

Plus/minus 100 % - Fehlerabschätzung bei der Altlastenerkundung

Klaus Bücherl

Diplom-Geologe BDG,
Obmann des ITVA-Fachausschusses F2 Probenahme
LUBAG GmbH, Regensburg

1 Einführung

Kaum ein technischer Bereich muss sich mit derartig hohen Unsicherheiten und Fehlern beschäftigen wie die Altlastenbranche. Schlagzeilen über Kostenexplosionen bei der Altlastensanierung werden von Insidern achselzuckend hingenommen. Für Investoren sind diese scheinbar unkalkulierbaren Risiken aber ein Schreckgespenst und werden so zum Investitionshemmnis.

Die Verunsicherung auf beiden Seiten veranschaulicht folgendes Beispiel: Bei der Altlastenerkundung eines Altstandortes für die Errichtung eines Einkaufszentrums war eine bis zu 7 m mächtige künstliche Auffüllung angetroