (ASTM 2019). These documents are very brief and offer a geostatistical subject matter expert a comprehensive checklist and reminder of issues to consider as you are moving forward. These assume a strong

714 knowledge of geostatistical modeling and simulation but are still useful in providing a roadmap
715 at any level of expertise.

716 The International Standards Office has produced ISO 18557_2017 which aligns very well with
717 U.S. Regulatory Guidance such as MARSSIM, NUREG-7021, NUREG-1757 and so forth. In
718 addition to mapping out best practices in general for radiological contaminated sites, they
719 provide an Appendix on Geostatistical data processing and examples of good practices. At 48
720 well written pages, the document is entirely accessible and quite recognizable to those familiar
721 with U.S. peer guidance documents. Appendix A provides a very short treatment of what
722 geostatistics is and how to arrive at appropriate sample sizes for geostatistics based on required
723 resolution in the final model and on real-world applications at significant sites (e.g., Fukushima).

724 **2.13 NUREG-7021 Rev 1 Approach**

725 The present revision aligns strengthens the recommendations of NUREG-7021 with follow-on
726 documents into a revised workflow that begins with an HSA and terminates in a hybrid FSS.

727 This revision emerges in an evolving regulatory environment where tension remains between
728 the need to handle complex situations responsibly and handle them simply. Complex situations
729 require complex approaches, and simplification of the problem in order to use simpler statistical
730 tests may be insufficient or even misleading regarding the nature of contamination and the
731 spatial uncertainty about it. For complex sites, stakeholders may need to manage costs by
732 relying on rigorously modeled spatial compliance supplemented by well-placed sampling
733 locations in the FSS. This revision is mindful of the complexity in modeled spatial compliance
734 and draws from significant efforts to make advanced geostatistical approaches more tractable to
735 reasonably trained geospatial analysts. The following are within the scope of this revision:

- 736 1. Retention of previous NUREG-7021 that remain applicable in light of progress since its
737 first release.
- 738 2. Leverage progress in Stewart's (2011) GEM model.
- 739 3. Responses to relevant recommendations in the public records described in this section,
740 including principles of multiple DCGLs created by multiple scenarios and/or multiple
741 depths varying DCGLs by depth, conservative single-DCGL approaches, open pit
742 excavation guidance, and approaches for survey design under varied conditions and
743 different phases of investigation.
- 744 4. Indication of recent developments in geostatistical modeling and/or decision support in
745 the peer-reviewed literature including the two NRC workshop results.

746 In this revision, the entire guidance approach is referred to as GEM+ to reflect a continued
747 repositioning of geospatial methods that indicate subsurface compliance. The acronym GEM+
748 refers to the integration of the GEM methodology with the previous approaches in the first
749 edition.

750

3 GEM+ Foundations

751 3.1 Overview

752 GEM+ is a hybrid strategy that applies standard hypothesis testing for compliance of survey unit
753 wide volumetric compliance ($DCGLV_w$) and virtual scan test for local volumetric activity limits
754 ($DCGLV_{local}$). GEM+ FSS was engineered to simultaneously accommodate multiple and depth
755 dependent values for both $DCGLV_w$ and $DCGLV_{local}$. In particular, survey unit compliance with
756 depth-specific (d), survey unit wide (w), radionuclide-specific (i), and volumetric (v) average
757 activity limits $DCGLV_{w,i}^{(d)}$ is evaluated by vertical layer using standard Sign and WRS hypothesis
758 tests. Survey unit compliance with depth-specific (d), localized (local), radionuclide-specific (i)
759 volumetric (v) activity limits $DCGLV_{local,i}^{(d)}$ is evaluated with a virtual scan of the Geostatistical
760 Site Model for soils that are to remain in place.

761 A survey unit demonstrates compliance when both $DCGLV_w$ hypothesis tests and $DCGLV_{local}$
762 scan tests pass with required statistical confidence, and the underlying geostatistical model is
763 demonstrably stable. The GEM+ workflow aims to develop a stable GSM over the phases of
764 environmental investigation consistent with the principles of the EPA TRIAD approach
765 (Crumbling 2001a). This chapter presents the foundations of GEM+, and the following chapter
766 discusses their use within the GEM+ workflow. The discussion continues with delineating
767 subsurface survey units.

768 3.2 Subsurface Survey Units

769 A Subsurface Survey Unit (SSU) is defined by its 3D study area boundaries of surface, floor,
770 and wall geometry (Figure 3.1). Adequate sampling, modeling, or other authoritative evidence to
771 conservatively estimate the floor depth and wall geometry of the survey unit volume to ensure
772 that elevated residual activity beyond the wall or below the floor is highly unlikely and arguably
773 are not significant (NUREG-1757 Volume 2, Revision 2, G-5). Initial location of walls and floor
774 are made during HSA and adjusted as necessary by new information gained during scoping,
775 characterization, and remedial action. MARSSIM and the DUWP-ISG-02 (USNRC 2023, p. 2–
776 10) emphasize that selection of boundaries (walls and floor) should be based on several factors
777 including area with similar operational history or potential for residual activity, exposure pathway
778 modeling assumptions and site-specific conditions, and should not have highly irregular
779 (gerrymandered) shapes unless the shape is appropriate for the site operational history or other
780 relevant conditions. Later on, geostatistical modeling could be used to adjust the position of
781 walls and floors if the phases of investigation reveal new information. Figure 3.1 shows a cross
782 section of three example survey volumes to demonstrate that wall and floor configurations are
783 adaptable to site specific circumstances. Bear in mind walls and floors are not required to be
784 physical walls and floors as revealed in an excavation scenario. They are flexible 3D boundaries
785 defining a defensible spatial scope in 2D and 3D of the FSS.

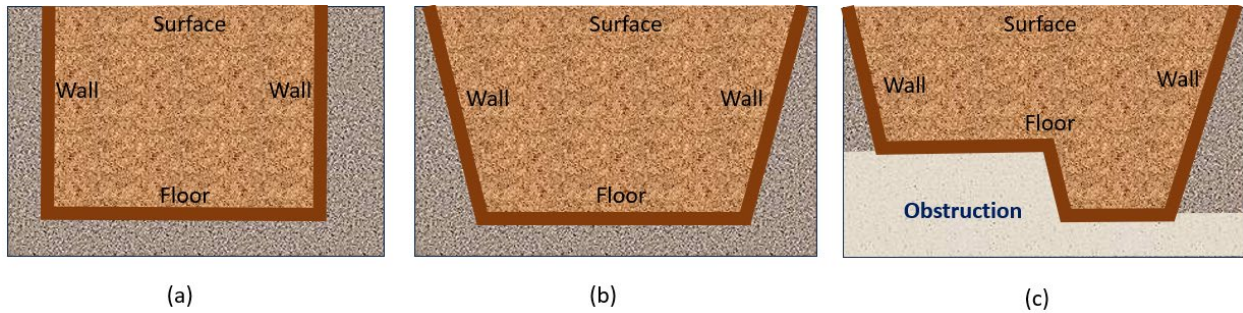


Figure 3.1. Cross sections of some example SSU configurations. (a) A shear wall and flat floor configuration, (b) wall lay back requirements for safety consideration, (c) wall and floor irregularity are possible with complex substructures.

Within the SSU, there may be excavation and backfill processes as well. Figure 3.2 shows example configurations including the location of walls, floor, original surface, excavation surfaces, and backfill. It is recommended that the floor and walls extend further than excavation in order to test soil activity levels just beyond the excavated surface.

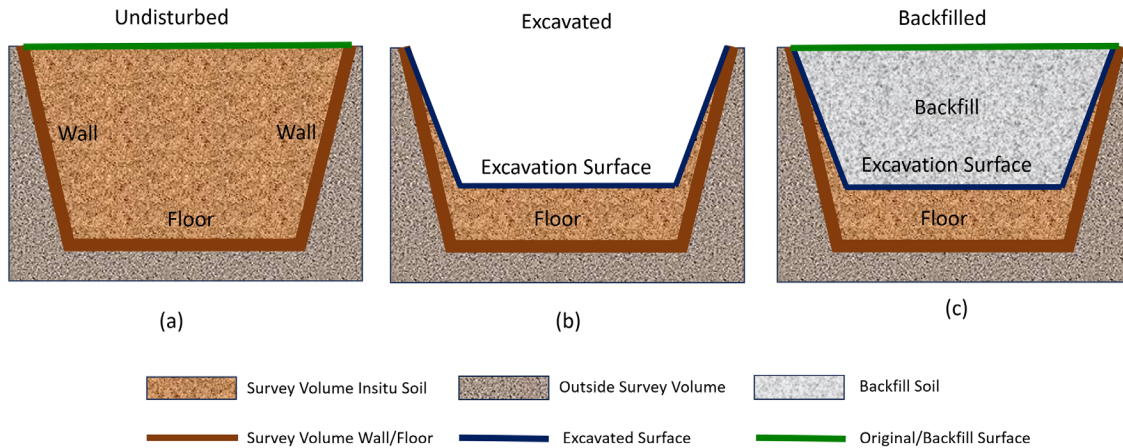


Figure 3.2. Example configurations for excavated subsurface survey units. (a) Undisturbed soils the license may wish to leave behind, (b) an excavation that will remain open, and (c) an excavation that will receive backfill.

3.3 Exposure Mechanisms

Multiple exposure scenarios may be simultaneously under consideration which underscore multiple and possibly depth varying DCGLvs (NUREG 1757 Vol 2 Rev 2, DUWP-ISG-02). These include (but are not limited to) shallow soil root uptake, excavation (now or in future), well digging, and groundwater contributions (Figure 3.3). This document does not provide guidance on how to compute these, but points readers to guidance issued under NUREG 1757 Vol 2 Rev 2 Appendix G and J as well as the DUWP-ISG-02 (USNRC 2023). In general, this guidance has been written in anticipation of broad subsurface scenarios included in those guidance documents.

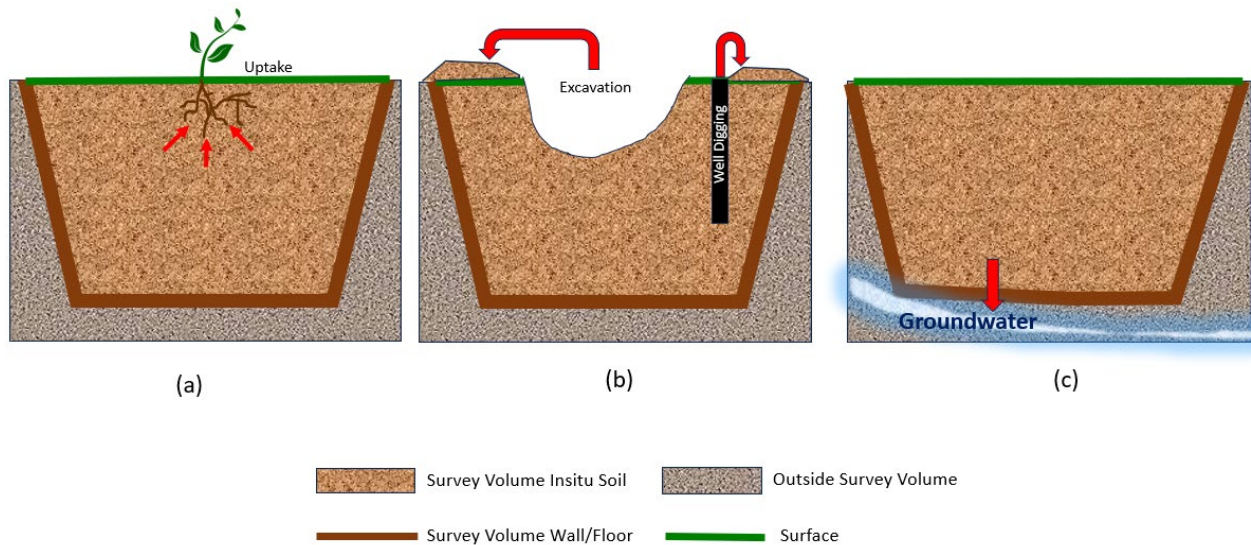


Figure 3.3. Exposure mechanisms anticipated in this guidance.

For survey unit wide $DCGLv_{w,i}^{(d)}$ values, the geometry is the entire horizontal extent plus the vertical layer depth. This is in keeping with the MARSSIM principle of limiting survey unit wide exposure but allows for variability by depth. For local $DCGLv_{local,i}^{(d)}$ values, a 3D scan geometry is required, and its determination will depend on the associated exposure scenario. In computational language, this geometry is simply the scan kernel geometry (e.g., cubic). The set of local DCGLv values can depend on any number of exposures, each which can vary by depth. Examples include:

- a) $DCGLv_{EMC,i}^{(d)}$: Exposure driven, local elevation limit for radionuclide of concern (ROC) i at depth d with specified local exposure scan geometry. The scenario of drilling with soil uplift and surface spread could fall into this category.
- b) $DCGLv_{NTE,i}^{(d)}$: Not to exceed (NTE) activity limit for ROC i at depth layer d with a given scan geometry. DUWP-ISG-02 (USNRC 2023, p. 2–5) indicates a licensee may establish a single cleanup level at a fraction of the applicable dose standard (e.g., unrestricted release standard) to account for multiple contaminated media or exposure pathways.
- c) $DCGLv_{GW,i}^{(d)}$: DCGLv related to groundwater (GW) exposure scenarios for ROC i at depth layer d with a given scan geometry.
- d) $DGCL_{ops}$: An operational DCGLv value conservatively selected to cover all radiological and depth specific DCGLvs.
- e) $scanv_i^{(d)}$: A scanning threshold of interest for ROC i at depth layer d with a given scan geometry.

GEM+ does not provide guidance on selection or method for calculating these. Rather, given this set of constraints, GEM+ simultaneously searches for likely exceedances of these using the Geostatistical Site Model and identifies the number and location where likely exceedances may occur. This is discussed in detail further down in the virtual scanning section. For multiple

834 radionuclides, the sum of fractions approach could be used or an operational $DGCL_{ops}$ might be
835 used to conservatively account for cumulative risk. Furthermore, virtual scanning assumes that
836 values have been selected after considering background levels.

837 **3.4 Subsurface Classification**

838 Class designations will remain the same here as in MARSSIM (USNRC, 2000).

839 **Class 1** applies to areas with the highest potential for contamination. Class 1 designations apply
840 to areas that have, or had prior to remediation, a potential for radioactive contamination or
841 known contamination. Areas containing contamination in excess of any $DCGLv_{w,i}^{(d)}$ for any depth
842 (d) or radionuclide (i) prior to remediation should be classified as Class 1.

843 **Class 2** areas have, or had prior to remediation, a potential for radioactive contamination or
844 known contamination, but are not expected to exceed any $DCGLv_{w,i}^{(d)}$ over any depth (d) or
845 radionuclide (i). There should be a high degree of confidence that no individual measurement
846 would exceed any $DCGLv_{w,i}^{(d)}$ to justify classification of an impacted area as a Class 2 versus a
847 Class 1 area.

848 **Class 3** areas are impacted areas that are not expected to contain any residual radioactivity or
849 are expected to contain levels of residual radioactivity that in all cases are a small fraction of the
850 $DCGLv_{w,i}^{(d)}$ based on the site operating history and previous radiological surveys. Class 3 areas
851 include areas where there is insufficient information to justify designation of the area as non-
852 impacted.

853 **3.5 Preliminary Site Model**

854 Under GEM+, if the HSA identifies the possibility of contamination, stakeholders are
855 encouraged to develop an initial preliminary site model (PSM) that describes where experts
856 believe contamination may exist. The PSM is a preliminary map that explicitly describes by
857 qualitative or quantitative means where contamination may be, and where further sampling
858 should be collected. The aim of the PSM is to kick start the formal mapping process that will
859 transition to the GSM when sufficient data is available.

860 In most cases, few or no subsurface contamination samples are available to engage a
861 geostatistical approach at first. In this case, judgement maps based on site process knowledge
862 and possibly limited data can be used to articulate and map concerns about possible
863 contamination locations. Eliciting stakeholders for the historical PSM falls under a broader
864 concept known as expert elicitation, which enjoys a lengthy and well-studied track record for
865 empowering workflows and decision processes with codified human judgement. A wealth of
866 materials describes how to engage stakeholders and pull the best unbiased judgement possible
867 from them. A light introduction to the broader subject of elicitation is available in Burgeman
868 (2016). Deeper information about human biases during elicitation and engagement with formal
869 statistical processes can be found (e.g., Kahneman 2011; Kahneman et al. 2021). Helpful
870 elicitation materials may be found in Gosling (2017), but the kind of formal elicitation found there
871 may not be warranted.

872 To facilitate consensus building, one or more meetings with stakeholders should be held to
873 collaboratively develop a PSM. This dialogue with stakeholders likely finds its greatest value in
874 surfacing differing opinions about site conditions, thereby facilitating an opportunity for

discussion, alignment, and collective PSM that captures group knowledge. During this step, any geographic tool may be used. The best tool may be a printed map laid out on table with pens for marking and writing. It could be a GIS software system that is edited on the spot. The map should contain any monitoring or past sampling results, previous characterizations that may exist, and any other relevant information. Figure 3.4 shows example map units that reflect expert judgment on the location of contamination. The units range from 1 to 5: 1 indicates high confidence that no problems will be found, 5 indicates high confidence that problematic concentrations will be (or have been) found, and 3 indicates real uncertainty either way. Other ranges and categories are possible.

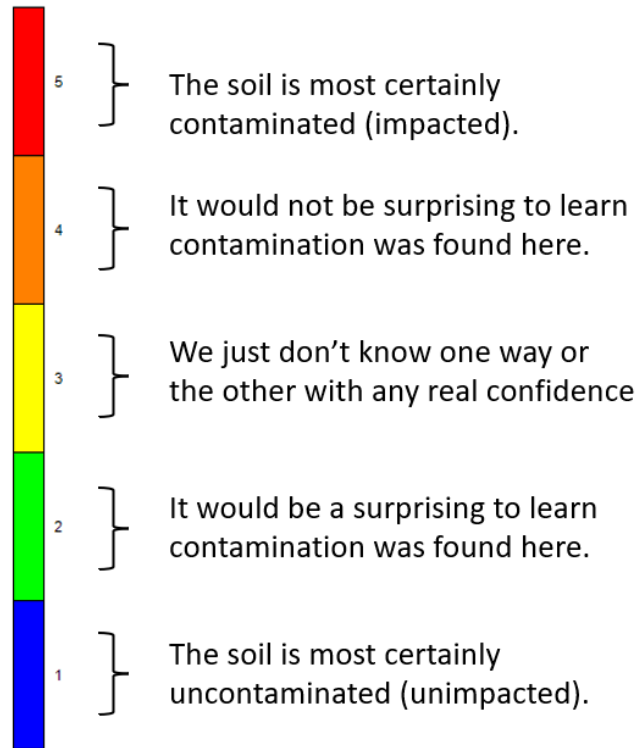


Figure 3.4. Example expert judgement legend.

In Figure 3.5, a continuous legend example is taken from (Stewart 2011) where a value of 1 means lowest possible concern, and 5 indicates highest possible concern. This expert opinion is essentially “painted” onto the foundational grid using graphical tools. Tools such as SADA (Stewart, 2009) allow for specification and directly “painting” values onto the foundation grid. GIS software packages likely also have options for editing grid values. GEM+ does not require that expert judgment necessarily be collected by depth. It is sufficient to indicate where soil sampling at any depth will likely find elevated activity levels.

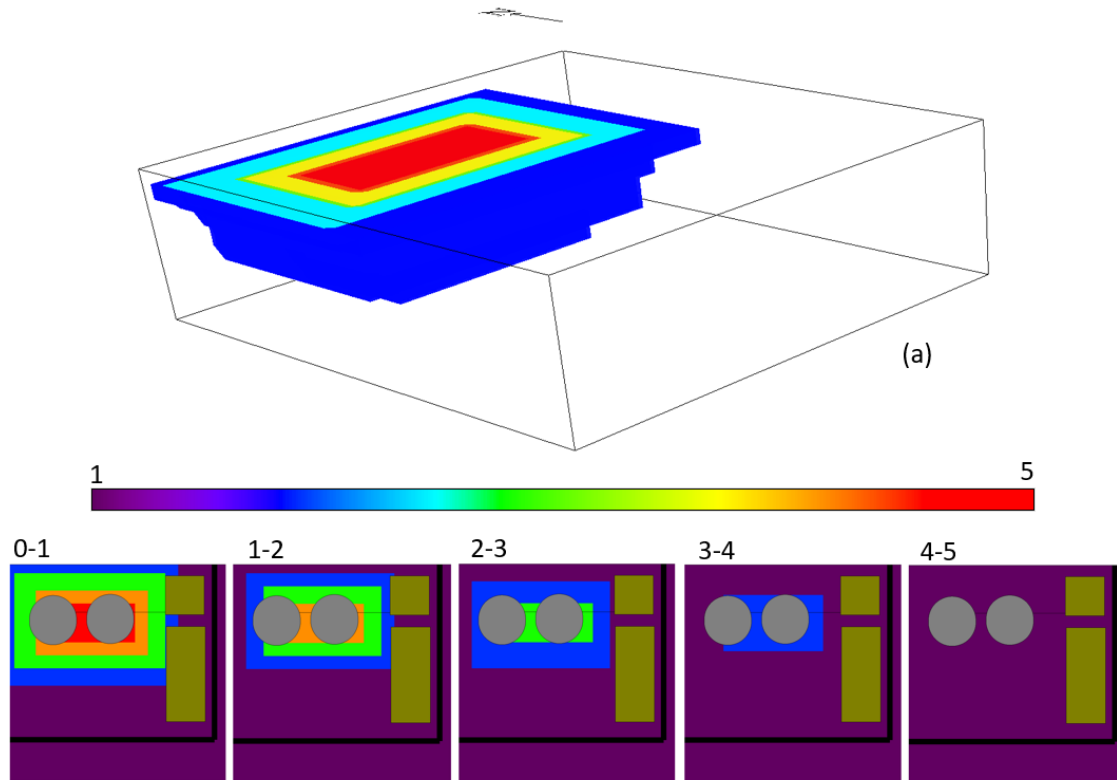


Figure 3.5. PSM for cesium-137 contamination (Stewart 2011). (a) Shown in the SADA 3D viewer for levels 2 and higher and (b) shown layer by layer in the SADA 2D viewer.

Sufficient historical sampling or monitoring may be available to construct the PSM with geospatial models right away. In this case, EPRI (2016) or Stewart et al. (2009) discuss a full array of approaches and software, ranging from deterministic inverse distance methods to full geostatistical simulation. If HSA findings are prepared to support a formal spatial model, then geostatistical simulations may be considered instead (discussed below).

A hybrid or semiquantitative PSM is also possible and will have a mixture of qualitative and quantitative sources. Some sampling may have been done, but the data may be insufficient to entirely create a useable data driven PSM. Several possibilities exist. In one scenario, some site areas may be constructed from the data in a local modeling effort and others partially from elicited judgment. In another scenario, investigators consider the available but limited sample values along with Likert scores to pragmatically create a PSM from both information types.

3.6 Geostatistical Modeling

Geostatistics is a cornerstone of the GEM+ framework. Geostatistics can provide a fine scale, 3D spatial model of radiological activity throughout the subsurface unit. There are two broad categories of geostatistical modeling relevant here: kriging and simulation.

3.6.1 Simulation

Geostatistical *simulation* is the preferred choice when scanning multiple $DCGLv_{local,i}^{(d)}$ with exposure geometries (a.k.a. different spatial supports). Furthermore, simulations permit a robust

estimation of compliance uncertainty anywhere in the subsurface volume for any $DCGLV_{local}$ based exposure geometry. This aligns well with the concept of power in the hypothesis test counterpart.

Geostatistical simulation provides many fine scale *joint realizations* of how subsurface contamination could be distributed given the data at hand. Figure 3.6 shows three subsurface simulations. In practice, 100s or 1000s are generated. Each realization represents a single scenario of how contamination could be distributed given available data. Generating multiple realizations allows stakeholders to access a range of possible outcomes and quantify the uncertainty about contamination distribution.

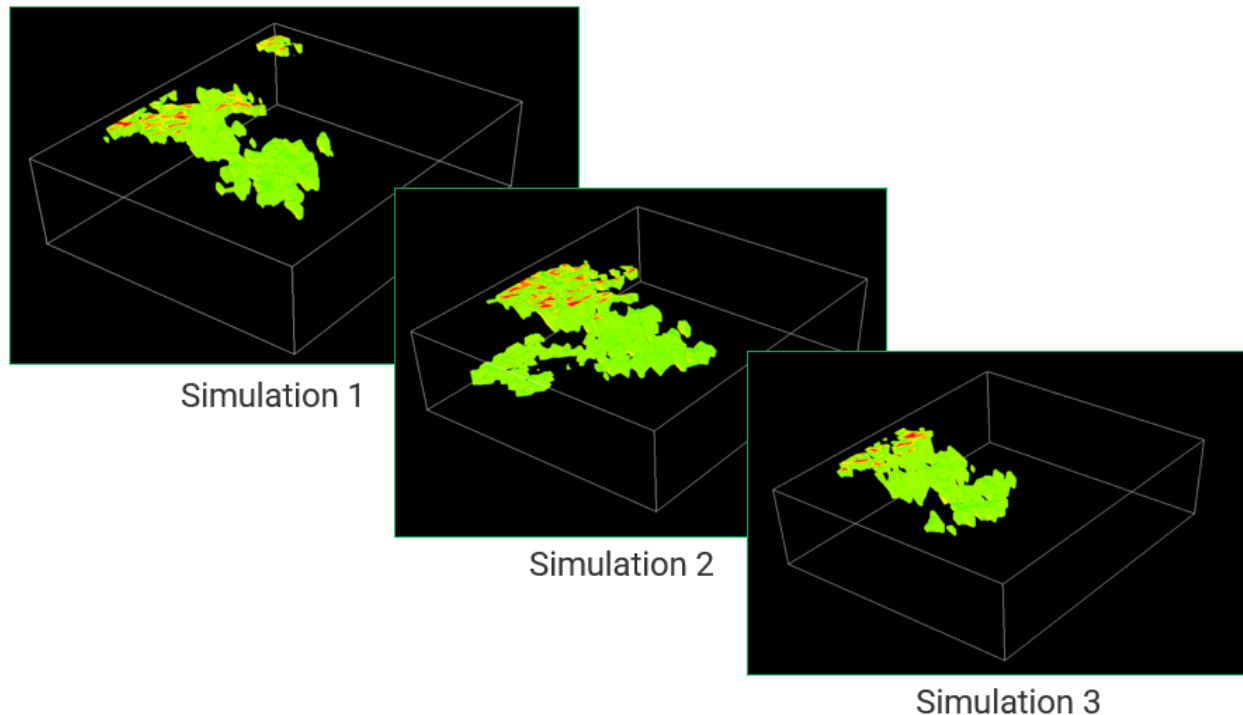


Figure 3.6. Three example subsurface simulations from the geostatistical simulation model.

3.6.2 Kriging

Another common geostatistical approach is kriging. Kriging is a geostatistical approach that produces the single best unbiased linear estimate at each point. That is, if the single best estimate is needed for any point, then the kriging mean is that estimate. A kriging map is therefore deterministic: one smooth “best guess” surface. Investigators may also adopt the kriging variance as a *heuristic* of estimation uncertainty to introduce a model of probability of uncertainty about the best point estimate value. The term heuristic is used here since the kriging variance for any point estimate is unfortunately not based on the actual value of nearby samples but rather the distance to them. The result is that variance is only a function of the spatial distribution of points and not their values (Goovaerts 1997; Deutsch and Journel 1992) as one might expect. It is possible then to observe the same kriging variance in situations where data values range widely and data values range minimally if the spatial configuration of samples is the same.

3.6.3 Simulation vs. Kriging

The challenge is that, unlike simulation, kriging does not allow consideration of the many plausible joint distributions of contamination that are possible across the site. Simulation by contrast, produces multiple equally probable realizations of the field. These are like random joint realizations or possibilities of true but unknown radiological activity could appear in the random field rather than just local means as in kriging. Each simulated map assigns plausible *joint configurations* across estimation points, reflecting both point uncertainty and spatial or local area uncertainty (e.g., a local scan unit geometry). Because of this, simulations naturally support virtual scanning. At any scale or window geometry, one can aggregate the simulated values to see how compliance outcomes would vary. In short:

- Kriging describes uncertainty about the average or best estimate surface but scanning it cannot explore possible real world contaminant configurations that could lead to compliance issues.
- Simulation describes uncertainty about the possible configurations of radiological activity, and therefore directly supports scanning and compliance evaluation at any DCGLv scale. Furthermore, postprocessing of simulations can produce average or expected values at each point (as in kriging) that retain greater heterogeneity than the kriging counterparts (Figure 3.7)

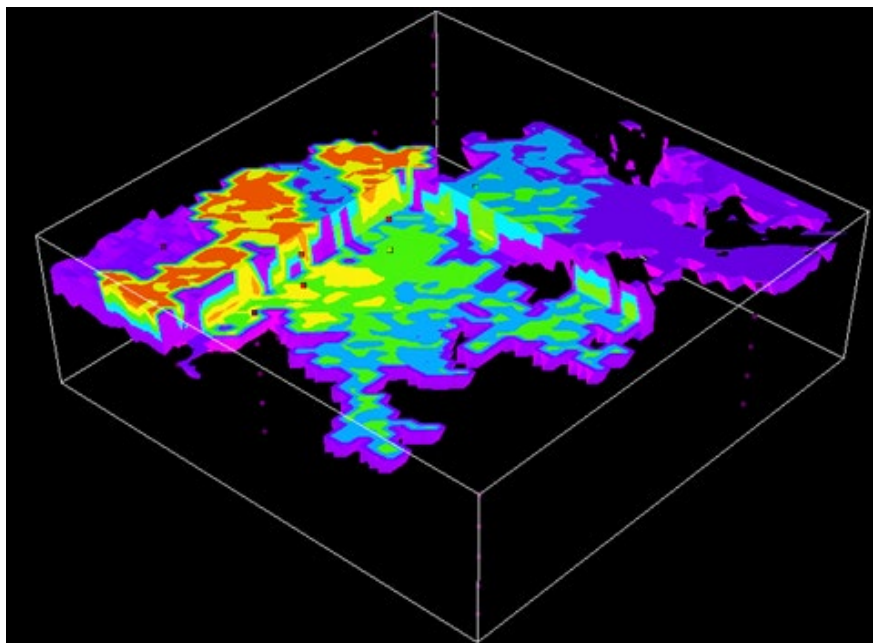


Figure 3.7. Map of expected activity levels developed by post-processing simulations. Simulation base map retains greater heterogeneity than kriging estimates.

For the GEM+ framework, kriging can still play an important role. Investigators could first calibrate their geostatistical model by parameterizing and testing the kriging counterpart to their simulation choice. For example, sequential indicator simulation (SIS) is the simulation counterpart to indicator co-kriging. Model parameter needs for both kriging and simulation are very similar: simulation requires only a handful of additional parameter values above and beyond those required by its kriging counterpart.

3.6.4 Further Resources

EPRI (2016) provides a useful roadmap for developing a geostatistical model. Major steps include preliminary setup, exploratory data analysis, structural analysis (variogram), geostatistical modeling, and postprocessing (e.g., cross validation, sensitivity testing). These steps are well connected to the language of decommissioning and the DQO process and are highly recommended as investigators implement the GEM+ workflow. EPRI (2022) provides an excellent review of geostatistical modeling in the context of plant decommissioning. That work builds an understanding of geostatistics in general, including basic kriging and conditional simulation. The document reviews geostatistics in industry and at decommissioning sites.

Stewart et al. (2009) provide a very thorough discussion of geostatistics, including kriging and simulation in the context of the SADA 5.78. The discussion includes a review of principles, practical recipes for setting up the model, and examples performing geostatistical analysis. That document may be quite useful to understanding and implementing geostatistical modeling in a hands-on experience. Many types of simulation are available, including turning band simulation, sequential Gaussian simulation, SIS, simulated annealing, and filter simulation (EPRI 2016; Stewart 2011; Goovaerts 1997; Deutsch and Journel 1992). One simulation algorithm cannot be claimed to be best for all cases (Goovaerts 2001). The simulation method should be selected based on the kind of site-specific circumstances that investigators may face.

While GEM+ is agnostic to the choice of geostatistical simulation, one simulation approach stands out for its flexibility and robustness: SIS, which is an attractive approach in GEM+ owing to its ability merge hard (lab quality samples) and soft (field, scanning) data and makes no distributional assumptions (e.g., Gaussian). In this case, the conditional cumulative distributions at each point and over each area is empirically constructed directly from the data. Stewart et al. (2009) provides a practical overview and recipe for implementing the model, and Stewart (2011) provides a deeper discussion of SIS and imperfect field measurements. Stewart (2011) emphasizes two highly favorable properties of SIS:

- It uses nonparametric approaches to modeling (Goovaerts 1997, p. 284). Nonparametric methods are preferred in MARSSIM, evidenced by the emphasis on sign and WRS tests (USNRC 2000).
- It provides an accessible way to encode different kinds of information into the model (Goovaerts 1997, p. 395). This information includes both hard (laboratory) and soft (field detection measurements) data and other relevant foundation grid attributes such as geophysical measurements.

More information about geostatistical modeling can be found in the literature (Deutsch, Isaaks, and Srivastava 1989; Journel 1992; Cressie 1993; Goovaerts 1997; Wackernagel 2003; Chiles and Delfiner 2012; Schabenberger and Gotway 2017; Metahni et al. 2019; Ortiz 2020; Hockett et al. 2022; EPRI 2022). The latter two are contemporary literature surveys with tight connections to the challenges of subsurface contamination.

3.7 Geostatistical Site Model

Under GEM+ nomenclature, the geostatistical simulation model aided by postprocessing of simulation values is referred to as the GSM, and it facilitates a variety needs including virtual scans for likely exceedances of any $DCGLV_{local,i}^{(d)}$ or production of a single representative estimate of local activity levels (Figure 3.7). Figure 3.8 shows the GSM as a geostatistical

simulation model plus the post-processing step that can produce virtual scans as well as a range of other products including single estimate and local variance maps (similar to kriging). GSM can support readiness for FSS by estimating the probability of passing hypothesis testing as well (discussed in Chapter 9).

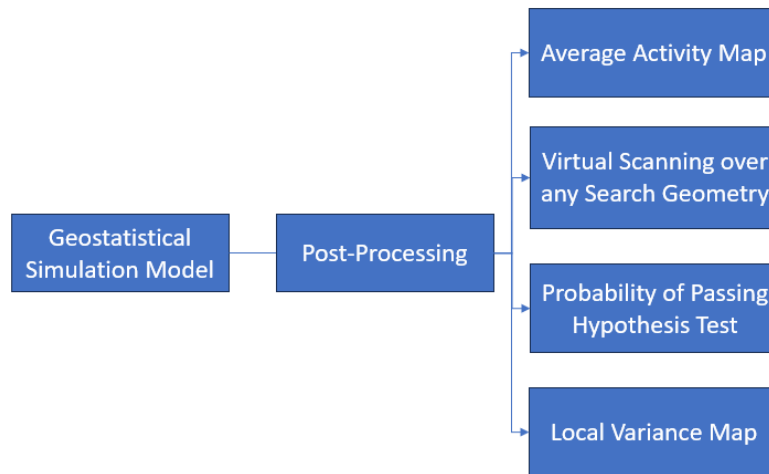


Figure 3.8. GSM is capable of producing multiple renderings (average map, local variance, virtual scans, hypothesis readiness testing) as needed.

Following the lead of Triad (Crumbling 2001a), the GSM is similarly evolved across the phases of environmental investigation, as new data becomes available through scoping, characterization, and remedial sampling. The GSM can assist in discovering $DCGLV_{local}$ exceedances during virtual scanning and aid in driving decisions about where to sample and where to remediate to meet compliance. In the end, the GSM, updated by FSS samples, serves as the basis for the virtual scan in the FSS. Virtual scans of the GSM are discussed later in the virtual scan section.

GSM depends on a foundational 3D grid system. In GSLIB (Deutsch and Journel, 1992), SADA (Stewart et al. 2009) and VSP (Matzke et al. 2014) this regularized 3D grid is specified by origin (x_0, y_0, z_0) defined within an area preserving the projection system either in meters or feet (not degrees) such that grid cell sizes $(\Delta x, \Delta y, \Delta z)$ have direct and interpretable meaning in terms of volume. The units should match the coordinate projection system of the CSM. Each node within the grid is identified by the center of the grid cell with a spatial coordinate $\mathbf{u} = (x, y, z)$, and an index value of 1 is assigned to the lower left grid cell. The index increases by 1, cycling first on x , then y , and then z (Deutsch and Journel 1992). The subsurface foundation grid serves as the finest spatial resolution in the investigation. It is the basis of the geostatistical simulations, which ultimately aid in sampling designs, remedial designs, and compliance checking that will follow. Analysts should confirm how grids are handled in their particular GIS software.

For nonrectangular-shaped sites with nonlinear boundaries or surficial variations, investigators can choose to mask out cells that lie above the surface or outside site boundaries or that contain subsurface structures (geophysical or manmade) that should be eliminated for various reasons (Figure 3.9). In software such as SADA, exclusionary polygons or layer designs help to accomplish any necessary masking.

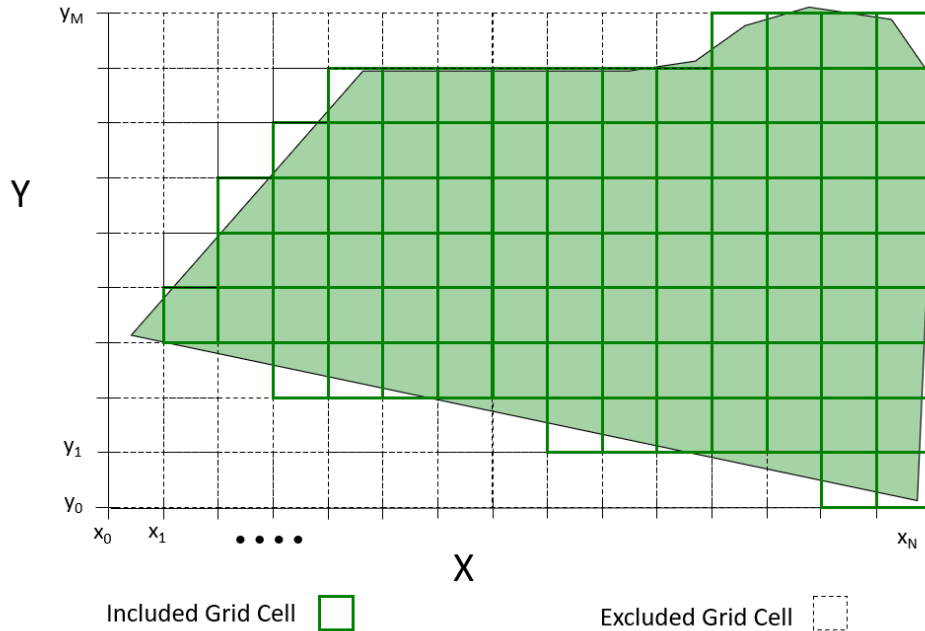


Figure 3.9. 2D visualization of included and excluded (masked) cells in the foundation grid.

3.8 Virtual Scanning

Given a geometry specification for each $DCGLV_{local,i}^{(d)}$, GSM will conduct a systematic kernel search through all simulations for all $DCGLV$ specifications to identify the number of exceedances and their location. Figure 3.10 shows two local $DCGLV$ specifications processed over a GSM stack of 50 simulations.

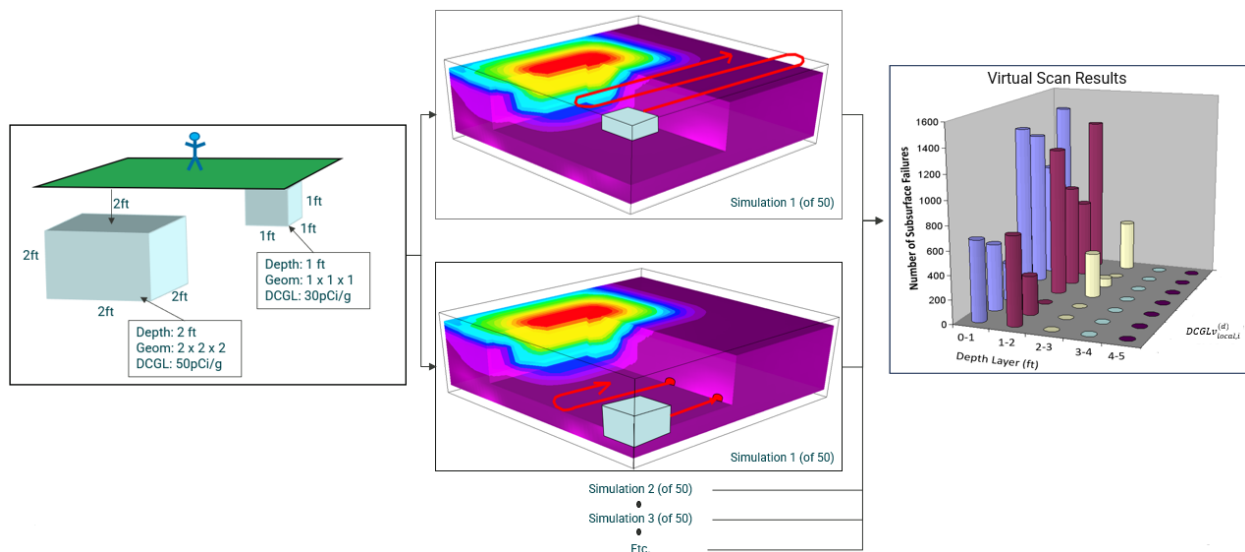
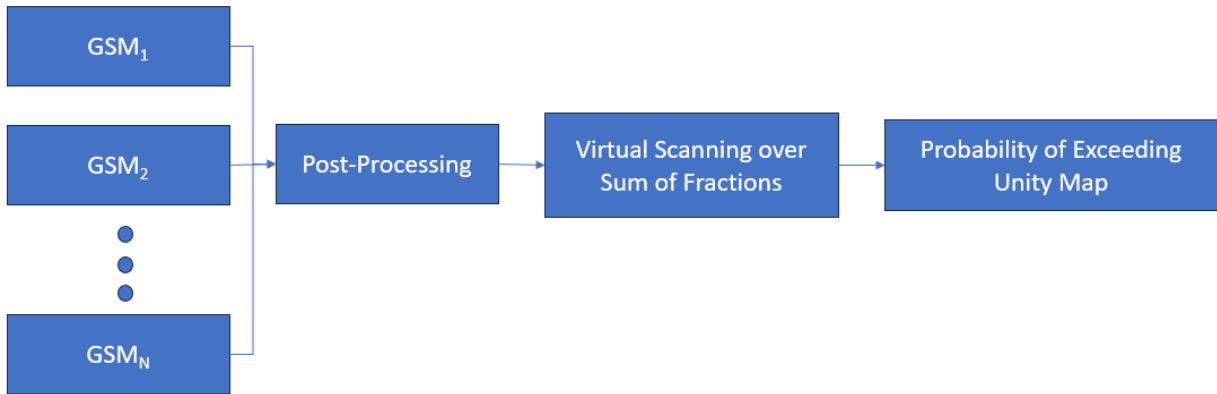


Figure 3.10. GEM+ scan of two local $DCGLV$ specifications through 50 simulations. Tallying the number of exceedances over all simulations.

1048 The virtual scan be thought of simply as a spatial search for noncompliant exposure volumes
 1049 located anywhere in the subsurface.

1050 **3.9 Multiple Radionuclides**

1051 In the case of multiple radionuclides GEM+ supports scanning for unity rule violations across
 1052 multiple GSMs, one per radionuclide. The requirement here is that each radionuclide has
 1053 equivalent scan geometries per DCGLv specification. For example, the $DCGLv_{local,i}^{(d)}$ scan
 1054 geometry specification for cesium-137 and cobalt-60 scan need to be identical to allow cross
 1055 radiological sum of fraction scans. For example, cesium-137 might have a 7 pCi/g $DCGLv_{local}$
 1056 and cobalt-60 a 5 pCi/g $DCGLv_{local}$ but over a common scan geometry of 1 ft × 1 ft × 1 ft at
 1057 depth of 2 ft. The GEM+ unit rule scan will then automatically compute the DCGL fraction per
 1058 radionuclide per $DCGLv_{local}$ and quickly scan for instances where the unity rule is violated. The
 1059 result will be a probability of unity rule violation map per scan geometry. The GSM unity
 1060 workflow is depicted in Figure 3.11.



1061
 1062 **Figure 3.11. GEM+ unity scan rule.** GSMs (one per radionuclide) are cumulatively scanned for
 1063 likely exceedances of the unity rule.

1064 Specifically, for a single scan point (Figure 3.13) we have for a collection of $DCGLv_{local,i}^{(d)}$ with
 1065 common geometry ($3 \times 3 \times 1$) and depth = d the sum of fractions for a particular scan location
 1066 (red box in Figure 3.12) is given by Eq. (1):

$$1067 \quad S_j = \frac{1}{|W|} \sum_{k \in W} \left(\sum_{i=1}^R \frac{C_{i,k}^{(j)}}{DCGLv_{local,i}^{(d)}} \right), \quad (1)$$

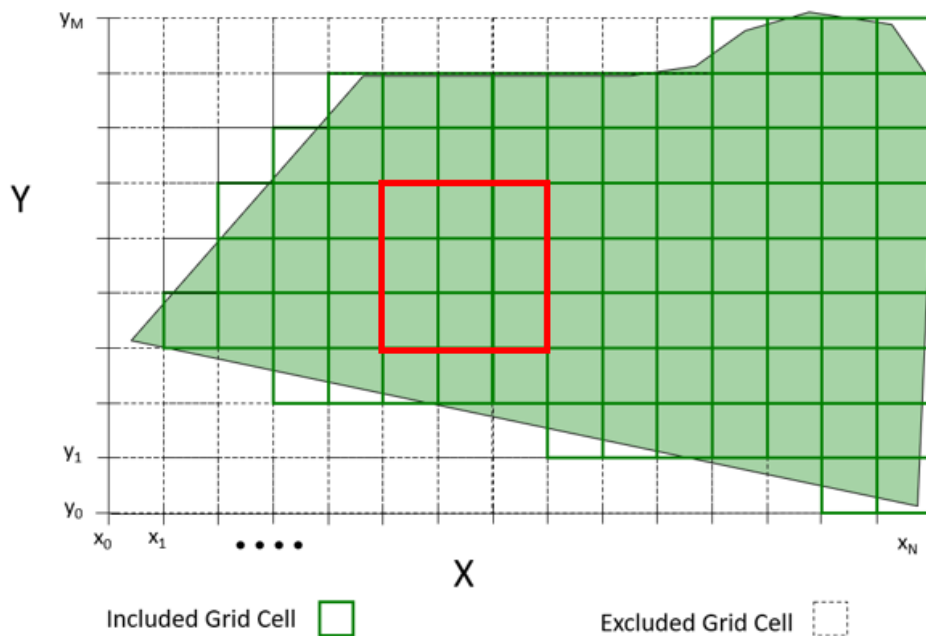
1068 where S_j is the sum of fractions of the j th simulation over window W (red box), $C_{i,k}^{(j)}$ is the
 1069 simulated concentration of radionuclide i at location k in simulation number j , and $|W|$ is the
 1070 number of spatial locations in window W (red box).

1071 From there we can compute the number of violations over N simulations, giving Eq. (2):

$$1072 \quad p(\text{red box fail}) = \frac{1}{N} \sum_{j=1}^N I_j, \quad (2)$$

1073 where

1074
$$I_j = \begin{cases} 1 & \text{if } S_j > 1 \\ 0 & \text{otherwise} \end{cases}.$$



1075

1076

Figure 3.12. A single 3 × 3 × 1 geometry scan point (red box).

1077 Postprocessing of the GSMs for unity rule violations is easily automated and computational
1078 times are rather short even on modern laptops or PCs. It amounts to a fast kernel search for
1079 instances where simulated radionuclide activity levels aggregated over a common scan
1080 geometry and divided by the DCGL exceed 1.

1081 Alternatively, it is possible to first convert samples to sum of fractions and then model a single
1082 GSM over the sum of fractions. This would allow stakeholders to only develop and evolve one
1083 GSM. It certainly allows a cleaner, easier approach. However, there are several drawbacks in
1084 this method. Stakeholders often need to know “Which radionuclides drive the exceedance, and
1085 where?” A direct sum of fractions map can’t apportion contributions or guide isotope-targeted
1086 remediation/ALARA tradeoffs. Each radionuclide may have different spatial scales and covariate
1087 structures. A single SoF variogram forces one covariance model onto a mixture of processes,
1088 potentially misfitting all of them. Secondly detection limits, counting stats, and quality control
1089 vary per isotope. Aggregating to sum of fractions before modeling masks could mask this
1090 heteroskedasticity. Finally, if you only model the sum of fractions, then you cannot use (or even
1091 diagnose) the cross-covariances between radionuclides that may strengthen overall
1092 assessment when one or more spatially covary together. Modeling the sum of fractions as a
1093 GSM is cautiously possible so long as radionuclides share similar spatial ranges/structures;
1094 censoring is minimal; you only need one scenario’s pass/fail probability.

1095

4 GEM+ Workflow

1096

1097

1098

1099

1100

1101

1102

The GEM+ workflow leads site investigators through the phases of investigation from the HSA to the FSS with two primary objectives. First, initialize and evolve (in the manner of Triad) a GSM that can increase efficiency in sampling, characterization, and remedial designs against a final status criterion. Secondly, the GEM+ leverages the GSM to estimate the likelihood of exceeding any $DCGLV_{local}$ through virtual scanning in the FSS evaluation. Additionally, GEM+ can leverage GSM to estimate the probability of passing a traditional hypothesis testing regarding any $DCGLV_w$ for indicating when a SSU may be prepared for FSS.

1103

1104

1105

1106

1107

1108

1109

1110

1111

1112

1113

1114

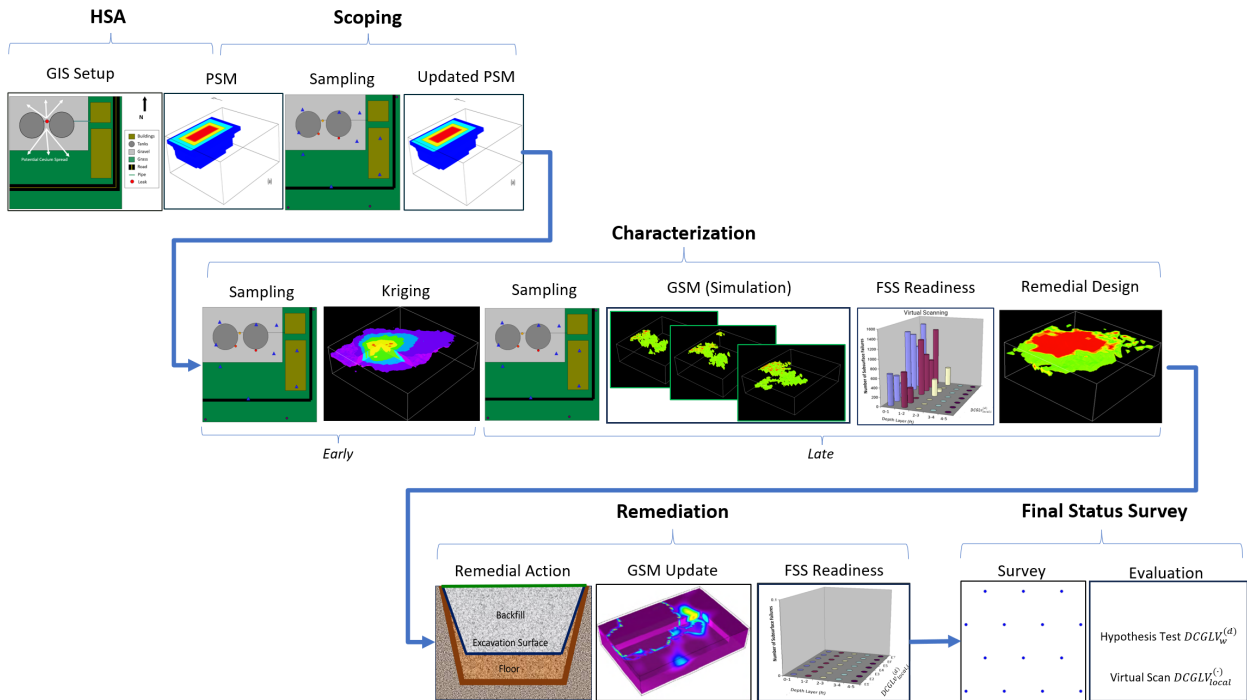
1115

1116

1117

1118

Figure 4.1 shows the GEM+ workflow across the major phases of environmental investigation. The workflow begins with an HSA that assesses whether the site is contaminated, establishes GIS mapping, and provides an early indication of where contamination in the PSM. The scoping phase picks up the PSM, collects scoping samples and updates the PSM for use designing and modeling contamination in the Characterization Phase. Based on the PSM, site knowledge and other inputs, early sampling efforts may yield a full GSM simulation model, however a simpler kriging model would be suitable. Later in characterization as more sample potentially become available the kriging model is evolved into a simulation model capable of virtual scanning, FSS readiness tests, remedial design, and other critical GEM+ functions. Based on the results of Characterization, the remediation phase takes action on the site and records any samples or soil removal in the GSM. FSS readiness tests can be reapplied to assess the likelihood of demonstrating compliance in the real FSS that follows. In the FSS, a two-part sampling strategy to support both hypothesis testing and virtual scanning is proposed. Hypothesis testing is conducted per layer in the traditional MARSSIM sign or WRS test. Virtual scanning is conducted for potential elevated areas within the subsurface after FSS samples have applied a final update to the GSM.



1119

1120

Figure 4.1. GEM+ workflow.

1121 The remainder of the document discusses each of these phases and GEM+ processes in detail.

1122 **5 HISTORICAL SITE ASSESSMENT**

1123 In GEM+, the HSA is the same well-documented process, questions, steps, and philosophies
1124 presented in Chapter 3 of MARSSIM. Readers are encouraged to concurrently read Chapter 3
1125 of MARSSIM. The purpose of the HSA briefly summarized is:

- 1126 1. to identify potential, likely, or known sources of radioactive material and radioactive
1127 contamination based on existing or derived information,
- 1128 2. to distinguish between areas that need further action and those that pose no threat to
1129 human health,
- 1130 3. to provide information useful in scoping and characterization surveys, and
1131 4. to develop an initial PSM.

1132 From the GEM+ vantage, the primary goal of the HSA is to create a PSM mapping likely
1133 contaminated areas (volumes) based on site operating history, historical sampling, and any
1134 other relevant information. Construction of the PSM is flexible and can be driven by qualitative
1135 or quantitative means as best suits the particular situation. The goal of the PSM is to clarify
1136 potential location of contaminants of concern, indicate likely Class levels, and support survey
1137 sampling in the next phase. Later in the characterization phase, GEM+ will pivot from this mixed
1138 methods model to a formal geostatistical site model as more empirical measurement become
1139 available.

1140 In the HSA, all available information regarding the operational history of the site, geography,
1141 geology, licensed materials, and so forth would be collected. For NRC licensees, there should
1142 be records maintained throughout their operations (MARSSIM 2000). This information is an
1143 excellent start for the HSA and may lead to a more rapid evaluation.

1144 The content gathered in this phase can be both qualitative and quantitative. The documentation
1145 of aspects such as facility operations, contamination events, waterway locations, soil types, and
1146 nearby affected populations comprises an important qualitative investigation. Investigators may
1147 also have access to ongoing monitoring data, historical surveys, formal geophysical surveys,
1148 and other geological descriptions. In some cases, information may be incomplete, potentially
1149 gathered for non-HSA purposes, or may provide unclear descriptions of accuracy. At this point,
1150 any information is welcome, whether it is qualitative or quantitative and regardless of questions
1151 about accuracy and current relevancy. However, a particular emphasis is placed on location in
1152 space and the use of geographic tools, such as GIS, to map site knowledge where possible.

1153 **5.1 Quantitative Data**

1154 Quantitative forms of data may be already available during the HSA. Direct measurements of
1155 radioactive contaminants may be available from existing monitoring programs or historical site
1156 evaluations conducted for other reasons. For well documented sites, there may be a
1157 quantitative understanding of geology or hydrogeology. Quantitative data may include
1158 groundwater modeling efforts, geophysical surveys, direct geological characterization, land
1159 cover data, and so forth.

1160 For this framework, a particular emphasis is placed on location in space and the use of
1161 geographic tools, such as GIS, to represent site circumstances. Recall that the common ground

1162 for decision making among many lines of evidence will be the PSM, which is spatially and
1163 numerically defined. Therefore, obtaining sample values as well as sample locations is
1164 imperative.

1165 A precaution regarding historical measurements cannot be overlooked. Conditions may have
1166 changed on the site since those measurements were conducted, radionuclides with short half-
1167 lives may have already decayed substantially, additional releases may have occurred, or
1168 decontamination may have already been performed. It is important to document how relevant
1169 data are for the current HSA, although no data should be omitted from the report. Guidance on
1170 these evaluations is provided by the data quality assessment process outlined in USEPA
1171 (2006a).

1172 **5.2 Qualitative Data**

1173 Qualitative data is highly valuable in this phase and can directly impact the conceptual site
1174 model and associated PSM. Examples include the acquisition and review of licenses, site
1175 permits, special authorizations, and operating records. These records can shed light on site
1176 activities, contamination control procedures, demolition, effluent releases, and various
1177 discharges. Records on storage, infrastructure failures, leaks, onsite landfills, decommissioning
1178 activities, research activities, test activities, fires, and so forth are highly relevant to this
1179 framework. Contacts, interviews, site visits and so forth provide considerable insight as well
1180 (USNRC 2000). Logs and notes from these efforts should be included in the body of knowledge.

1181 For the PSM, particular emphasis is placed on location and extent. If qualitative data can be
1182 expressed as location on a map, this is particularly useful. Obvious and direct examples include
1183 site plats, blueprints, sketches of structures, photographs, and modern GIS data, including
1184 roads, land cover, processing locations and so forth. Investigators should not be reluctant to
1185 manually add their knowledge to the map, even though they may not currently be in a spatial
1186 format. This could include bright lining areas that appear suspicious following a site visit,
1187 expressing expert judgment about where contamination might have moved, and so forth.
1188 Geospatial projects just as these encourage the use of qualitative data, informal notes,
1189 metadata, use of multimedia (video, animation) and so forth. There may be some reluctance to
1190 include information that is uncertain in nature or may not seem relevant at the time; however,
1191 these data should be included as their benefit and relevancy may appear only later in the
1192 process. The discussion now turns to more specific types of information that should be
1193 collected.

1194 **5.3 Radionuclides of Concern**

1195 Efforts should be made to identify the radionuclides used at the site. This list of potential ROCs
1196 can be evaluated to assess the potential for residual contamination. For sites with long
1197 operational histories, certain short-lived radionuclides may have already decayed sufficiently
1198 below any threshold for human health effects. Knowing which radionuclides may be present can
1199 also influence factors important to the subsurface, such as mobility.

1200 **5.4 Locations of Concern**

1201 An important task is distinguishing between impacted and non-impacted areas. Table 3.1 in
1202 MARSSIM offers some guidance in this process by providing a set of questions and
1203 commentary that can motivate or stimulate the investigation in the right direction. A shorter list
1204 of questions is produced here, based on Table 3.1, and when possible, questions are combined

1205 into more general statements. The purpose of this more generalized list is to set the stage for
1206 the kind of processes and evaluations that comprise this framework. If the answer to any of
1207 these questions is yes, it indicates a higher probability that the area is impacted (USNRC 2000).

- 1208 • Was the site ever licensed for the manufacture, use, or distribution of radioactive
1209 materials?
- 1210 • Did the site ever have permits to store, dispose or incinerate radioactive materials?
- 1211 • Was the site used to conduct research or perform tests that included the use of
1212 radioactive materials?
- 1213 • Was the site used for decontamination, maintenance, or storage of radioactively
1214 contaminated equipment?
- 1215 • Was the site involved in using, processing, storing, or disposing of naturally occurring
1216 radioactive materials?

1217 If the recommendations found herein are used as a technical basis for a future guidance
1218 document, Table 3.1 of MARSSIM should be revisited for the greater level of detail it provides.

1219 Non-impacted areas are identified through knowledge of site history or previous survey
1220 information, specifically those areas where there is no reasonable possibility for residual
1221 contamination, including some quantitative evidence from historical surveys or knowledge about
1222 decay rates.

1223 Impacted areas are those where operational history suggests contamination may exist or where
1224 historical survey results indicate a problem exists. Operational histories that suggest a more
1225 detailed investigation be conducted include (USNRC, 2000):

- 1226 • Locations where radioactive materials were used and stored,
- 1227 • Records that indicate locations of spills, discharges or other unusual occurrences, or
- 1228 • Locations where radioactive materials were buried or disposed.

1229 Areas immediately surrounding or adjacent to these locations are included as well, because
1230 contamination may have spread. The meaning of adjacent here is an important subject matter,
1231 particularly for subsurface. In subsurface, adjacency includes not only horizontal but vertical
1232 space. The size of the adjacent region can greatly impact the cost and extent of the sampling
1233 design.

1234 **5.5 Impacted Media**

1235 MARSSIM provides a considerable level of guidance on how to assess the media types that
1236 have been impacted. This document is largely concerned with subsurface soil but may have a
1237 role to play in groundwater evaluations as well. The questions found in this section draw on
1238 MARSSIM regarding subsurface evaluation.

- 1239 • Are there areas of known or suspected surface soil contamination? Given the type of
1240 radionuclide involved and the underlying geology, could migration to the subsurface
1241 reasonably occur?
- 1242 • Is there a reasonable chance for enhanced mobility of radionuclides? Certain other
1243 analytes can facilitate mobility in the soil.
- 1244 • Has the surface or subsurface been disturbed? Inexpensive geophysical surveys or
1245 simple visual inspection may provide clues.

1246 • Are there buried pipes? Trenches? Subsurface structures (basements)? Buried
1247 pits/landfills?

1248 Similar questions exist for groundwater, and the reader is directed to Chapter 3 of MARSSIM for
1249 a discussion of those.

1250 **5.6 Developing the PSM**

1251 Information identified in the HSA is used to create the PSM (see GEM+ Foundations for a
1252 description). This model describes where likely contamination exists, establishes classification
1253 levels, and establishes initial subsurface survey units (SSU). The PSM is only a preliminary map
1254 and will likely change as scoping data and characterization indicate. The benefit of the PSM is
1255 (1) to represent stakeholder understanding of where impacted, non-impacted, and Class 1, 2,
1256 and 3 volumes may exist, and (2) to support survey sampling and characterization until a
1257 geostatistical model can be established.

1258 Partitioning a site into survey units should be done with care. In geospatial reasoning, the goal
1259 is to characterize soil status by considering the continuum of radiological spread across the full
1260 subsurface. Separating a site into independently analyzable survey areas may disable valuable
1261 spatial context and information potentially necessary for the geospatial model to function as
1262 intended. For example, the PSM in Figure 5.1 expresses concern about contamination near the
1263 center of the site. The high Likert scores in the middle of the figure (red) indicate high concern
1264 that contamination exists there. The region of high concern is spatially bounded by a “possibly”
1265 contaminated region (Likert = 3), followed by a “unlikely” zone (Likert = 1). Here, and later in the
1266 GSM, separating a geospatial model into Class 1, Class 2, and Class 3 would be undesirable
1267 because data collected in each of the color-coded areas in Figure 5.7 would most likely inform
1268 conditions in all of them, depending on the spatial autocorrelation structure. In this case, the
1269 entire area should be considered Class 1 to include the unlikely regions as a way of bounding
1270 and refining the geospatial model, which, here, models the contamination process as a whole.

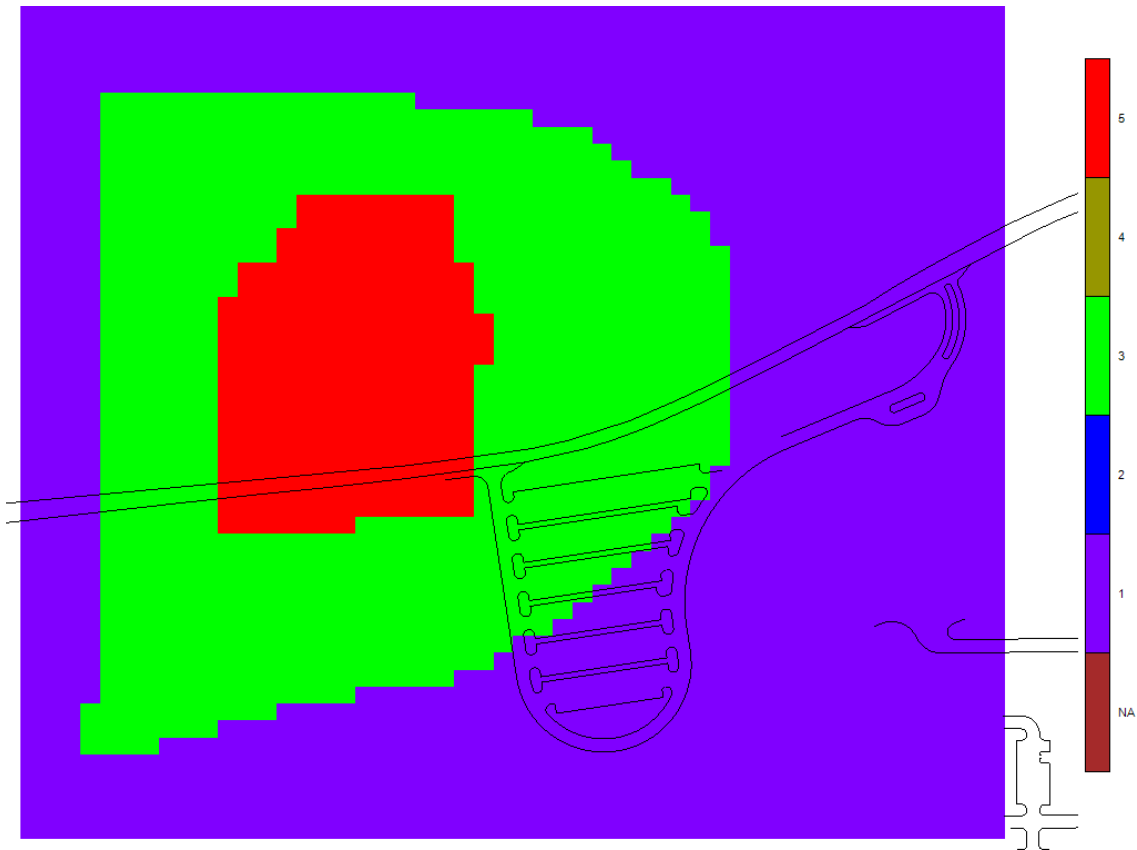


Figure 5.1. Initial PSA based on HSA.

Alternatively, Figure 5.2 shows a PSM that includes two separate areas of possible concern on the site. These areas are well separated from one another by an area considered impacted or in this case a Class 3 area. In this case, it makes sense to classify the area in the upper left-hand corner as Class 1 owing to the high Likert values (or concentrations) and the lower right as Class 2 owing to the midrange Likert values.

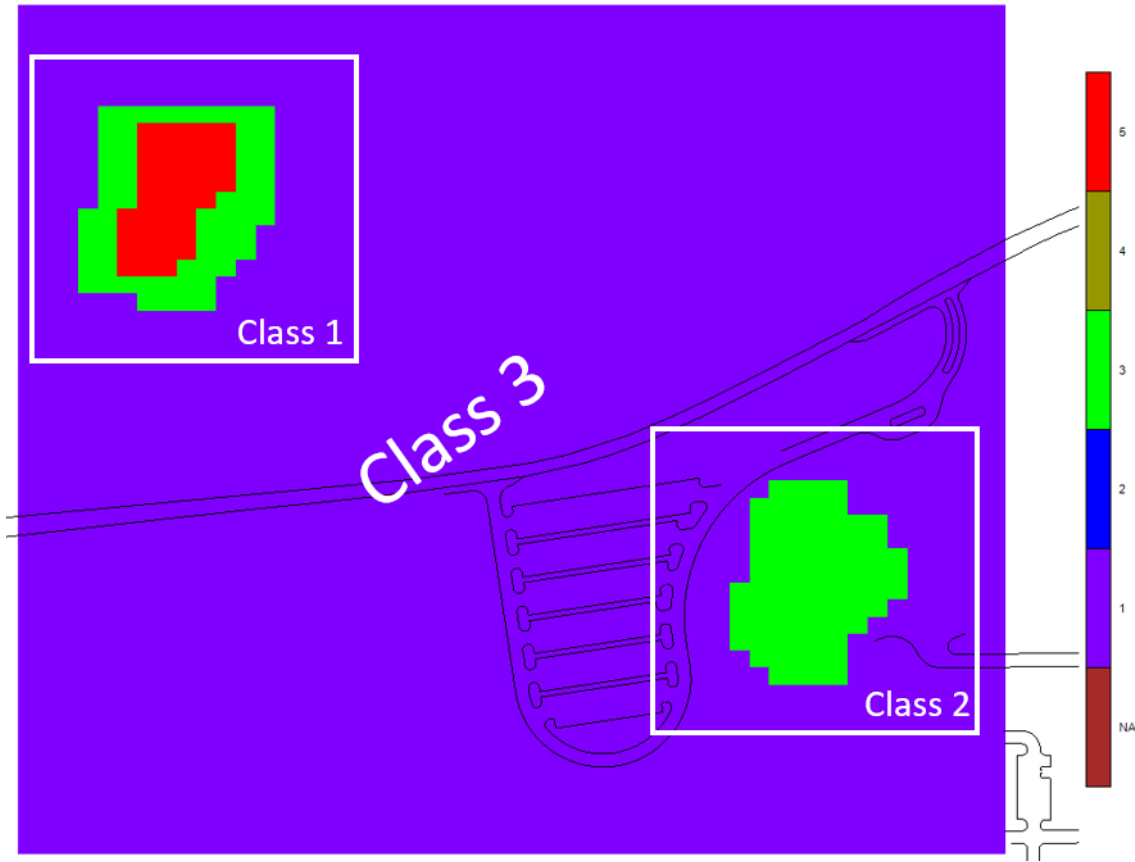


Figure 5.2. Spatially separable concerns in a PSM.

In both the Class 1 and Class 2 designations, the bounding box has been sized to adequately bound the area of concern (AOC) with Class 3 space or non-impacted space, allowing a proper sampling and modeling unit for detecting and delineating the modeled boundaries of contamination. This configuration requires sampling inside and outside the AOC.

This classification guidance is well aligned with the intent of NUREG 1757 Vol 2 Rev 2 (p. A-4):

For soil, survey units should be areas with similar operational history or similar potential for residual radioactivity to the extent practical. Survey units should be formed from areas with the same classification to the extent practical, but if areas with more than one class are combined into one survey unit, the entire survey unit should be given the more restrictive classification. Survey units should have relatively compact shapes and should not have highly irregular (gerrymandered) shapes unless the unusual shape is appropriate for the site operational history or the site topography.

In this context, the GEM+ process of classifying and clustering into larger units reasonably strikes a balance between partitioning a site into manageable decision units while retaining the spatial fidelity of models operating in those units.

1295

6 SCOPING PHASE

1296 Gogolak et al. (2022, Section 5.3) discuss the scoping phase in the context of subsurface
1297 contamination, including the importance of site conceptual models (covered in the GEM+
1298 preparation phase), developing a sampling and analysis plan for subsurface soils with the DQO
1299 process in mind, and determining preliminary scenarios and corresponding DCGLvs. This
1300 present guidance leverages and builds on this progress. The scoping survey should be
1301 conducted with site-specific conditions and knowledge in mind while following the recommended
1302 DQO process.

1303 **6.1 Sampling**

1304 Numerous sampling strategies given in the literature are suitable for the scoping survey
1305 (Huckett et al. 2022; Stewart 2011; Stewart et al. 2009; Johnson et al. 2005; USEPA 2002;
1306 Gilbert 1987). At a minimum, enough samples should be collected to confirm the presence or
1307 absence of elevated activities in areas indicated by the HSA and articulated through the PSM.
1308 Ideally, enough samples should be collected at sufficient depths to provide empirical clarity on
1309 the location of SSU walls and floors and position the effort for geostatistical modeling. Readers
1310 will find a comprehensive but non-exhaustive list of possible sampling strategies in the
1311 Appendix. Practical suggestions from the list for the scoping phase include (but are not limited
1312 to) judgmental designs, check and cover, and high value. These strategies assume expert
1313 knowledge either expressed in the manual location of samples (Judgmental) or in the PSM (high
1314 value, check and cover). Both offer considerable flexibility in the practical use of baseline
1315 knowledge. Sample types can vary widely including laboratory quality samples, field
1316 measurements, or core scans. These can be used effectively in updating the PSM and
1317 quantitatively leveraged in developing the GSM. Figure next shows a Check and Cover design of
1318 9 samples distributed around judgmental samples (colored circles) from

1319 **6.2 Updating the PSM**

1320 Regardless, for GEM+, the scoping phase data should be used to update the PSM with
1321 quantitative evidence of impact/non-impact and reinforce HSA determinations of Class level.
1322 Depending on the total number of collected core holes (historical data or scoping efforts) three
1323 possibilities exist for updating the PSM.

1324 In a qualitative scenario, investigators interpret the results of scoping in a continued qualitative
1325 PSM. HSA proposed SSU unit wall and floor boundaries are revised and Classification levels for
1326 SSUs are adjusted if needed. This is likely a preferred choice when a small number of core
1327 samples have been taken (e.g., 5–10).

1328 In a quantitative scenario, the number of survey core hole data combined with relevant historical
1329 data is sufficient to produce a quantitative PSM using spatial models such as kriging, simulation
1330 or other. The PSM could be a single map based on straightforward activity levels, sum of
1331 fractions aggregation, or collection of PSMs, one per radionuclide. Shifting to a geospatial
1332 model at the end of the scoping survey is a strong prepositioning for development of the GSM.
1333 This is likely a preferred choice when the number of core holes is 15 or more.

1334 In a hybrid scenario, it is possible to computationally pivot from a qualitative to a quantitative
1335 PSM using covariate geospatial models (e.g., Markov–Bayes, co-kriging). In this scenario, the
1336 prior qualitative PSM is used to condition the model, along with the hard sample values. These

kinds of models are found in a variety of geostatistical packages (e.g., SADA, GSLIB). One problem occurs if the prior PSM is significantly out of agreement with the scoping results. In this case, investigators might revisit those prior assumptions about site processes and develop a new qualitative PSM that reconciles the differences. If there is disagreement, the updated PSM may appear unusual, as the model attempts to reconcile the two irreconcilable perspectives.

Figure 6.1 (taken from Stewart [2011]) demonstrates the progress of a hypothetical site from an HSA understanding (a, b) through a scoping survey (c) and a shift to a geostatistical PSM (d) although the latter may not be common or necessary in practice.

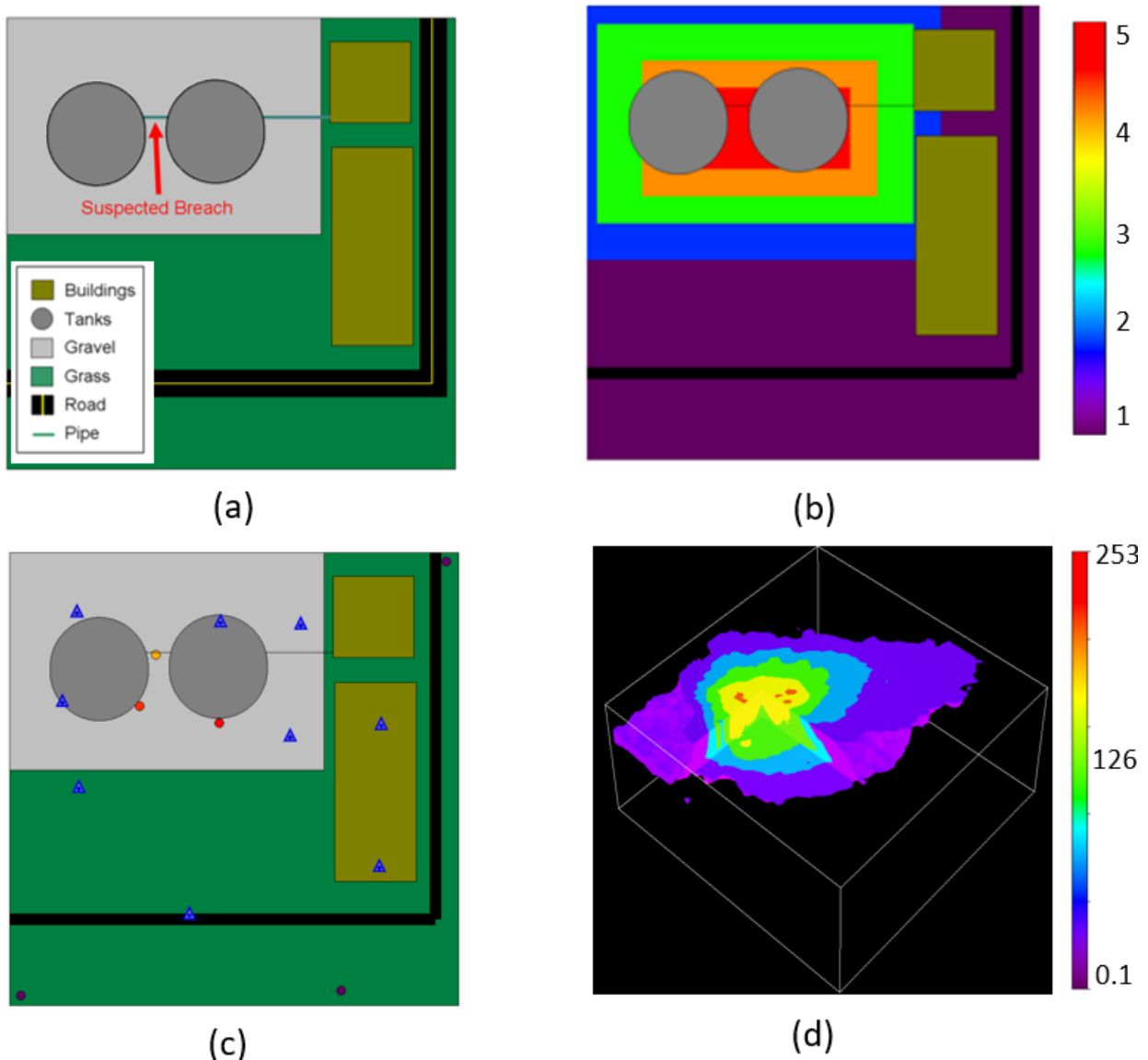


Figure 6.1. Progression of updating an HSA to a PSA. (a) HSA map of likely problem area, (b) PSM of likely contamination spread, (c) check and cover adds new core locations (blue triangles) to judgmental samples (circles), and (d) quantitative PSM/GSM based on geostatistical model.

7 CHARACTERIZATION PHASE

The goals of this phase are to determine the nature and extent of subsurface contamination, evaluate remedial alternatives and technologies, provide additional input to pathway analysis/dose or risk assessment models for determining site-specific DCGLVs, and demonstrate compliance against applicable regulations (Gogolak 2022; USNRC 2022b). During this phase, the GEM+ primary goal is to develop a stable GSM that supports robust virtual scanning and can support remedial design choices. This can be achieved by leveraging any data collected during this phase but especially those suggested by the model for their stabilizing value.

Stability of a geostatistical model can be assessed along several lines. First the underlying variogram specification is reasonable and further precision in its parameter values results in inconsequential variation in the virtual scanning outcome. Secondly, the spatial search neighborhood geometry, informed by the structure of the variogram and cognizant of the spacing of core holes, is selected so to produce robust model estimates and with small parameter variations incurring inconsequential differences in the virtual scanning outcomes. Cross validation is a valuable tool as well, however, with limited data sets (e.g., fewer than 30 core holes) its results should be proportionally considered. Due to spatial sparsity, some core holes are critical to supporting the model in the region and estimation in their absence may suggest a worse estimation than actually is occurring. Other factors concerning stabilizing the GSM may come into play and it is important a qualified analyst is involved to ensure proper modeling and stability assessment.

Sample data coming out of the scoping phase can vary widely. On one hand, historical and scoping samples constitute a healthy sample base that immediately supports construction of a viable GSM. In other cases, only a handful of examples exist, and information may be somewhat subjective and still expressed through a qualitative PSM. For this reason, the characterization phase is divided into an early and later phase to aid in choosing sampling strategies and enabling purposeful development of a stable GSM. Figure 7.1 shows the evolution of site information across the Early and Late Characterization stages. The concept of early and late are fuzzy in nature and clearly a transition between the two occurs. In the end, the characterization phase should have a GSM stable enough to support a virtual scan test and even produce a remedial map designed to support compliance in the FSS.

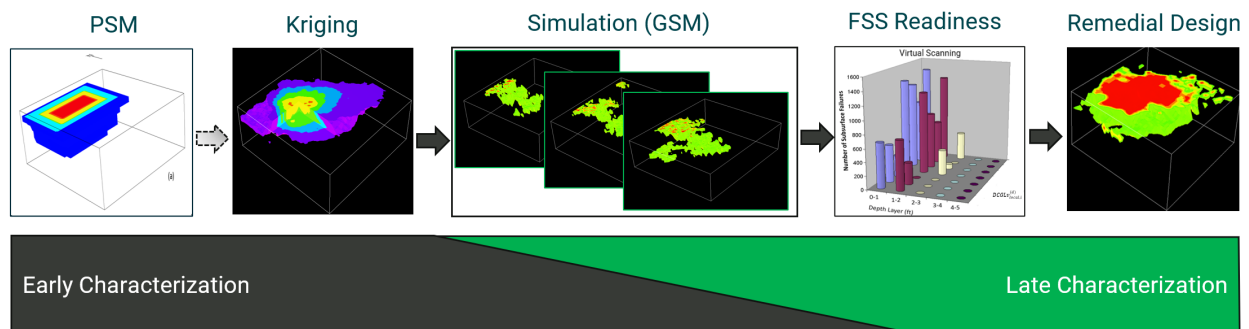


Figure 7.1. GEM+ in early-to-late characterization phase.

7.1 Early Characterization

The reasonable goal in early characterization is to develop a stable kriging model, ideally an indicator kriging model. As mentioned before, indicator methods (indicator kriging, SIS) easily integrate multiple forms of data and do not yield smoother Gaussian based estimates as in ordinary kriging or sequential Gaussian simulation. However, early on it may be useful to choose a simpler ordinary kriging model or indicator kriging model as a milestone for developing a simulation-based GSM. Kriging models run quickly allowing rapid parametrization of the variogram, search neighborhood, and so forth. Analysts may even consider easier deterministic methods such as Inverse Distance (Stewart 2009) as a pre-step to kriging. Figure 7.1 next shows that early characterization begins with the PSM, and the shaded arrow pointing to the kriging milestone indicates that this model qualitatively or quantitatively informs choices made to efficiently develop a kriging model. For example, a qualitative PSM might be leveraged with judgmental, random, or check-and-cover sampling to place new characterization samples. A quantitative PSM may already achieve the kriging milestone, allowing later stage secondary sampling designs immediately.

GEM+ core hole design strategies to consider in early characterization are shown in Table 7.1 and described in the Appendix. These are not an exhaustive list of design strategies that may be useful early in characterization.

Table 7.1. Early-stage design options.

Judgmental	Adaptive fill
Simple random	Bayesian Ellipgrid
Simple gridded	Check and cover
Simple unaligned grid	Threshold radial

Core hole sample depths should be selected to align with grid spacing to best support the GSM and the known scan level depths driven by DCGLvs. Vertical spacing need not be as fine scale as vertical layers can be in the GEM+ grid, especially if a full 3D approach is used. If a 2.5D geostatistical approach is used (sometimes called layer) where a separate model is run for each layer, then in turn, each layer would need to have a representative sample from each core hole. Layered modeling is not recommended as it can create artifacts on layers that core hole sample collections skip. Sample costs can be controlled by mixing lab, field, and scan results that can be combined later under an indicator-based model.

Samples are then used to develop or update a kriging model and give a first look at the distribution of subsurface activity levels.

7.2 Late Characterization

With a kriging model in hand, it is pragmatically easy step to shift to the simulation counter part by adding only a few additional parameters. Selection of tail parameters or number of simulated values to use are additional parameters (Deutsch and Journel 1992) that are relatively easy to choose and tend to be robust to model outcomes. It is possible to allow the GSM itself to suggest where new samples would be helpful to confirm boundaries, to confirm high value areas, or to reduce variance in the model. For example, postprocessing the GSM can produce maps of high-value areas or high-variance areas that can drive secondary designs such as high-value and variance-based designs. Other designs such as MrSDM (Stewart 2011) can

1422 directly position samples where new information has the greatest potential to reduce overall
1423 remedial volume.

1424 Table 7.2 lists later stage sample strategies driven by the model and/or the specific needs of the
1425 investigation.

1426 **Table 7.2. Later stage design options**

High value	AOC boundary
Adaptive fill	Structural
High variance	MrsDM

1427
1428 As always, sampling should be performed with the DQO criteria and site-specific circumstances
1429 in mind. General entry points for broad geospatial based sampling design include Lark (2016),
1430 Delmelle (2012), Li and Zimmerman (2015), and Marchant et al. (2013). Samples collected in
1431 later characterization should be used to update the geostatistical simulation model.

1432 **7.3 Remedial Design Support**

1433 As a GSM is taking shape and stabilizing, it is useful to engage in virtual scanning right away.
1434 Scanning each DCGLv across the GSM can produce maps of compliance failures given the
1435 information currently in hand. There would be one scan map per radionuclide, a single map for
1436 sum of fractions, or a single map that conflates compliance failure locations (union of maps).

1437 Whether an area is considered a compliance failure depends on the risk stakeholders are willing
1438 to take. For example, a particular scan cell may have a 1% chance of exceeding the
1439 corresponding DCGLv value but would not be included in the AOC if the stakeholder indicated
1440 5% or more probability. It would be included if the stakeholder indicated 1% or more. Revisit
1441 Virtual Scanning in the GEM+ foundations for more information.

1442 When a multiscale DCGLv site fails compliance, investigators may wish to know what soils to
1443 remediate or remove to bring the site into compliance. In particular, they are interested in the
1444 minimum volume of contaminated soil that must be removed, replaced, or cleaned to move the
1445 site into compliance. In addition, investigators will want to know where this volume is positioned
1446 on the site. GEM+ assesses compliance over multiple exposure unit volume sizes. Some of
1447 those smaller failing units may be located within larger units, others may overlap one another,
1448 and any number of configurations could occur. The MrDM takes advantage of these overlapping
1449 failures to optimize the remedial design. Although the optimization algorithm is dense, the idea
1450 is rather intuitive: using the foundation grid and treating each grid cell as a single remedial
1451 volume, the algorithm strategically selects just enough cells for remediation to bring all exposure
1452 failures into compliance. Stewart (2011) provides a detailed account of the approach, noting
1453 limitations and opportunities for improvement.

1454 Generally, given a GSM and a set of failing georeferenced exposure unit volumes, MrDM will
1455 strategically develop a remedial design map that will allow site investigators to achieve
1456 compliance over all exposure volume scales, depths, and DCGLvs, while removing the least
1457 amount of soil volume. Figure 7.2 shows the results of MrDM run on the Cesium Site from
1458 Stewart (2011). This remedial design will ensure that all exposure scenarios at all scales and
1459 depths will be satisfied with a minimal amount of soil removal.

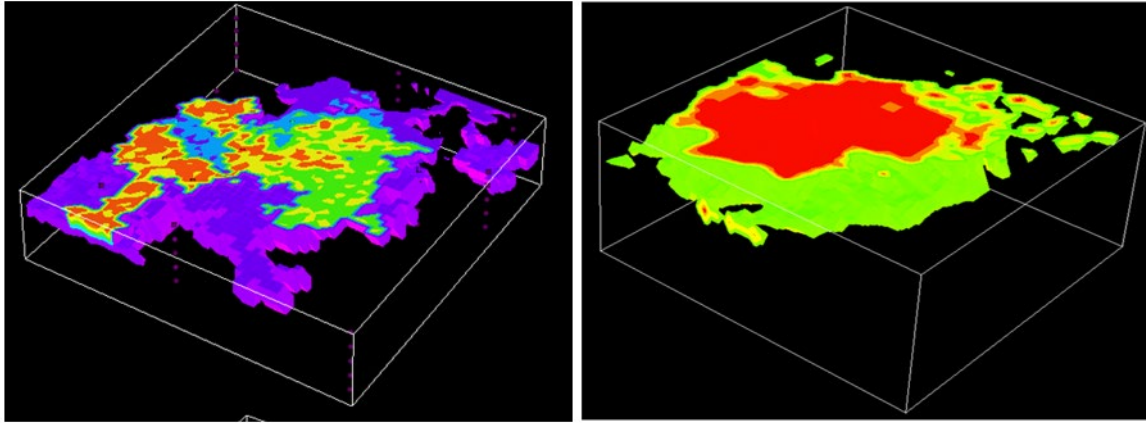


Figure 7.2. Cesium Site MrDM from Stewart (2011). (left) GSM average concentration and (right) MrDM design simultaneously bringing the site into compliance over seven different exposure scenarios. Taken from Stewart 2011.

For multiple radionuclides, the remedial design could also be based on the sum of fractions rather than per radionuclide remedial design. To accomplish this, a GEM+ Unity Scan is used rather than a radionuclide specific GSM scan under a MrDM application.

8 REMEDIATION

During characterization, GEM+ can develop a volume of concern map that indicates where remediation could mitigate exceedances of DCGLvs. In practice, this volume of concern may or may not be directly equivalent to the final remedial design or action owing to complications in the subsurface, considerations of safety, engineering requirements and so forth. It is necessary that the GSM is updated by whatever sampling or removal actions have occurred during remediation.

Backfill soils in particular must be treated carefully in the GEM+ process. Geostatistical modeling is properly applied to in situ soils where contaminating processes have ostensibly created a spatial auto-correlation structure that can be detected, modeled, and leveraged to quantify the spatial distribution of radiological activity in the subsurface. These in situ soils may be removed and replaced by backfill soils that have arrived from a separate local, likely have undergone some mixing, and otherwise exhibit zero or at least different autocorrelation structure. Given the guidance recommendations in NUREG 1757 Vol 2 Rev 2 and the DUWP-ISG-02 (USNRC, 2023) the most likely scenario is that considerable mixing has occurred as part of processing, lifting, moving, storing, and substantially scanning reuse soils. From a geostatistical modeling perspective, these soils artificial replacements of existing insitu soils in the model and they should not play a role in estimating activity levels in remaining insitu soils. Former in situ measurements should be retained in the simulation and modeling of geostatistical distributions radiological activity in the in situ soils. Reuse or backfill soil samples should not interact with the underlying geostatistical activity distribution model in any way. When backfill is involved there are now two frames of reference: a *geostatistical frame* where new and past in situ measurements remain in analysis (Figure 8.1a) and a *compliance frame* where sampling and assessment accomplished on the final configuration does not include excavated in situ measurements or volumetric estimate (Figure 8.1b). In Figure 8.1, these frames are represented separately. Blue dots and geostatistical model outcome represent values derived from in situ

soils. Gray dots and region represent the substitution of backfill estimates and measurements for readiness and final compliance checks.

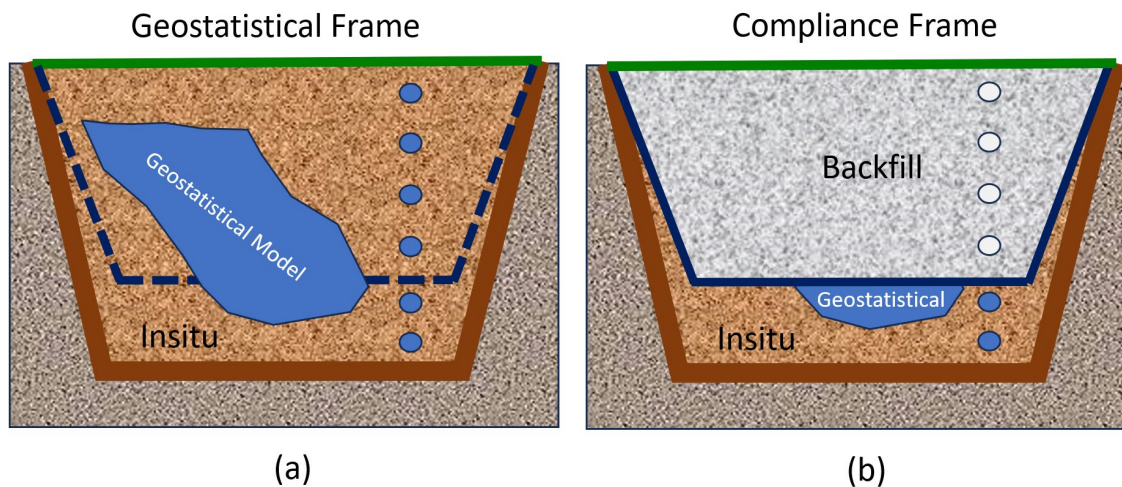


Figure 8.1. Frames of reference for geostatistical modeling and compliance under a backfill scenario. (a) A geostatistical frame includes the entire SSU soils (removed or otherwise) in order to support geostatistical modeling (b) Compliance is over the actual configuration which includes backfill and remaining insitu soils. Data from mixed backfill should not be used in the update of the geostatistical model. Backfill data is used in the final compliance.

From a 3D modeling perspective, the compliance frame is created by excavated insitu soils with measured or estimated backfill concentrations prior to FSS virtual scans. Detailed guidance of how deal with backfill soils is provided in NUREG 1757 Vol 2, Rev 2 and in the DUWP-ISG-02 (USNRC 2023). Readers should refer to this guidance when dealing with backfill and substituting the outcomes into the 3D compliance frame here.

8.1 Updating Remedial Actions in the GSM

For GEM+ the remedial update occurs in two steps. First, the GSM is updated using samples that have been collected during soil removal or sampled from the open excavation surface. Next, the soil removal (and replacement) is encoded in GSM by masking the affected areas with a representative backfill concentration value. Figure 8.2 shows the remedial update workflow.

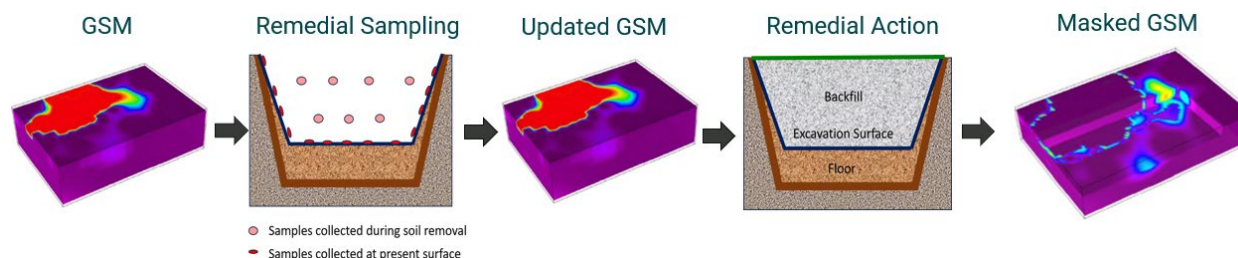


Figure 8.2. GSM remedial update workflow.

Successful integration with GEM+ will require proper documentation of the horizontal and vertical locations of samples and/or scan results (Gogalak et al. 2022). Without proper position

1515 data, use of GEM+ in a geospatial framework will be negatively impacted. Furthermore, properly
1516 positioning the excavation surface is important.

1517 **8.2 Open Pit Excavation and the DUWP-ISG-02 (USNRC 2023)**

1518 In many cases, investigators may opt to begin remediation by removing soil and sampling the
1519 soil and the surface of the excavation pit as digging progresses. US NRC (2023), Gogalak
1520 (2022), Hockett (2022), and USNRC (2022b) cover the details of this MARSSIM-based
1521 approach for dealing with this scenario. Generally, investigators begin excavation and soil
1522 removal early in the process and continue to excavate until the excavation surface
1523 demonstrates a MARSSIM surface compliance on the bottom and walls of the open pit.
1524 Generally, an agile relationship exists between remediation and characterization, and soil
1525 samples are periodically taken in the removed volume as well as the excavation surface as the
1526 dig continues. GEM+ can interact with this approach in three ways.

1527 First, GEM+ supports a priori estimation of excavation geometry during characterization. For
1528 example, deep excavation pits may require significant preparatory work to perform and safely
1529 access the excavation (USNRC 2022b). Pumping may be required to lower the water table.
1530 Depending on safety circumstances, regulators may be contacted to decide whether alternative
1531 scanning techniques can be used (USNRC 2022b). Estimating the likely geometry of the
1532 excavation ahead of time is critical to anticipating these problems and planning for them.

1533 Second, samples collected in the soil volume and on the surface of the excavation can be
1534 provided to the GEM+ process to rapidly update the simulation stack or kriging map to inform
1535 the remedial geometry in near real time. This capability is useful for anticipating changes to the
1536 remediation plans ahead of discovery in the field.

1537 Finally, GEM+ supports compliance evaluation by indicating compliance below the excavation
1538 surface. The continual interplay between GEM+ remedial geometries and the samples collected
1539 in the field should indicate whether any problems might exist below the MARSSIM-tested
1540 surface at a particular excavation surface point.

1541 **9 FINAL STATUS SURVEY**

1542 Four strategic objectives converge and complement one another under the GEM+ FSS
1543 workflow: 1) traditional sample design and hypothesis testing against the set of depth specific
1544 $DCGLV_w$, 2) supplemental sampling and virtual scanning against the set of depth specific local
1545 elevation criteria including $DCGLV_{emc}$, $DCGLV_{gw}$, $DCGLV_{nte}$, etc., for soils left intact and 3)
1546 application of the GEM+ unity workflow for multiple radionuclides and 4) assessment of the
1547 underlying GSM and decision stability. Note that if stakeholders have elected to model sum of
1548 fractions rather than individual radionuclide concentrations, Step 2 and Step 3 conflate into a
1549 single step of scanning the sum of fractions GSM. See Section 3.9 before carefully considering
1550 use of a single sum of fraction GSM.

1551 This GEM+ FSS approach provides the following benefits to stakeholders:

1552 **Integrated Hypothesis Testing:** A framework for systematically assessing radionuclide and
1553 depth-specific $DCGLV_w$, $DCGLV_{EMC}$, $DCGLV_{nte}$, $DCGLV_{gw}$ and others by integrating decision-
1554 based hypothesis-based design and testing on $DCGLV_w$ criteria with model-based (virtual scans)
1555 sampling and evaluation on the remainder.

1556 **Virtual Scanning:** A modeled scan of the subsurface (GSM) related to local exposure concerns
1557 $DCGLV_{local}$ for subsurface soils inaccessible to physical scans. This is extendable to multiple
1558 radionuclides through the GEM+ unit workflow.

1559 **Decision Stability:** Spatial contextualization around hypothesis testing and methods for
1560 evaluation of virtual scan (model) stability and to qualify suitability of use for compliance
1561 determination.

1562 **Readiness Checks:** Prior readiness checks to help assess whether a site is ready for a
1563 successful FSS by predicting the outcome and evaluation of FSS sampling, hypothesis testing,
1564 and virtual scanning.

1565 **Integration with Existing Guidance:** Integration with existing guidance including NUREG-1757
1566 Vol 2 Rev 2 (USNRC, 2022), DUWP-ISG-02 (USNRC, 2023), NUREG 7021 (USNRC, 2012),
1567 and NUREG 1575 (USNRC, 2000), and with consideration of recent advances in guidance and
1568 methodology including (but not limited to) Hockett et al. (2022), Gogolak et al. (2022), ASTM
1569 (2019), ASTM (2018b), ASTM (2018c), EPRI (2016), and Stewart (2011).

1570 Note that GEM+ may not be applicable in every subsurface case. For example, if the SSU is too
1571 small for statistical sampling (e.g., 300 m²) the entire volume may be treated as a potentially
1572 elevated area. The licensee may use other thresholds or professional judgment instead of
1573 GEM+ but these should be documented in the decommissioning plan (USNRC 2023, p. 2–29).
1574 We continue now with discussion of the foundational elements of the GEM+ FSS.

1575 **9.1 Final Status Survey Readiness Check**

1576 Prior to the FSS stage, it is recommended that licensee confidence in meeting the objectives for
1577 both $DCGLV_{w,i}^{(d)}$ and $DCGLV_{scan(\cdot)}^{(d)}$ be met by applying an FSS readiness check. The readiness
1578 check simulates the process of hypothesis testing and virtual scanning given the pre-FSS
1579 geostatistical model. Virtual hypothesis test predicts the outcome of a Sign/WRS hypothesis test
1580 and evaluation by using estimated activity samples drawn from the geostatistical model (instead
1581 of real samples). Readiness can be assessed for different spatial configurations (gridded or
1582 random) and different strategies (comprehensive or graded). The results can report the
1583 *probability of passing* the Sign/WRS test by observing what percentage of simulations that pass.
1584 This can be fully automated and on most commodity computers, it normally will only take a few
1585 minutes to complete.

1586 For virtual scanning (Section 3.3), the pre-FSS model is scanned to produce the scan summary
1587 tables and corresponding GSM maps. If the probability of failing a hypothesis test or virtual scan
1588 in the FSS is too high, licensees are recommended to first consider further remediation or other
1589 alternatives. Indeed, characterization phases (Chapter 7) and remedial design phase (Chapter
1590 8) aim to prepare the survey unit volume for successful FSS evaluation as part of their design
1591 strategy.

9.2 Hypothesis Testing

9.2.1 Overview

The first objective is to design and conduct a hypothesis test against $DCGLV_{w,i}^{(d)}$: per layer using the traditional Sign or WRS test and a corresponding sample design applied to core hole location and sampling. DUWP-ISG-02 (USNRC 2023). P. 2-13 says that identification and quantification of local elevated areas may not be part of the SSU investigation study objective. The following considers the case where indeed local elevation is not a concern.

9.2.2 Sample Design

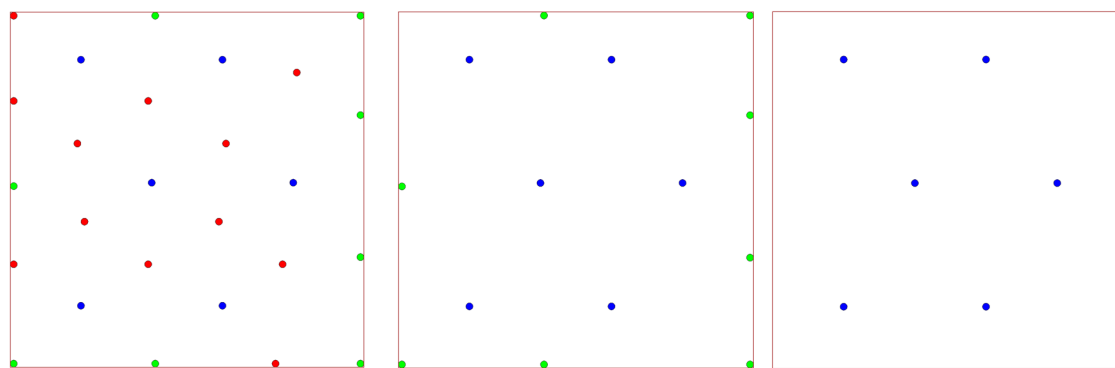
Under Class 1 and Class 2 core hole designs utilize a gridded sampling strategy, whereas Class 3 utilize a random design. Per NUREG 1757v2, G-7 depth dependent hypothesis test designs are developed to assess soil volume activity levels against survey wide, depth specific $DCGLV_w^{(di)}$ for $i = 1, \dots, N_{\text{depths}}$ in order to fully evaluate total volume activity (1757v2 G-10)². Extending the WRS/Sign sample design to multiple subsurface layers requires core hole location and core sampling across a stratum of possibly different DCGLVs and variance estimates. Treating each depth layer as unique and independent design could generate a cost prohibitive number of core holes. A couple of options exist.

A *conservative approach* is to determine a core hole design for the most restrictive DCGLV/variance combination and apply the design from the surface to the floor. This design will likely oversample other layers creating a higher than required confidence for those layers.

An alternative, *graded approach* begins with the core design strategy (gridded or random) at the bottom most layer. Based on the Sign/Wilcoxon Rank Sum sample size requirements a grid or random sample design is applied at this layer. In the next layer up, based on the requirements of the Sign/Wilcoxon Rank Sum, two possibilities exist. If more samples are required, the adaptive fill sample design method distributes additional core locations to maximize sample coverage. If fewer samples are required, core sections from select core holes are selected for archive and remain unsampled to reduce cost. The full set of core locations from lower level or newly added at the present layer become the anchor design for the layer one higher and this process continues until the first layer is reached. A couple of examples will help clarify.

Consider the scenario where as depth increases the number of required samples decreases (Figure 9.1). This could be because of less restrictive $DCGLV_w$ s or smaller variances or both. Beginning with the deepest layer, we learn that only six samples are required and distribute these in random start triangular grid. Layer two needs 14 samples. Six of these are satisfied by the 6 existing core holes designed by layer 3. An additional 8 are added by adaptive fill to maximize spatial coverage. The first layer requires 26 samples, 14 of which are satisfied by core samples accumulated during Layer 1 and Layer 2 designs. An additional 12 samples are added by adaptive fill. In total, the site requires 26 cores with varying depth and only 46 samples. Note that a conservative design would have required 46 cores and 138 samples.

² As per 1757v2 G-10 "The number of cores to be taken is initially the number (N) required for the WRS or Sign test as appropriate."



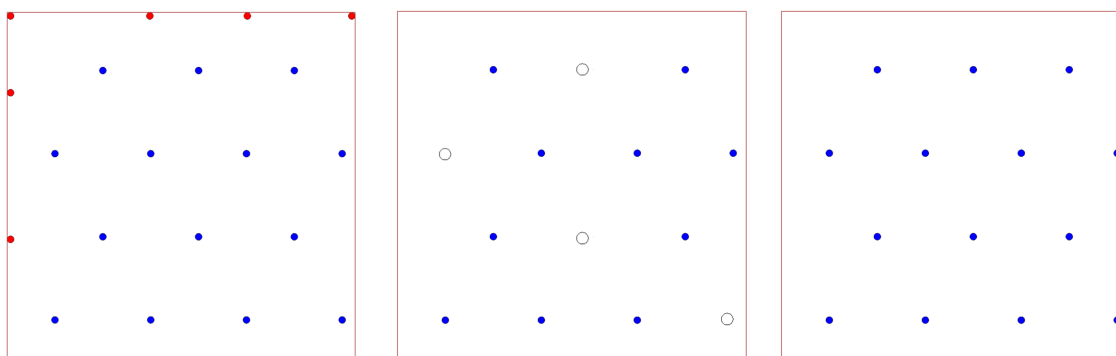
Layer 1: 26 Cores

Layer 2: 14/26 Cores

Layer 3: 6/26 Cores

Figure 9.1. An example graded core hole design for subsurface DCGL_v compliance. Beginning with the Layer 3 six-sample gridded design (blue), adaptive fill designs add 8 additional cores (total of 14) for Layer 2 and 12 additional for Layer 1.

Another scenario (Figure 9.2) might incur higher sampling requirements in the first and lowest level owing to exposure mechanisms associated with direct exposure and groundwater contribution respectively. The same rules apply as in Figure 9.1. The anchor design begins at layer 3. In layer 2 where fewer samples are required, the core hole samples are archived (empty circles). In layer 1, additional samples are added as required by hypothesis testing at the top layer. In this scenario, 20 cores are drilled and a total of 44 samples (20 + 10 + 14). In the conservative scenario 20 cores would be drilled and 60 samples collected (20 per layer).



Layer 1: 20 Cores

Layer 2: 10/20 Cores

Layer 3: 14/20 Cores

Figure 9.2. An example of bottom-up core hole design where samples may increase and decrease per layer. Beginning with Layer 3, 14 gridded samples anchor the design. Layer 2 has fewer sample requirements and so some randomly selected (empty circles) are archived

Samples taken from the core at each layer should be able to support the exposure scenario as well as any geostatistical model involved. NUREG 1757 encourages depth layers to be no more than one meter. In the case of multiple radionuclides, the number of sample requirements at each layer would be driven by the most restrictive radionuclide and every radionuclide would be sampled at each sample location per layer in order to apply the unity rule.

From a statistical viewpoint the graded design is well aligned with hypothesis testing. The design (at all three levels) is established with a random start point (bottom layer) and maximal

spatial coverage (adaptive fill) although the final complete design may not appear explicitly gridded above the bottom layer. Variations between completely conservative and completely graded are possible as well. For example, stakeholders may be willing to sustain conservative oversampling in the first few depth layers before shifting to a graded approach in deeper layers or vice versa. While different sidewall and floor configurations may not be planar, sampling locations should be established based on a flat, final graded area (NUREG-1757, G13).

9.2.2.1 Evaluation

Results of the hypothesis test for each radionuclide/depth layer should ideally be based exclusively on the results of the unbiased gridded or random design. However, licensees may petition for inclusion of other samples (e.g., from remediation, characterization) under the broader goals of a graded approach. The hypothesis test either Scenario A or Scenario B is carried out per layer as usual per MARSSIM. Application of the Unity Rule for sample points would apply per usual.

It is highly encouraged that spatial modeling be implemented to provide additional context and confidence in the assumptions of the hypothesis test and to assess pre-requisite decisions such as classification, and wall and floor position. When geostatistics is utilized, we recommend the model be updated with the results of the FSS samples to confirm underlying spatial distribution.

It is worth noting that the use of GEM+ framework implies a spatial autocorrelation structure is at play which would affect the confidence of the hypothesis test per Griffith (2005). Stakeholders may wish to factor autocorrelation and the effective sample size in order to obtain a more accurate confidence level. Future revisions of this guidance will include additional guidance on this topic.

9.2.2.2 Workflow

The following are the major steps in conducting a GEM+ survey design and compliance evaluation.

1. Estimate number of core holes for Sign/WRS Test $DCGLV_w^{(d)}$ evaluation.
2. Locate core holes with either a conservative or graded core hole design.
3. Collect core hole samples per the conservative or graded design.
4. Apply hypothesis testing for each decision layer $DCGLV_w^{(d)}$
5. Apply unity rule to sample results as required for multiple radionuclides.

9.3 Virtual Scanning

9.3.1 Overview

In cases with locally elevated values of a concern $DCGLV_{local}^{(d)}$, licensees should consider the combined workflow here. While the task may appear at first overwhelming, virtual scanning of multiple objectives is easily achieved through automated virtual scans (Stewart, 2011). In practice, the set $DCGLV_{scan}^{(d)}$ may only be comprised of a small subset of these objectives depending on site specific needs and exposure concerns. Regardless of the number of objectives, an automated virtual scan by software against the set $DCGLV_{scan}^{(d)}$ can yield two core products: (1) a tabular set of summaries per scan objective and (2) a map per objective of where scan issues exist. Table 9.1 illustrates what a summary table will look like. The scan objective column shows the kind of scan, the radionuclide column holds the radionuclide name,

Rad ID refers to that rad's i index, Depth is the depth layer ID, and the Scan value is the $DCGLV_{local}$ in picocuries per gram (pCi/g) or becquerels per gram (Bq/g). The results of the scan are in the last two columns. The $P(failure)$ indicates the virtual scan prediction that a true but unknown scan would fail to pass given the sampling available and the spatial uncertainty that exists between samples.

Table 9.1. Virtual scan results based on geostatistical simulation.

Scan objective	Radionuclide	Rad ID	Depth	Scan value	P(failure)
$DCGLV_{EMC,i}^{(d)}$	Cesium-137	1	2	5	2%
$DCGLV_{EMC,i}^{(d)}$	Cesium-137	1	3	8	1%
...
$DCGLV_{gw,i}^{(d)}$	Cobalt-60	1	4	3	20%
...
$DCGLV_{nte,i}^{(d)}$	Cesium-137	3	1–10	12	5%
...

Software (Stewart 2011) can be used to automatically generate these tables given a geostatistical model and a set of $DCGLV$ values. For each row, VSP can also produce a GSM map of where in the SSU virtual scanning identifies an issue (Figure 9.3).

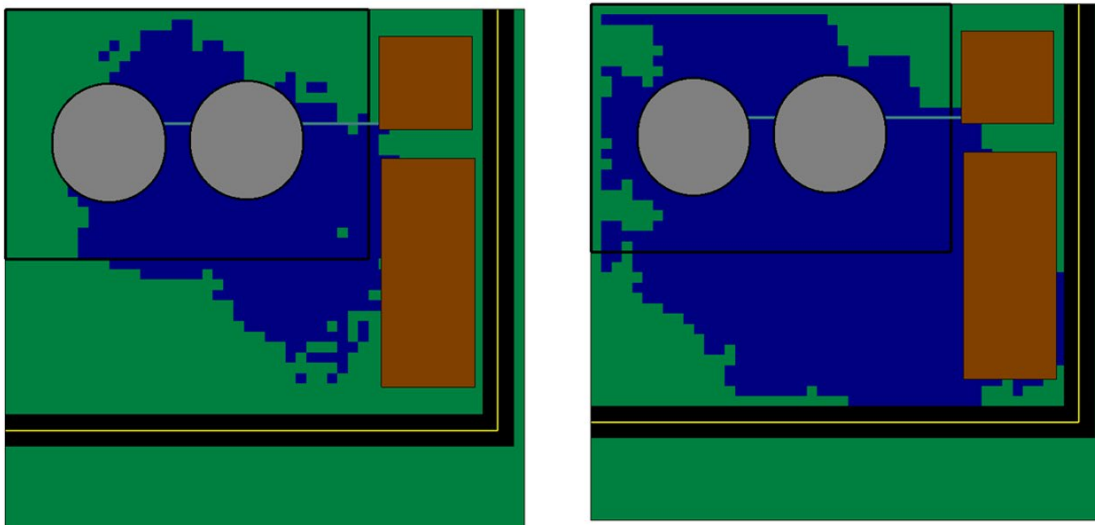


Figure 9.3. Map of two different $DCGLV_{local}$ scan criteria over the same place. Areas in violation are colored in blue. Taken from Stewart 2011.

Preparation for virtual scanning is found throughout the preceding chapters. Stakeholders who follow the GEM+ workflow should arrive at the FSS stage with some virtual scan readiness pre-checks already in place (see below).

1708 9.3.2 Sample Design

1709 Samples collected for hypothesis testing will also benefit the geostatistical model and
1710 corresponding virtual scan. However, the model may benefit from additional model-based
1711 sample design data as well. For example, an uncertainty may exist about the boundary between
1712 higher and lower values, adequate spatial coverage, or decision stability experienced during
1713 readiness checks. Several secondary or model-based sampling designs such as check and
1714 cover, AOC boundary design, high value, high variance, adaptive fill, and others (see Appendix
1715 A) may add additional cores to the design in support of the virtual scan. Among these, check
1716 and cover stands out as a viable FSS option owing to its balance between the goals of checking
1717 value areas while improving spatial coverage. A metric is available that indicates the number of
1718 supplementary core holes samples stakeholders may add for a combined sampling campaign.
1719 These additional samples are biased and should not be used in the independent hypothesis
1720 test. However, both hypothesis and biased design results should be used in the geostatistical
1721 model update where fine scale modeling mitigates clustering bias found in the design.

1722 9.3.3 Evaluation

1723 Data collected from both the hypothesis and virtual scan sampling designs are used to update
1724 the geostatistical model. The virtual scan is applied (Section 3.3) to determine the number of
1725 failures tabulated per layer and $DCGLV_{local,i}^{(d)}$. In this case, stakeholders will establish a
1726 maximum tolerance for the probability of failure and decide whether to move forward with
1727 compliance arguments or return to characterization and remediation.

1728 9.3.4 Workflow

1729 The following are the major steps in conducting a GEM+ survey design and compliance
1730 evaluation.

- 1731 1. Estimate number of additional model-based core holes beneficial for $DCGLV_{local,i}^{(d)}$.
- 1732 2. Locate and collect model-based core holes according to model-based design
- 1733 3. Update the model with both hypotheses testing and model-based samples
- 1734 4. Apply the Virtual scan against $DCGLV_{local,i}^{(d)}$ collection
- 1735 5. Apply the GEM+ Unity Rule Workflow Scan to assess cumulative risks
- 1736 6. Apply result of practical equivalency (ROPE) analysis to assess model stability

1737 9.4 Result of Practical Equivalency Stability Assessment

1738 9.4.1 Overview

1739 The third objective is to assess the stability of the geostatistical model by examining the stability
1740 of the virtual scan results as FSS data update the pre-FSS model into a post-FSS model. The
1741 ROPE assessment assesses how much the post-FSS virtual scan moved from the pre-FSS
1742 readiness scan in terms of shifting the probability of failure. If the pre- and post-FSS scans are
1743 comparatively equivalent in terms of the decision made on them, then the model is suitably
1744 stable. If the probability of failures or the scan issue maps shift dramatically, this may indicate
1745 the model may be inadequate for demonstrating compliance. More characterization sampling
1746 may be needed to be explored to suitable stabilize the model (Section 7.2).

1747 **9.4.2 Correlation Model Stability**

1748 A core component of geostatistical modeling is the correlation model which describes the
1749 underlying spatial structure. Various models are available including the semi-variogram and the
1750 correlogram. Parameters include range, sill, nugget, rotation parameters for anisotropic
1751 conditions. Stability in the correlation structure can be assessed through a practical comparison
1752 of the pre-FSS model against the experimental variogram values calculated with the full post-
1753 FSS data set (including pre-FSS samples from previous phases). In most cases small
1754 adjustments to the semi-variogram to reflect post-FSS data will have little effect on the final
1755 decision.

1756 **9.4.3 Geostatistical Model FSS Predictions**

1757 With the FSS values in hand, it is straightforward to assess how well the pre-FSS model
1758 predicted the true FSS sample values. Reasonably close predictions indicate a model that is
1759 capturing the important trends and is likely a stable model representation of subsurface
1760 conditions. What is considered “close” will depend on the survey unit circumstances.
1761 Reasonable metrics include assessing how many true values were within one or two standard
1762 deviations of the point prediction (e.g., using kriging variance or true simulation variance). In
1763 cases where fine scale screening values (essentially point screens) are relevant, assessment of
1764 how many times a change from “above screen” or “below screen occurred.”

1765 **9.4.4 Geostatistical Model Differencing**

1766 Another useful stability comparison is to evaluate the difference in expected values between the
1767 pre- and post-FSS model. Consider first producing an averaging over simulations to produce a
1768 single estimation map for both the pre- and post-FSS results. Subtraction of these two and
1769 comparison of resulting difference maps and summary statistics for the differences can help
1770 assess if the models are relatively comparable. The more data that is available the more precise
1771 the maps become so differences are certain to exist. What is of interest is trends that indicate
1772 fundamentally different processes than expected from the pre-FSS studies. Examples include
1773 systematic and significant shifting of values to higher or lower values.

1774 **9.4.5 Decision Stability**

1775 A driving decision for model stability is to reapply the virtual scan to the post-FSS geostatistical
1776 model and compare the results with the pre-FSS scan results. Violations of the post-FSS virtual
1777 scan would have already signaled an issue. However, even for successful post-FSS scans, it is
1778 worth examining if there were fundamental shifts in the decision metrics. There is certain to be
1779 slight variations in the scan probabilities across the $DCGLV_{local}$ values and this is normal.
1780 However, systematic and significant increases in scan probabilities across all $DCGLV_{scan}$ values
1781 could suggest systemic underestimation of survey volume sufficient to reverse a compliant
1782 decision if more samples were taken.

1783 If regulators agree that the confirmatory GSM and the pre-confirmatory GSM are ROPE, that
1784 FSS samples create no additional concern individually, and that regulatory review assess
1785 confidence application of the GEM+ workflow, then compliance is indicated.

1786

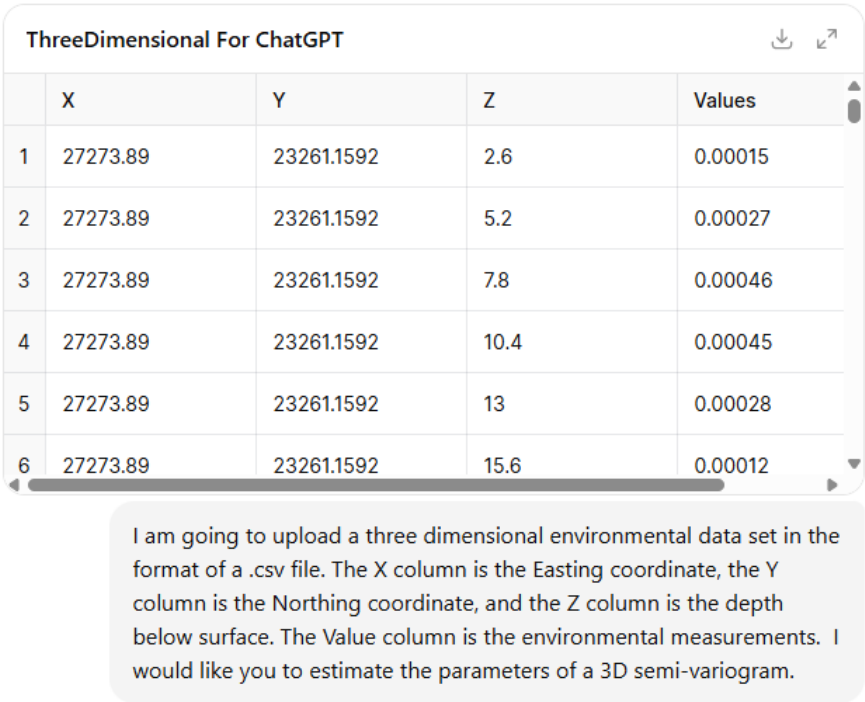
10 AI and Large Language Models

1787
1788
1789
1790
1791
1792
1793

Future editions of this document should consider recent breakthroughs in Large Language Models (LLMs) or more broadly generative AI. LLMs in particular have had an extraordinary and disruptive impact on industry, government, and academia. At the time of this writing, these sectors are working quickly to understand these phenomena and how to ensure their reliable and proper use within society. As of August 24, 2025, we provide a light and informal demonstration of where LLMs are at the time of this writing and the potential they have in aiding compliance efforts and geostatistical modeling in particular.

1794
1795

As an example of the current state of LLMs with respect to the problem of fitting correlation structures, we posed a prompt to ChatGPT 5.0 on August 24, 2025 (Figure 10.1).



1796
1797

Figure 10.1. Screenshot of ChatGPT prompt.

1798
1799
1800

After about four minutes, ChatGPT returned with a table of estimated spherical model parameters (Table 10.1), a table of lag and semi-variance values (Table 10.2) and an image of the best fit semi-variogram model (Figure 10.1).

1801

Table 10.1. ChatGPT model selection and parameterization.

Model	Spherical
Nugget	1.26×10^{-3}
Partial Sill	1.49×10^{-2}
Total Sill	1.62×10^{-2}
Range Parameter (a)	97.7
Practical Range (95%)	97.7
Weighted SSE	2.40
Pairs Used (sum over bins)	1.03×10^5
Bins Used	14

1802

1803

Table 10.2. ChatGPT lag distance and semivariance backing information.

Lag distance (h)	Semivariance $\hat{\gamma}(h)$	Pairs in bin
37.70	9.34×10^{-3}	3840
188.49	2.06×10^{-2}	4128
263.89	1.74×10^{-2}	7813
339.28	1.56×10^{-2}	6995
414.68	2.12×10^{-2}	8371
490.07	2.12×10^{-2}	6692
565.47	1.83×10^{-2}	10,382
640.87	9.79×10^{-3}	8540
716.26	1.93×10^{-2}	10,411
791.66	1.03×10^{-2}	9156
867.05	8.44×10^{-3}	6362
942.45	2.10×10^{-2}	10,079
1017.85	6.53×10^{-3}	5213
1093.24	1.77×10^{-2}	5061

1804

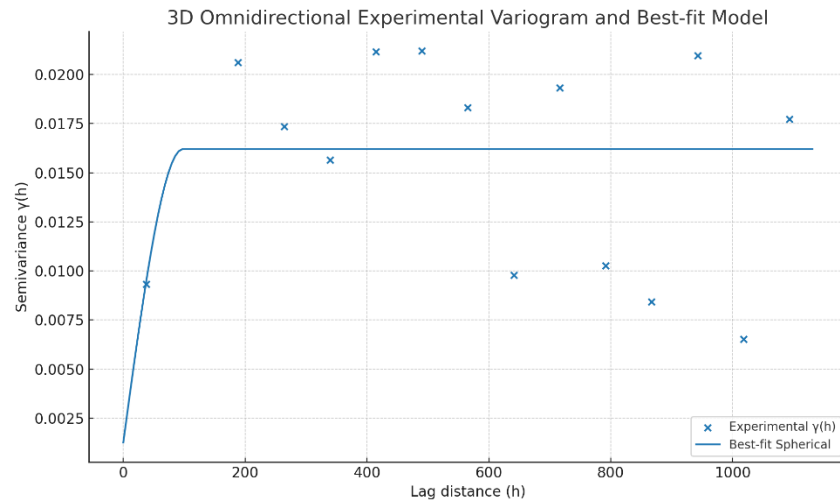


Figure 10.2. ChatGPT graphic visualization of best-fit spherical model.

The following is a script of the remaining response offered by ChatGPT 5.0.

“Thought for 2m 37s

*Done. I loaded your CSV, built a 3-D **omnidirectional** experimental semi variogram (using Euclidean distance over X,Y,Z), and fit standard models by weighted least squares (coarse-to-fine grid search across nugget/sill/range). I mapped your “**Values**” → “**Value**” column automatically.*

I just opened two tables for you:

Variogram Parameters (Estimated) — the best-fit model and parameter values (nugget, partial/total sill, range parameter, practical range, SSE).

Experimental Variogram Points — lag centers, $\gamma(h)$, and pair counts.

You’ll also see a plot of the experimental points with the fitted curve.

Files you can grab:

Download parameters JSON

Download experimental variogram points CSV

Download plot PNG

What I did (brief)

Used all 3D pairs (randomly subsampled if needed) and limited the tail to the 90th percentile of pairwise distances to avoid sparse bins; 15 lag bins; kept bins with ≥ 30 pairs.

*Fit **Spherical**, **Exponential**, and **Gaussian** models; selected the one with the lowest weighted SSE, then refined around the best parameters.*

*Reported **nugget**, **partial sill**, **total sill**, **range parameter (a)**, and **practical range** (95% of sill: a for spherical, $\sim 3a$ for exponential, $\sim \sqrt{3} \cdot a$ for Gaussian).*

1830
1831
1832
1833

Notes / next options
*This is an **omnidirectional** fit. If you want to check **anisotropy** (horizontal vs. vertical or azimuthal), I can compute directional variograms (set azimuth/dip tolerances) and estimate anisotropy ratios and rotated ranges.*

1834
1835

*If X/Y and Z are in **different units** (e.g., meters vs. feet), we should rescale Z before fitting; I treated all axes in the same unit.*

1836
1837
1838

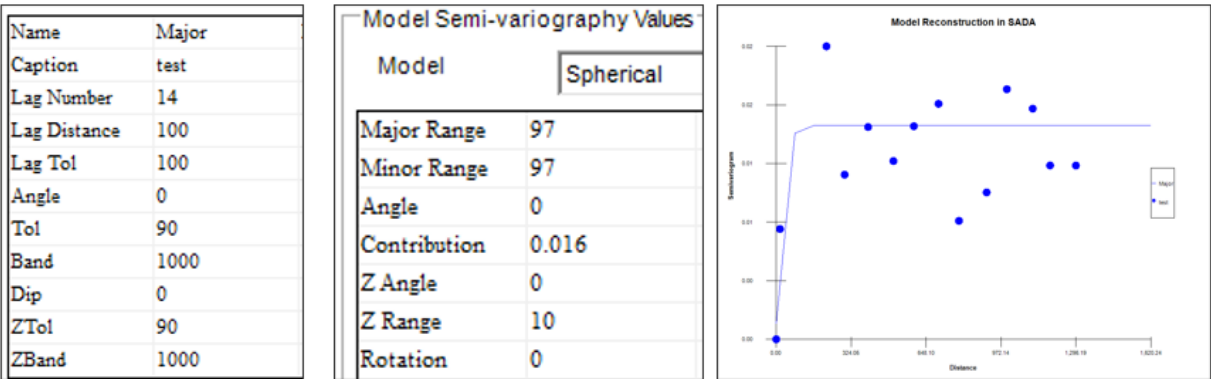
We can also add nested structures, try robust estimators (e.g., Cressie–Hawkins), de-trend, or transform values (e.g., log) if the residuals look nonstationary.

1839
1840

If you want anisotropy or nested structures, just say the directions/tolerances you care about and I'll run those fits on this same data.” – ChatGPT 5

1841
1842
1843
1844
1845
1846
1847
1848

While this is an informal example, and drawing any conclusions about reliability is too soon, it is difficult to ignore some remarkable features of the exchange. First, ChatGPT understood what a semi-variogram was, how to estimate it, that multiple models were available to fit and how to choose them. It recognized that anisotropy might be at play and inquired about deeper analysis. Finally, it understood to draw a model to supply data back to the prompter and understood to provide some explanation for its actions. Using SADA 5.0 to explore the results, the same environmental file was loaded into SADA and with the parameters approximating ChatGPT 5.0 response as close as possible. The results are shown in Figure 10.3.



1849
1850
1851

(a) (b) (c)
Figure 10.3. Approximation of ChatGPT 5.0 omnidirectional correlation model parameters in SADA 5.0.

1852
1853
1854
1855

While this report does not provide guidance on LLM use at this time it is a capability well worth experimenting with to support a subject matter expert in quickly building a geostatistical model. Future versions of this report should consider formal evaluation of this new disruptive capability.

1856

11 SUMMARY

1857
1858
1859

This revision of NUREG/CR-7021 advances the geospatial approach to subsurface compliance by extending the clear and defensible principles of MARSSIM into environments that are more complex and less accessible. The GEM+ workflow builds on the foundations of the first edition,

1860 the Triad model, and subsequent geostatistical research, while also incorporating insights from
1861 NRC workshops, public comments, and related technical guidance. The result is a structured
1862 but flexible framework that is scientifically rigorous, transparent, and capable of adapting to a
1863 wide range of subsurface conditions.

1864 Subsurface investigations present persistent challenges. Unlike surface media, subsurface
1865 contamination cannot be scanned comprehensively, and sampling costs often preclude the
1866 densities needed to drive uncertainty to negligible levels. Contamination may be distributed
1867 heterogeneously across three dimensions and affected by excavation, hydrogeology, and
1868 engineered structures. GEM+ addresses these challenges by repositioning geostatistical
1869 modeling at the center of decision support. Probabilistic simulations, kriging, and hybrid
1870 inference methods provide a means to quantify both expected conditions and the uncertainty
1871 that surrounds them, allowing compliance decisions to be based on the totality of evidence
1872 rather than isolated measurements.

1873 A major improvement in this revision is the hybrid FSS. The GEM+ FSS combines statistically
1874 selected soil samples with geostatistical virtual scans of the subsurface to produce multiple lines
1875 of evidence. This dual approach strengthens confidence in compliance outcomes and provides
1876 a structured path for assessing multiple DCGLv values that may vary by depth, radionuclide, or
1877 exposure pathway. By incorporating approaches such as the Multi-Scale Remedial Design
1878 Model (MrDM) and Multi-Scale Remedial Sample Design Model (MrsDM), GEM+ also provides
1879 adaptive tools for designing efficient remedial and sampling strategies that are responsive to
1880 site-specific complexity.

1881 The guidance presented here acknowledges the continuing challenge in maintaining scientific
1882 rigor and achieving practical feasibility. For simple contamination problems, conservative
1883 excavation and removal strategies may remain the most appropriate and cost-effective path to
1884 compliance. For more complex cases, GEM+ offers a graded approach that integrates
1885 advanced modeling into each phase of investigation, remediation, and final status assessment.
1886 In all cases, the framework emphasizes transparent reasoning, cumulative evidence, and
1887 reproducibility.

1888 This revision is not intended as a final word but as a significant step in an ongoing process.
1889 Advances in computing, geophysical sensing, real-time analysis, and visualization will continue
1890 to expand the scope of what can be achieved. In particular, LLMs or similar generative AI
1891 capabilities are certain to have an impact. Likewise, regulatory policy will continue to evolve in
1892 response to new lessons learned, stakeholder input, and changing risk landscapes. GEM+ has
1893 been deliberately structured to remain adaptable to these developments

1894 In conclusion, GEM+ provides a more complete, integrated, and scientifically defensible
1895 pathway for demonstrating compliance with NRC's License Termination Rule in subsurface
1896 settings. It unites the proven strengths of MARSSIM with modern geospatial methods,
1897 strengthens confidence in decision outcomes, and offers flexibility to match the scale and
1898 complexity of each site. As more complex sites move into decommissioning, GEM+ equips
1899 regulators, licensees, and stakeholders with a forward-looking framework that balances the
1900 need for technical rigor with the realities of cost, uncertainty, and public trust.

1901

1902

12 REFERENCES

- 1903 Ahmed, Shakeel, Dewashish Kumar, et al. (2008). Application of Geostatistics in Optimal
1904 Groundwater Monitoring Network Design. *Groundwater Dynamics in Hard Rock*
1905 *Aquifers*: 179-190.
- 1906 Back, Par-Erik (2006). *Value of Information Analysis for Site Investigations in Remediation*
1907 *Projects*, Dissertation. Chalmers University of Technology, Sweden.
- 1908 Back, Par-Erik (2007). A model for estimating the value of sampling programs and the optimal
1909 number of samples for contaminated soil. *Environmental Geology*. 52:573–585.
- 1910 Baraba`s, N., P. Goovaerts, and P. Adriaens. (2001). Geostatistical assessment and validation
1911 of uncertainty for three-dimensional dioxin data from sediments in an estuarine river.
1912 *Environmental Science and Technology*. 35:3294–3301.
- 1913 Bolstad, P. GIS Fundamentals: A First Text on Geographic Information Systems, 6th Edition,
1914 (2019), XanEdu Publishing, Inc, ISBN-13: 978-1593995522.
- 1915 Brakewood, Liv H. and Domenico Grasso. (2000). Floating Spatial Domain Averaging in Surface
1916 Soil Remediation. *Environmental Science and Technology* 34(18); 3837–3842
- 1917 Brockett, P. and A. Levine. (1984). *Statistics and Probability*. New York, NY, CBS College
1918 Publishing.
- 1919 Burke, R. (2003). *Getting to Know ArcObjects*. Redlands, CA, ESRI Press.
- 1920 Butt, T. E., E. Lockley, et al. (2007). Risk Assessment of landfill disposal sites – State of the art.
1921 *Waste Management* 28(6): 952-964.
- 1922 Byrd, D. and R. Cothorn. (2000). *Introduction to Risk Analysis*. Rockville, MD, Government
1923 Institutes.
- 1924 Carlon, C., L. Pizzol, et al. (2008). A spatial risk assessment methodology to support the
1925 remediation of contaminated land. *Environment International* 34(3): 397-411.
- 1926 Cember, Herman and Thomas Johnson. (2009). *Introduction to Health Physics*, McGraw-Hill.
- 1927 Coleman, R. (2009). Uncertainty Associated with On-Site Analysis of Unprepared Soil Samples
1928 Containing Cs-137 Using High Resolution Gamma Spectrometry. Knoxville, TN, A
1929 *Report Prepared for the University of Tennessee in Support of Nuclear Regulatory*
1930 *Commission Contract Number NRC-06-066*, "Sparse Radiological Subsurface Survey
1931 *Data Analysis*." 2006-2009.
- 1932 Crumbling, Deana (2001a). Current perspectives in site remediation and monitoring: Using the
1933 Triad Approach to improve the cost-effectiveness of hazardous waste site cleanups.
1934 EPA 542-R-01-016. Office of Solid Waste and Emergency Response. Washington, D.C.
- 1935 Crumbling, D. (2001b). Using the TRIAD approach to improve the cost-effectiveness of
1936 hazardous waste site cleanups. EPA 542-R-01-016, Washington, DC.

1937 Crumbling, D. (2002). In Search of Representativeness: Evolving the Environmental Data
1938 Quality Model. *Quality Assurance: Good Practice, Regulation, and Law* 9(3): 179 - 190.

1939 Crumbling, D. (2004). Summary of the TRIAD Approach. Washington, D.C., Office of Superfund
1940 Remediation and Technology Innovation, United States Environmental Protection
1941 Agency.

1942 Crumbling, D. M. Griffith, J., Powell, D. (2003). Improving decision quality: Making the case for
1943 adopting next-generation site characterization practices. *Remediation Journal* 13(2): 91-
1944 111.

1945 D'Or, D. (2005). Towards a real-time multi-phase sampling strategy optimization. In
1946 *Geostatistics for Environmental Applications*. P. Renard. (ed). Berlin, Heidelberg,
1947 Springer-Verlag: 355-366.

1948 Dakins, Maxine E., John E. Toll, et al. (1996). Risk-Based Environmental Remediation:
1949 Bayesian Monte Carlo Analysis and the Expected Value of Sample Information. *Risk*
1950 *Analysis* 16(1): 67-79.

1951 Davidson, JR. 1994. ELIPGRID-PC: A PC Program for Calculating Hot Spot Probabilities.
1952 ORNL/TM-12774. Oak Ridge National Laboratory, Oak Ridge, Tennessee.

1953 Davidson, JR. 1995a. Monte Carlo Tests of the ELIPGRID-PC Algorithm. ORNL/TM-12899. Oak
1954 Ridge National Laboratory, Oak Ridge, Tennessee.

1955 Davidson, JR. 1995b. ELIPGRID-PC: Upgraded Version. ORNL/TM-13103. Oak Ridge National
1956 Laboratory, Oak Ridge, Tennessee.

1957 Deutsch, C.V. and A.G. Journel (1992). *GSLIB: Geostatistical Software Library and User's*
1958 *Guide*. New York, Oxford University Press.

1959 Delmelle, E.M.. and P. Goovaerts (2009). Second-phase sampling designs for non-stationary
1960 spatial variables. *Geoderma* 153(1-2): 205-216.

1961 Demougeot-Renard, H., Chantal de Fouquet, et al. (2004). "Forecasting the Number of Soil
1962 Samples Required to Reduce Remediation Cost Uncertainty." *J Environmental Quarterly*
1963 33(5): 1694-1702.

1964 Doll, William, Gamey, T., Watson, David, and Jardine, Phil. (2002). Geophysical Profiling in
1965 Support of a Nitrate and Uranium Groundwater Remediation Study. 10.4133/1.2927099.

1966 Eckerman, K. and Jeffrey R. (1993). External Exposure to Radionuclides in Air, Water, and Soil
1967 (FGR12). Washington, D.C., Office of Radiation and Indoor Air, United States
1968 Environmental Protection Agency. EPA 402-R-93-081.

1969 Emery, Xavier (2008). Statistical tests for validating geostatistical simulation algorithms.
1970 *Computers & Geosciences* 34(11): 1610-1620.

1971 England, Evan and Naser Heravi (1992). "Conditional Simulation: Practical Application for
1972 Sampling Design Optimization." In *Geostatistics Troia*, ed A.Soares. Kluwer Academic
1973 Publ., Dordrecht, the Netherlands

- 1974 Englund, E., and N. Heravi. (1993). Conditional simulation: Practical application for sampling
1975 design optimisation. p. 613–624. In *Geostatistics Troia*, ed. A. Soares. Kluwer Academic
1976 Publ., Dordrecht, the Netherlands.
- 1977 Englund, E., and N. Heravi. (1994). Phased sampling for soil remediation. *Environmental*
1978 *Ecology. Stat.* 1:247–263.
- 1979 EPRI (2016). Guidance for Using Geostatistics in Developing a Site Final Status Survey
1980 Program for Plant Decommissioning.
- 1981 Franklin, R.B., Mills, A.L., (2003). Multi-scale spatial heterogeneity for microbial community
1982 structure in an eastern Virginia agricultural field. *FEMS Microbiology Ecology* 44:335-
1983 346.
- 1984 Freeze, R. A., J. Bruce, J. Massman, T. Sperling, and L. Smith. (1992). Hydrogeological
1985 decision analysis. 4. The concept of data worth and its use in the development of site
1986 investigation strategies. *Ground Water* 30:574–588.
- 1987 Gass, Saul. (1985). *Linear Programming Methods and Applications*. New York, NY, McGraw-Hill
1988 Book Company.
- 1989 Gilbert, Richard (1987). *Statistical Methods for Environmental Pollution Monitoring*. New York,
1990 NY, Van Nostrand Reinhold.
- 1991 Gogolak, C. C Wiblin, and S Bland (2022) Guidance on Surveys for Subsurface Radiological
1992 Contaminants White Paper. Last Accessed at
1993 <https://adamswebsearch2.nrc.gov/webSearch2/main.jsp?AccessionNumber=ML22088A>
1994 [219](https://adamswebsearch2.nrc.gov/webSearch2/main.jsp?AccessionNumber=ML22088A)
- 1995 Goovaerts, P. (1997). *Geostatistics for Natural Resource Evaluation*. New York, Oxford
1996 University Press.
- 1997 Goovaerts, P. (1999). Geostatistics in soil science: state-of-the-art and perspectives. *Geoderma*
1998 89(1-2): 1-45.
- 1999 Goovaerts, P. (2001). Geostatistical modelling of uncertainty in soil science. *Geoderma* 103(1-
2000 2): 3-26.
- 2001 Gorr, W and K Kurland, GIS Tutorial for ArcGIS Pro 3.1, (2023) ESRI Press, ISBN:
2002 9781589487390.
- 2003 Gosling, J. (2017) SHELF: The Sheffield Elicitation Framework
- 2004 Groenigen, J. W., W. Siderius, et al. (1999). Constrained optimisation of soil sampling for
2005 minimisation of the kriging variance. *Geoderma* 87(3-4): 239-259.
- 2006 Griffith, Daniel A. (2005). Effective Geographic Sample Size in the Presence of Spatial
2007 Autocorrelation. *Annals of the Association of American Geographers* 95(4): 740-760.

2008 Holland, J., Bartell, S.M., Hallam, T.G., Purucker, S.T., Welsh, C.J.E., (2003). Role of
2009 Computational Toolkits in *Environmental Management. Ecological modeling for resource*
2010 *management*. Ed. V. H. Dale. New York, Springer-Verlag.

2011 Hortal, J., M. B. Araújo, et al. (2009). Testing the effectiveness of discrete and continuous
2012 environmental diversity as a surrogate for species diversity. "*Ecological Indicators* 9(1):
2013 138-149.

2014 Hockett, J., D Fagan, Z Weller, M. Obiri, L Newburn, F Day-Lewis and D Peeler (2022)
2015 Subsurface Radiological Survey Design and Geospatial Analysis Tool
2016 Recommendations, PNNL-33647.

2017 Isaaks, E., Srivastava, R.M. (1989). *An Introduction to Applied Geostatistics*. New York, Oxford
2018 University Press.

2019 ITRC (2003). *Technical and Regulatory Guidance for the Triad Approach: A New Paradigm for*
2020 *Environmental Project Management*. Washington, D.C., Interstate Technology and
2021 Regulatory Council.

2022 ITRC (2008). *Decontamination and Decommissioning of Radiologically Contaminated Facilities*.
2023 Washington, DC, Interstate Technology & Regulatory Council.

2024 ISO (2017), Characterization principles for soils, buildings, and infrastructures contaminated by
2025 radionuclides for remediation purposes. ISO 18557:2017(E).

2026 James, B. R. and S. M. Gorelick. (1994). "When enough is enough: the worth of monitoring data
2027 in aquifer remediation design." *Water Resource Res.* 30:3499–3513.

2028 Johnson, R. (1996). A Bayesian/Geostatistical Approach to the Design of Adaptive Sampling
2029 Programs. In *Geostatistics for Environmental and Geotechnical Applications*. R.
2030 Srivastava . et. al. (eds). Philadelphia, PA, American Society for Testing and Materials.

2031 Johnson, Robert, John Quinn, et al. (1997). Adaptive sampling and analysis programs for
2032 contaminated soils. *Remediation Journal* 7(3): 81-96.

2033 Johnson, Robert, Karen Smith, et al. (1999). *The Application of Adaptive Sampling and Analysis*
2034 *Program (ASAP) Techniques to Norm Sites*. Argonne National Laboratory
2035 Environmental Assessment Division Argonne, Illinois. DOE/BC/W-31-109-ENG-38-9.

2036 Johnson, Robert, David LePoire, et al. (2005). Bayesian Approaches for Adaptive Spatial
2037 Sampling: An Example Application ANL/EAD/TM-05-01. Argonne, Illinois, prepared by
2038 the Environmental Assessment Division, Argonne National Laboratory.

2039 Johnson, Robert and Lisa Durham (2004). Triad Case Study: Rattlesnake Creek. *Remediation:*
2040 *The Journal of Environmental Cleanup Costs, Technologies and Techniques* 15(1).

2041 Juang, Kai-Wei, Wan-Jiun Liao, et al. (2008). Additional sampling based on regulation threshold
2042 and kriging variance to reduce the probability of false delineation in a contaminated site.
2043 *Science of The Total Environment* 389(1): 20-28.

2044 Kahneman, D. (2011) Thinking Fast and Slow, Farrar, Straus, and Giroux.

- 2045 Kahnmean, D, Sibony O, and Sunstein C (2021) Noise: A Flaw in Human Judgment. Little,
2046 Brown Spark; Hachette Book Group.
- 2047 Lark, R. M. (2002). Optimized spatial sampling of soil for estimation of the variogram by
2048 maximum likelihood. *Geoderma* 105(1-2): 49-80.
- 2049 Linkov, I., A. Varghese, et al. (2004). *Comparative Risk Assessment and Environmental*
2050 *Decision Making*. Springer Netherlands: 15-54.
- 2051 Lively (2012). "Method for Implementing Subsurface Solid Derived Concentration Guideline
2052 Levels (DCGL)—12331," WM2012 Conference, February 26–March 1, 2012, Phoenix,
2053 AZ.
- 2054 Lyon, B.F., S.T. Purucker, et al. (1994). The Value of Perfect Information: How Much is a
2055 Crystal Ball Worth? International Specialty Conference, Pittsburgh, PA, Air & Waste
2056 Management Association.
- 2057 Matzke, B.D., L Newburn, J Hathaway, L Bramer, J Wilson, S Dowson, L Sego, B Pulsipher
2058 (2014), Visual Sample Plan Version 7.0 User's Guide, PNNL-23211.
- 2059 McConnell, Steve. (1993). *Code Complete: A practical handbook of software construction*.
2060 Redmond, WA, Microsoft Press.
- 2061 McNulty, G., B. Deshler, and H. Dove. IT Corporation. (1997). Value of Information Analysis;
2062 Nevada Test Site. CONF-970335–37, US Department of Energy, Washington, DC,. 12.
2063 pp.
- 2064 Meyer, P. D. and E. D. Brill, Jr. (1988). A Method for Locating Wells in a Groundwater
2065 Monitoring Network Under Conditions of Uncertainty. *Water Resource Res.* 24.
- 2066 Microsoft Developers Network. (2011). Parallel Programming in the .NET Framework.
2067 Microsoft.com. (<http://msdn.microsoft.com/en-us/library/dd460693.aspx>) Accessed on 15
2068 April 2011.
- 2069 Miller, Harvey and Shih-Lung Shaw (2001). *Geographic Information Systems for Transportation*.
2070 London, Oxford University Press.
- 2071 Miller, Irwin, John Freund and Richard Johnson. (1990). *Probability and Statistics for Engineers*,
2072 4th Edition. Englewood Cliffs, NJ, Prentice Hall.
- 2073 National Research Council (2007). *Models in Environmental Regulatory Decision Making*.
2074 Washington, D.C., The National Academies Press.
- 2075 Norberg, T. and L. Rosén (2006). Calculating the optimal number of contaminant samples by
2076 means of data worth analysis. *Environmetrics* 17:705–719.
- 2077 Norrman, J., Purucker, S.T., Stewart, R.N., Back, P.-E., Englelke, F., (2008). Framework for
2078 optimizing the evaluation of data from contaminated soil in Sweden. Proceedings of
2079 ConSoil 2008, 10th International Conference on Soil-Water Systems; Milan, Italy.

2080 Ostresh, Lawrence M. (1977). On the Convergence of a Class of Iterative Methods for Solving
2081 the Weber Location Model. *Operations Research* 26(4):597-609.

2082 O'Sullivan, David and David Unwin. (2003). *Geographic Information Analysis*. Hoboken, NJ,
2083 John Wiley and Sons, Inc.

2084 Puckett, G. and T. Shaw (2004). Triad case study: Former small arms training range.
2085 *Remediation Journal* 15(1): 79-87.

2086 Purucker, S. T., C. J. E. Welsh, et al. (2007). Use of habitat-contamination spatial correlation to
2087 determine when to perform a spatially explicit ecological risk assessment. *Ecological*
2088 *Modelling* 204(1-2): 180-192.

2089 Rügner, H., M. Finkel, et al. (2006). Application of monitored natural attenuation in contaminated
2090 land management--A review and recommended approach for Europe. *Environmental*
2091 *Science & Policy* 9(6): 568-576.

2092 Saito, Hirotaka and Pierre Goovaerts (2003). Selective Remediation Of Contaminated Sites
2093 Using a Two-Level Multiphase Strategy and Geostatistics. *Environmental Science and*
2094 *Technology* 37.

2095 Savelieva, E., Demyanov, V., Kanevski, M., Serre, M.L., Christakos, G., (2005). BME-based
2096 uncertainty assessment of the Chernobyl fallout. *Geoderma* 128(3-4): 312-324.

2097 Simbahan, Gregorio C. and Achim Dobermann (2006). Sampling optimization based on
2098 secondary information and its utilization in soil carbon mapping. *Geoderma* 133(3-4):
2099 345-362.

2100 Sinha, P., Lambert, M.B., Schew, W.A. (2007). Evaluation of a risk-based environmental hotspot
2101 delineation algorithm. *Journal of Hazardous Materials* 149.

2102 Stewart, R.N., Welsh, C., Purucker, T., and Stewart, D., (2009), An Introduction to Spatial
2103 Analysis and Decision Assistance (SADA): Environmental Applications for Version 5
2104 User Guide, www.sadaproject.net.

2105 Stewart, R.N., Purucker, S.T., (2006). SADA: A Freeware Decision Support Tool Integrating
2106 GIS, Sample design, Spatial Modeling, and Risk Assessment. Proceedings of the Third
2107 Biennial Meeting of the International Environmental Modelling and Software Society,
2108 Burlington, Vermont.

2109 Stewart, R.N., Purucker, S.T. Welsh, C. (2009). *An Introduction To Spatial Analysis and*
2110 *Decision Assistance Environmental Applications for Version 5 User Guide*. Washington,
2111 D.C., United States Nuclear Regulatory Commission.

2112 Stewart, R. N., Purucker, S.T. (2011). An Environmental Decision Support System for Spatial
2113 Assessment and Selective Remediation. *Environmental Modelling & Software*.

2114 Stewart, R.N. (2011), "A Geospatial Based Decision Framework for Extending MARSSIM
2115 Regulatory Principles into the Subsurface. " PhD diss., University of Tennessee, 2011.
2116 https://trace.tennessee.edu/utk_graddiss/1130.

2117 Sun, Xian-He and Yong Chen. (2010) Reevaluating Amdahl's law in the multicore era. *Journal*
2118 *of Parallel and Distributed Computing* 70(2).

2119 Sunila, R., E. Laine, et al. (2004). Fuzzy model and kriging for imprecise soil polygon
2120 boundaries. Proceedings of the 12th International Conference on Geoinformatics,
2121 Sweden.

2122 Trochim, William M. (2006) The Research Methods Knowledge Base, 2nd Edition, at URL:
2123 <http://www.socialresearchmethods.net/kb/>.

2124 USDOL (2008). Occupational Health and Safety Administration (OSHA), Sloping and Benching,
2125 Safety and Health Regulations for Construction (Excavations), 1926 Subpart P App B
2126 http://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=10932
2127

2128 USEPA (1980). Interim Guidelines and Specifications for Preparing Quality Assurance Project
2129 Plans. QAMS-005/80. Washington, D.C

2130 USEPA (1987a). Data Quality Objectives for Remedial Response Activities: Example Scenario.
2131 EPA/540/G-87/003. Office of Emergency and Remedial Response. Washington, D.C.

2132 USEPA (1987b). Data Quality Objectives for Remedial Response Activities: Example Scenario.
2133 EPA/540/G-87/004. Office of Emergency and Remedial Response. Washington, D.C.

2134 USEPA (1989a). Risk Assessment Guidance for Superfund: Volume I - Human Health
2135 Evaluation Manual. (Part A). EPA/540/1-89/002. Office of Emergency and Remedial
2136 Response, U.S. Environmental Protection Agency. Washington, D.C.

2137 USEPA (1989b). Methods for Evaluating the Attainment of Cleanup Standards, Volume 1: Soils
2138 and Solid Media. Washington, D.C., Office of Policy, Planning, and Evaluation. EPA-
2139 230/02-89-042.

2140 USEPA (1992a). Guidelines for Exposure Assessment. 600Z-92/001 U.S. Environmental
2141 Protection Agency, Risk Assessment Forum Washington, D.C.

2142 USEPA (1992b). Preparation of Soil Sampling Protocols: Sampling Techniques and Strategies.
2143 EPA/600/R92/128. Office of Research and Development. Washington, DC.

2144 USEPA (1994a). Statistical Methods for Evaluating the Attainment of Cleanup Standards,
2145 Volume 3: Reference Based Standards for Soils and Solid Media, Office of Policy,
2146 Planning, and Evaluation. EPA 230-R-94-004. Washington, D.C.

2147 USEPA (1994b). Guidance for the Data Quality Objectives Process. EPA/600/R-96/055.
2148 Washington, DC. EPA QA/G-4 (EPA/600/R-96/055). Washington, D.C.

2149 USEPA (1994b). An SAB Report: Review of EPA's Approach to Screening for Radioactive
2150 Waste Materials at a Superfund Site in Uniontown, Ohio. Washington, D.C., Prepared by
2151 the ad hoc Industrial Excess Landfill Panel of the Science Advisory Board (SAB). EPA-
2152 SAB-EC-94-010.

2153 USEPA (1997). Exposure Factors Handbook (Final Report). EPA/600/P-95/002F a-c. U.S.
2154 Environmental Protection Agency. Washington, DC

2155 USEPA (2000a). Data Quality Objectives Process for Hazardous Waste Site Investigations.
2156 Washington, D.C., Office of Environmental Information. EPA/600/R-00/007.

2157 USEPA (2000b). Guidance for the Data Quality Objectives Process. EPA QA/G-4 (EPA/600/R-
2158 96/055). Washington, D.C.

2159 USEPA (2001a). Resources for Strategic Investigation and Monitoring, Office of Solid Waste
2160 and Emergency Response. EPA 542-F-01-030b.

2161 USEPA (2001b). Data Quality Objectives Decision Error Feasibility Trials Software (DEFT) -
2162 USER'S GUIDE. Washington, D.C., Office of Environmental Information. EPA QA/G-4D.

2163 USEPA (2001c). Requirements for Quality Assurance Project Plans. EPA QA/R-5. Office of
2164 Environmental Information. Washington, D.C.

2165 USEPA (2002a). Calculating Upper Confidence Limits for Exposure Point Concentrations at
2166 Hazardous Waste Sites, 9285.6-10. Office of Emergency and Remedial Response U.S.
2167 Environmental Protection Agency Washington, D.C.

2168 USEPA (2002b). Guidance for Quality Assurance Project Plans for Modeling. Washington, D.C.,
2169 Office of Environmental Information. EPA/240/R-02/007.

2170 USEPA (2002c). Guidance on Choosing a Sampling Design For Environmental Data Collection.
2171 Washington, D.C., US Environmental Protection Agency. EPA QA/G-5S.

2172 USEPA (2003). Using the TRIAD Approach to Streamline Brownsfield Site Assessment and
2173 Cleanup. Washington, D.C., Office of Solid Waste and Emergency Response.

2174 USEPA (2003b). Guidance for Geospatial Data Quality Assurance Project Plans. Washington,
2175 D.C., USEPA. EPA QA/G-5G.

2176 USEPA and Pennsylvania Department of Environmental Protection (2003). STATEMENT OF
2177 WORK: APPLYING THE TRIAD APPROACH TO ADVANCE LAND REUSE MARINO
2178 BROTHERS SCRAP YARD SITE.

2179 USEPA (2005). Use of Dynamic Work Strategies Under a Triad Approach for Site Assessment
2180 and Cleanup – Technology Bulletin September 2005, Office of Solid Waste and
2181 Emergency Response. EPA 542-F-05-008.

2182 USEPA (2005a). Case Study: Use of a Decision Support Tool: Using FIELDS and SADA to
2183 Develop Contour Maps of Contaminant Concentrations and Estimate Removal Volumes
2184 for Cleanup of Soil at the Marino Brothers Scrapyard Site, Rochester Borough,
2185 Pennsylvania. Washington D.C., Office of Superfund Remediation and Technology
2186 Innovation; Brownfields and Land Revitalization Technology Support Center.

2187 USEPA (2006a). Guidance on Systematic Planning Using the Data Quality Objectives Process.
2188 EPA QA/G-4. Office of Environmental Information. Washington, D.C.

2189 USEPA (2006b). Data Quality Assessment: A Reviewer's guide. Washington, D.C., US
2190 Environmental Protection Agency. EPA QA/G-9R.

2191 USEPA (2006c). Data Quality Assessment: Statistical Methods for Practitioners. Washington,
2192 D.C., Office of Environmental Information. EPA QA/G-9S.

2193 USEPA (2006d). Systematic Planning: A Case Study for Hazardous Waste Site Investigations.
2194 Washington, D.C. EPA QA/CS-1.

2195 USEPA (2008). Triad Issue Paper: Using Geophysical Tools to Develop the Conceptual Site
2196 Model. OSWER. Washington, DC, USEPA. 542-F-08-007.

2197 USEPA (2010) Triad Conceptual Site models
2198 http://www.Triadcentral.org/tech/dsp_sub.cfm?id=13 (last accessed 3/18/2010).

2199 USOMB (2007). Final Bulletin for Agency Good Guidance Practices. M-07-07. U.S. Office of
2200 Management and Budget, Section I(3). Washington, D.C.

2201 USNRC (2000). Multi-Agency Radiation Survey and Site Investigation Manual (NUREG-1575),
2202 Office of Regulatory Research. Washington, D.C.

2203 USNRC (2009). Title 10 of the Code of Federal Regulations, Washington, D.C.

2204 USNRC (2021). Proceedings of the Subsurface Soil Surveys Public Workshop July 14-15, 2021.
2205 RIL 2021-12, Rockville MD.

2206 USNRC (2022a). Summary of 2nd Annual Subsurface Investigations Workshop. 2022, Rockville
2207 MD.

2208 USNRC (2022b). Consolidated Decommissioning Guidance, NUREG-1757, Volume 2, Revision
2209 2.

2210 USNRC (2023) Draft Division of Decommissioning, Uranium Recovery, and Waste Programs
2211 Interim Staff Guidance DUWP- ISG-02.

2212 Van Groeningen, J.W., A. Stein, and R. Suurbier. (1997). Optimisation of environmental
2213 sampling using interactive GIS. *Soil Technology*10(2):83–97.

2214 Van Groeningen, J.W., G. Pieters, and A. Stein. (2000). Optimizing sampling for multivariate
2215 contamination in urban areas. *Environmetrics* 11:227–244.

2216 Vasát, R., G. B. M. Heuvelink, et al. (2010). Sampling design optimization for multivariate soil
2217 mapping. *Geoderma* 155(3-4): 147-153.

2218 Verstraete, S. and M. Van Meirvenne (2008). A multi-stage sampling strategy for the delineation
2219 of soil pollution in a contaminated brownfield. *Environmental Pollution* 154(2): 184-191.

2220 Viswanathan, Hari S., Bruce A. Robinson, et al. A geostatistical modeling study of the effect of
2221 heterogeneity on radionuclide transport in the unsaturated zone, Yucca Mountain.
2222 *Journal of Contaminant Hydrology* 62-63: 319-336.

- 2223 Watson, Alan G. and Randal J. Barnes (1995). Infill sampling criteria to locate extremes.
2224 *Mathematical Geology* 27(5): 589-608.
- 2225 Watson, Dave. (1999). The natural neighbor series manuals and source codes. *Computers &*
2226 *Geosciences* 25(4).

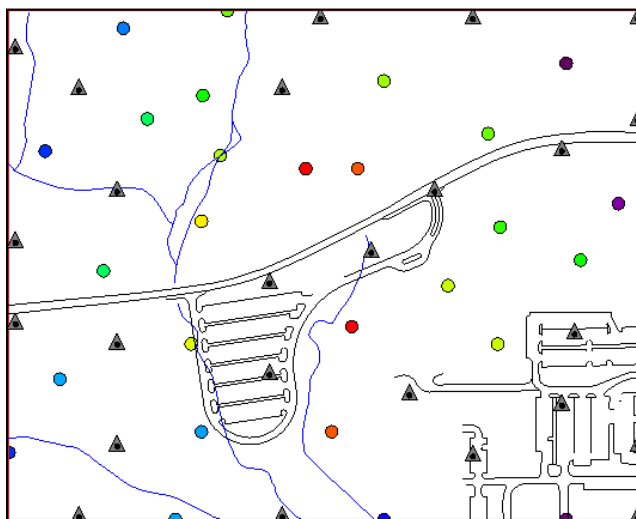
2227

13 APPENDIX: SAMPLE DESIGN STRATEGIES

2228

13.1 Adaptive Fill

2229 If historical sampling has been conducted for a scenario, then methods such as adaptive fill
 2230 (Stewart et al. 2009) aim to improve coverage by filling in spatial gaps from the largest to the
 2231 smallest in area. Figure 13.1 shows a 2D version with gray triangular samples fitted between
 2232 existing (circular) samples. Adaptive fill recognizes the presence of past samples but does not
 2233 consider activity levels measured by those samples. Adaptive fill may be useful in scoping if little
 2234 knowledge is recorded in the PSM or if the decommissioning team aims to sample a little more
 2235 in the wider gaps to better support decommissioning decisions. Other scenarios may apply as
 2236 well. Additional implementation details can be found elsewhere (Stewart et al. 2009, p. 470).



2237

2238 **Figure 13.1. Adaptive fill sample design.** Gray triangles indicate new samples adaptively situated
 2239 among existing (circular samples) to maximize coverage.

2240

13.2 Area of Concern Boundary Design

2241 In AOC boundary design, samples are placed along the boundary line between contaminated
 2242 and uncontaminated zones given by some decision criteria. In particular, those nodes that have
 2243 a value closest to the decision criteria are the targets of the design. They are selected in order
 2244 to more readily distinguish between contaminated and uncontaminated zones. This version is
 2245 currently available in SADA 5, and the reader is encouraged to review the user's guide and the
 2246 code for more information (Stewart et al. 2009). Figure 13.2 shows a 2D example.

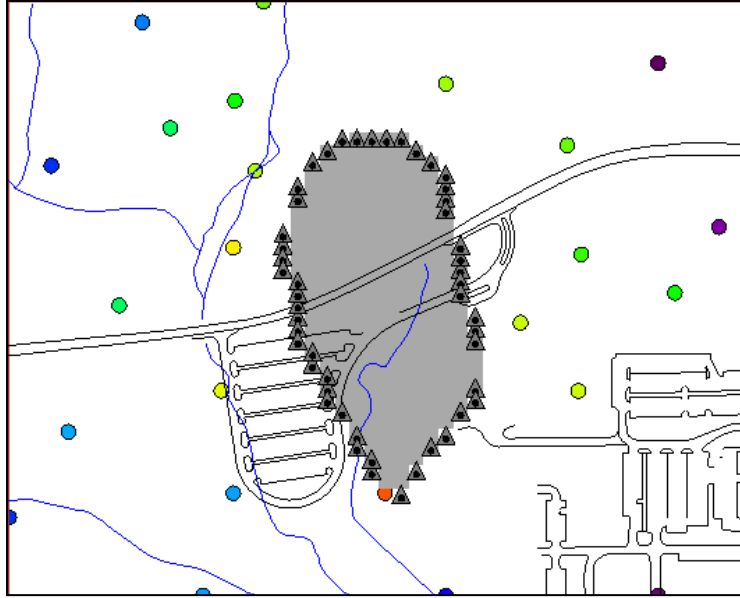


Figure 13.2. AOC boundary design showing a ring of new samples covering the boundary between contaminated (gray) and non-contaminated zones.

13.3 Bayesian Ellipgrid

Stewart et al. (2009) describe this approach as including a priori knowledge of finding a source in the calculation of an Ellipgrid-type search pattern. Here, areas where contamination is most likely to occur are favored with higher sampling than low-probability areas. This sampling approach is stratified, and stratification depends on likely contamination. This method is well aligned with the PSM, which would naturally provide the stratification needed via partitioning by Likert ranges. Those Likert values could be normalized (e.g., divided by 5) to a 0–1 scale and used within the Bayesian Ellipgrid approach. The challenge with search algorithms such as Ellipgrid and Bayesian Ellipgrid is they often require many samples to find small contaminant geometries. Figure 13.3 was taken from the SADA 5.0 users guide and shows an application of Bayesian Ellipgrid.

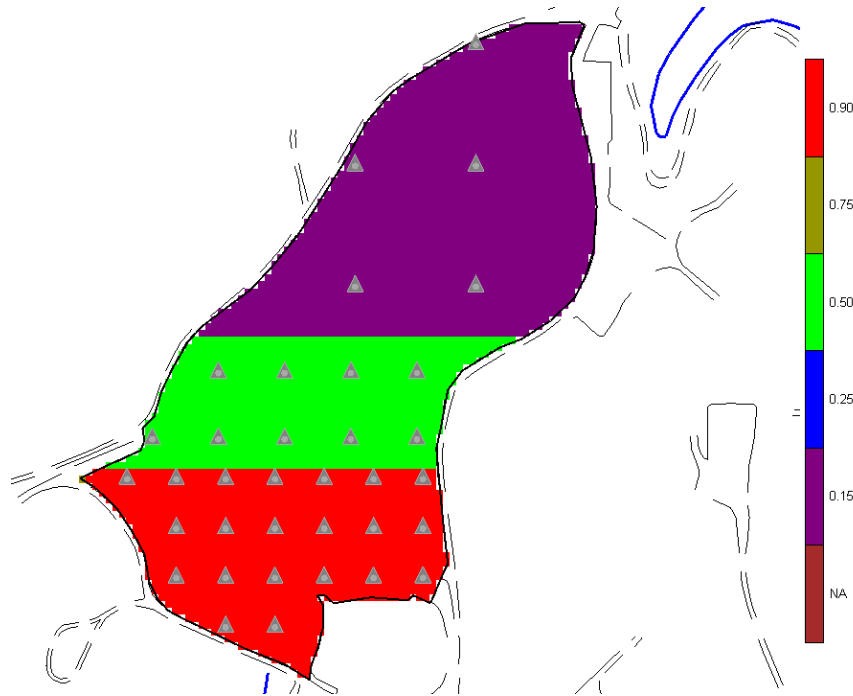


Figure 13.3. Bayesian Ellipgrid samples in tighter patterns over areas where the contamination is more likely to be found.

13.4 Check and Cover Sample Design

The check and cover sampling strategy, illustrated in Figure 13.4 and described by Stewart (2011), represents a pragmatic sampling approach useful at various stages in the process. Suppose that an investigator wishes to place samples where contamination is known or suspected to exist. This course of action appears to be reasonable, but validity depends on how close expert opinion or available lines of evidence coincide with reality. For this reason, investigators may wish to also place some samples in those locations believed to be uncontaminated. Sampling at both locations does two things. First, it mitigates the risk that current lines of evidence are wrong. Second, if contamination is encountered in anticipated locations, then samples in uncontaminated regions can provide some geographic limit to how widespread the contamination may be.

- Check: Sample where contamination is known or suspected to exist.
- Cover: Provide some sample coverage across the rest of the site.

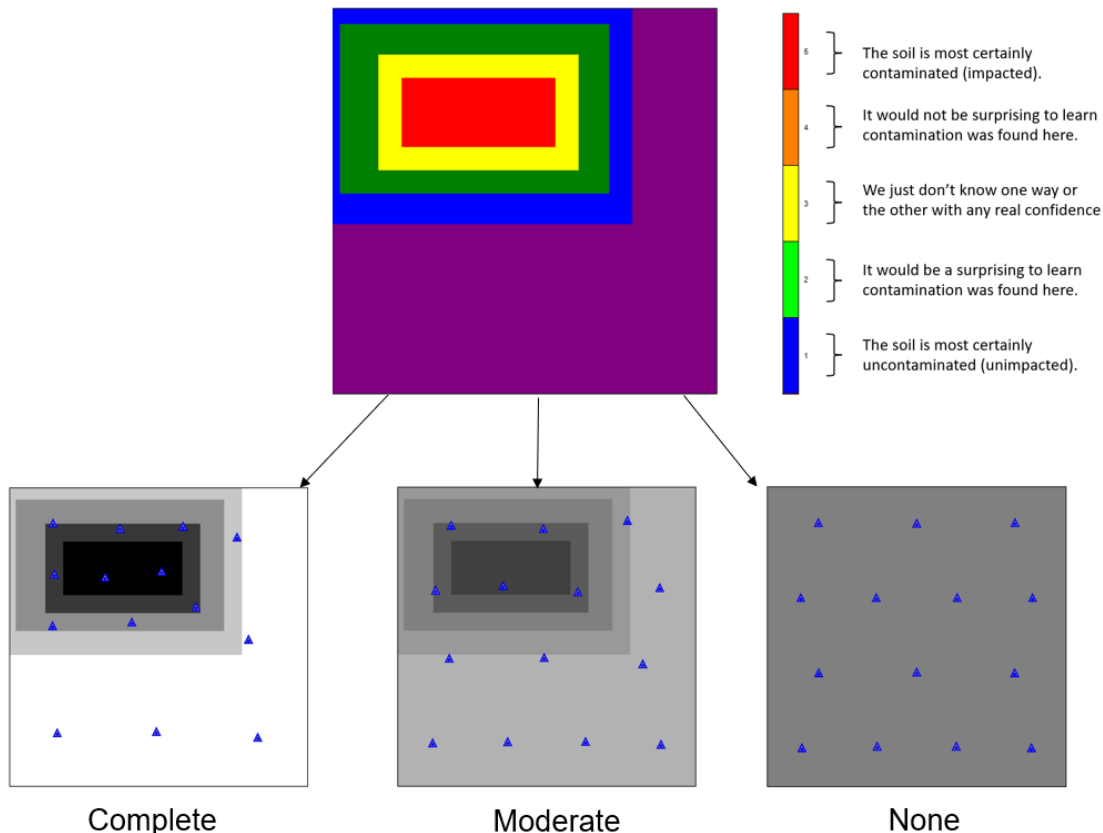


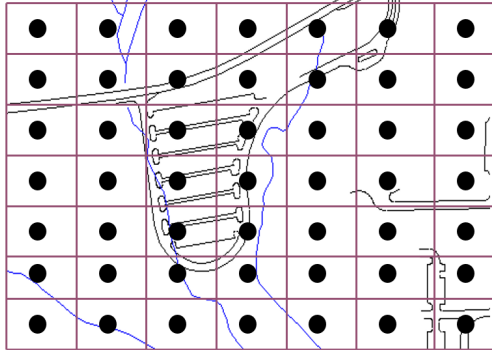
Figure 13.4. Check and cover sample design sample configurations under different confidence settings (from Stewart et al. 2011).

13.5 Final Status Survey Design

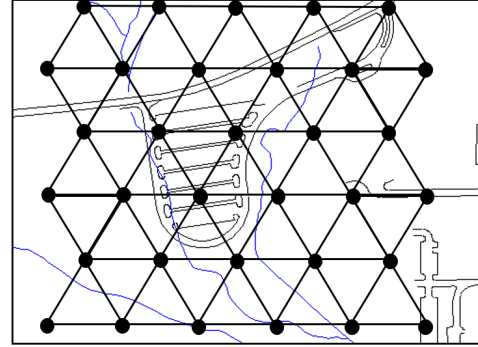
See Section 9.

13.6 Grid Designs

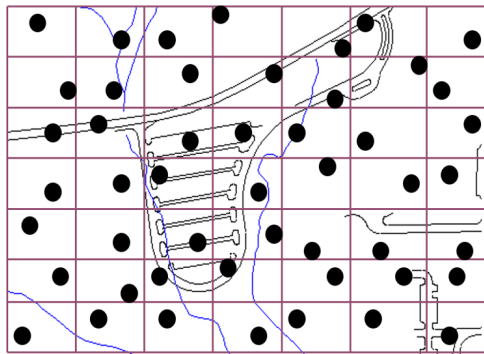
A large selection of gridded sample designs area is available for strategically covering or uniformly searching an area. For example, simple grids partition the site into an even set of cells that contain a single sample in the center of each cell. Simple unaligned grids randomize the sample location within the cell. Grid designs aimed at searching for material require specification of size, geometry, and tolerance for missing existing sources. Examples include Ellipgrid (Davidson 1994; 1995a; 1995b) and 3D Ellipgrid (Stewart 2009). These designs are generally thought of as uninformed, meaning that they do not use any prior site knowledge as in a PSM or other. Figure 13.5 shows example gridded designs



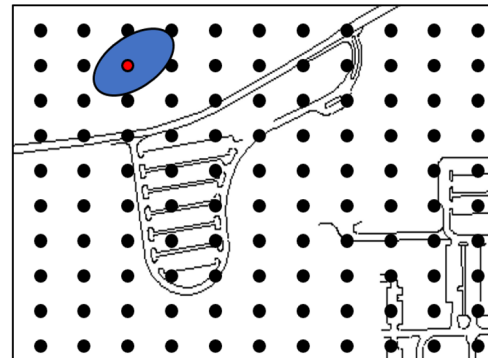
(a)



(b)



(c)



(d)

Figure 13.5. Gridded designs. (a) Rectangular grid, (b) triangular grid, (c) random gridded (randomized within grid cell), and (d) hot spot location based on search parameters and intended size of source (blue ellipse).

13.7 High-Value Designs

High-value designs (Stewart 2009, Chapter 41) aim to place sample locations where high values occur. For the PSM, this method would locate sample in areas with the highest Likert scores. For geostatistical values, this would be the highest kriged or averaged simulation values (Figure 13.6)

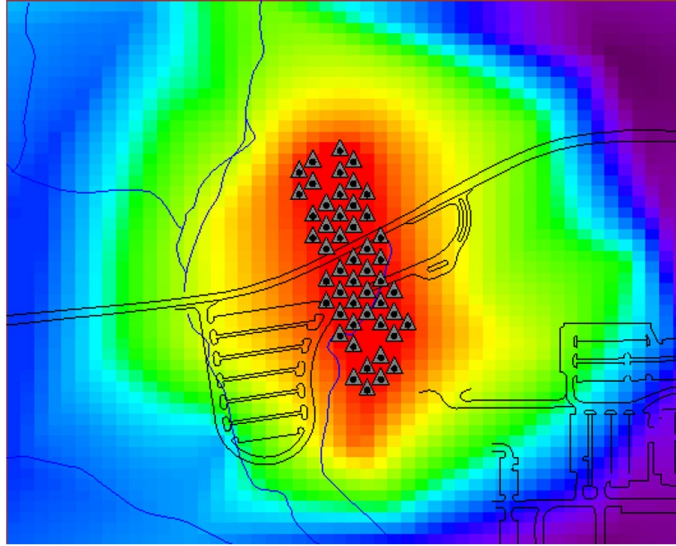


Figure 13.6. High-value design places samples in areas with the highest modeled values.

13.8 Judgmental Design

Judgmental design positions new samples where the decommissioning team judges they are most useful in terms of site conditions, the DQO process, and their relevance to the PSM update. These designs tend to be handcrafted and engineered heavily toward the particular scoping objectives at hand. For example, designs may place samples where known problems have occurred (should be represented with the PSM), to settle differing judgments about the presence and extent of contamination, or to specifically locate samples expected to be uncontaminated as a kind of confirmation. This type of approach is emphasized by Gogolak et al. (2020) for subsurface scenarios.

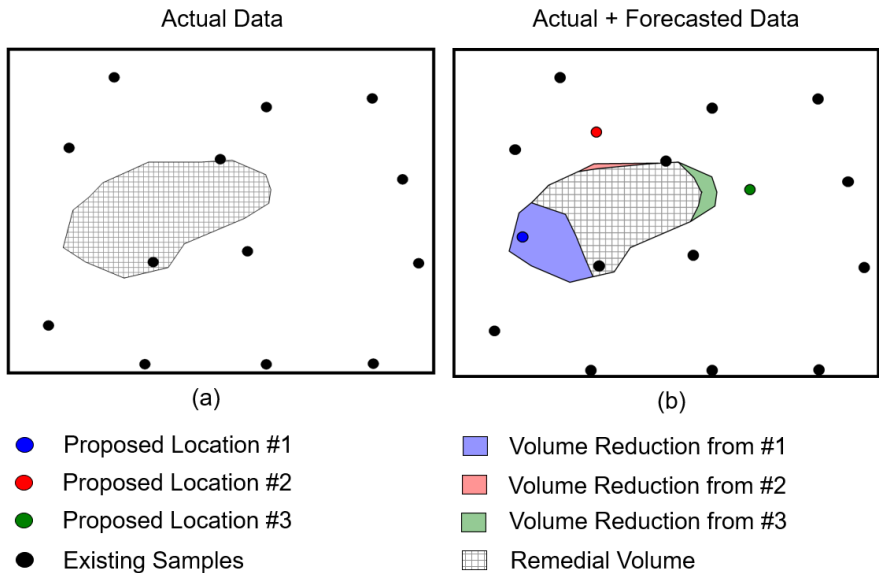
13.9 MrsDM

One way to maintain a high compliance standard while potentially reducing the volume of unnecessarily remediated soil is to strategically collect new samples. MrDM's remedial recommendations err on the side of safety by calling for remediation of soils that have relatively low probability of exceeding one or more DCGLvs. Strategically adding new samples can reduce low-risk but still failing areas by reinforcing model confidence with additional data values. The benefit of taking additional samples and the associated sampling cost must be weighed against the forecast reduction in the global remedial design and associated cost savings. Examples for a single-decision criterion over a single-exposure area are available in the literature (Pilger et al. 2001; Verstraete and Van Mervenne 2008; Freeze et al. 1992; England et al. 1992; Demougeot-Renard et al. 2004; Norberg et al. 2006; Back 2006, 2007). In Stewart (2011), MrsDM supplements the MrDM algorithm by identifying sample locations that may spatially increase decision confidence and reduce the remedial design requirements.

MrsDM begins with the set of proposed sample locations: for example, nodes of the foundation grid. For each candidate sample location, MrsDM estimates the effect on the remedial design strategy if the true measured values at that sample location are close to the value predicted by geostatistical simulations. Although the details are complicated, intuitively, MrsDM says if soil concentration values at select measured locations are as the model forecasts they are, then the

2329 decision confidence in that area could increase, thereby reducing the remedial volume
2330 associated with a MrDM design.

2331 An example may prove useful at this point. Suppose 13 existing core holes result in the MrDM
2332 remedial design shown as a gray set of remedial cells in Figure 13.7a. Investigators wish to
2333 choose the best two out of three proposed locations shown in Figure 13.7b along with their
2334 volumes of influence. Values for these three proposed locations are estimated and added to the
2335 full dataset, resulting in a new set of geostatistical realizations and a new MrDM, which is shown
2336 as the set of gray remedial cells in Figure 13.7. In addition to the new remedial area, the area
2337 reduced by the addition of the new proposed locations is shown in colors corresponding to the
2338 proposed location color.



2339

2340

Figure 13.7. MrsDM evaluation of proposed locations.

2341 In this example, proposed location #2 is assigned the least remedial reduction. Under MrsDM,
2342 the global reduction afforded by all three new sample locations is predicted to be the least
2343 diminished by the removal of proposed sample location #2. Therefore, the recommended new
2344 locations are #1 and #3.

2345 The cost associated with taking an additional sample is connected to the cost savings
2346 associated with the remedial reduction assigned to a proposed node. Such a relationship
2347 assists in determining the number of samples to collect. Indeed, when the cost of each
2348 additional sample exceeds the remedial benefit that sample adds to the total reduction, then the
2349 sample should be removed. This process can continue until the savings achieved by adding the
2350 proposed design outweigh the costs associated with taking the sample. Estimating these costs
2351 is beyond the scope of this report. However, the best next sample could be selected, and the
2352 remedial reduction associated with each removed sample design could be observed. Knowing
2353 both the cost of a sample and the savings imposed by the remedial design suggests when
2354 sampling is no longer economically viable. Under MrsDM, a stopping rule can be specified by
2355 cost consideration or by stating the number of desired samples.

2356 13.10 Random Designs

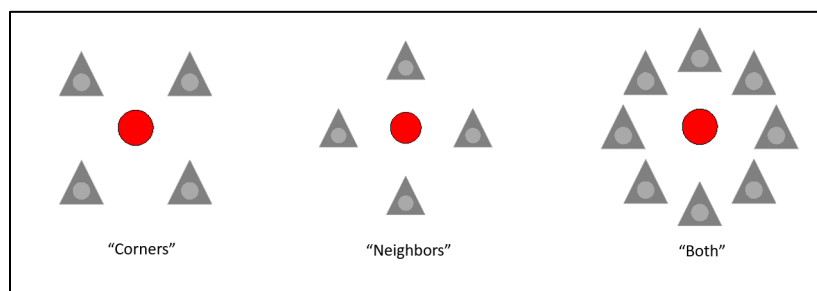
2357 As the name suggests, these designs randomly distribute new samples within a specified area.
2358 As shown in Figure 13.8, simple random designs scatter locations equiprobably throughout the
2359 study area. These are most appropriate for Class 3 where contamination is expected to be
2360 unlikely, and no known distributional patterns of radiological activity are expected.



2361
2362 **Figure 13.8. A 2D random sample design.**

2363 13.11 Threshold Radial

2364 In this design, samples that are excessively high (e.g., higher than DCGLVs) are geometrically
2365 locally bound by a set of samples. Options are available for bounding with four or more samples
2366 in select geometric patterns. A discussion of threshold radial can be found elsewhere (Stewart
2367 et al. 2009). Figure 13.9 shows threshold radial design configurations around a sample of
2368 interest (red).



2369
2370 **Figure 13.9. Three example geometric patterns for bounding an isolated hot-spot measurement.**

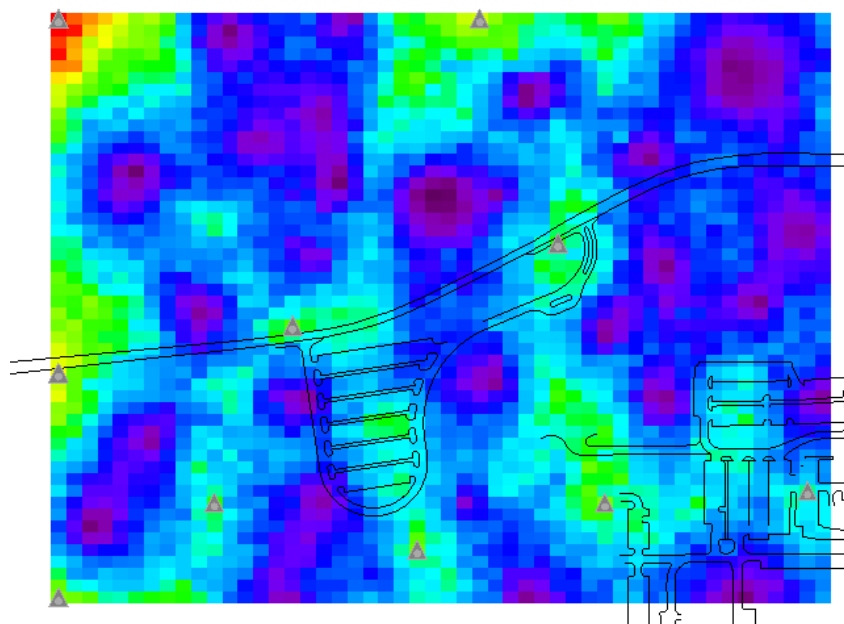
2371 13.12 Structural Sampling

2372 Several methods have been examined for sample designs aimed toward improving estimation
2373 of the covariance structure (e.g., semi-variogram) or of the entire geostatistical model. These
2374 methods are particularly helpful when experimental variograms are relatively difficult to control
2375 and additional sampling is needed to clarify the spatial structure. Strategies for variogram

2376 estimation are discussed elsewhere (Muller et al. 1999; Lark et al. 2002; Marchant et al. 2007;
2377 Lark et al. 2017; Zhu et al. 2005) (Picture not shown).

2378 **13.13 Variance Designs**

2379 A simulation-based variance map can be produced by simply processing the simulation stack,
2380 computing the variance for each node across all simulations at that node. High-variance sample
2381 design targets those areas with the highest simulation (or kriging) variance in an effort to settle
2382 what the true, but unknown, values in that region may be. Figure 13.10 shows 10 samples
2383 distributed to the highest-variance locations with a buffer restraint of 500 ft.



2384
2385 **Figure 13.10. Variance-based sample design shows new samples (gray triangles) in areas of**
2386 **highest model variance (e.g., kriging variance or simulation point variance)**

2387 This approach does not consider the estimated activity levels when placing samples. For
2388 example, the sample in the upper left is there because variance is highest often at the boundary
2389 outside of the sampling pattern. A level of judgment is necessary to screen these kinds of
2390 sampling artifacts from this sampled design.

2391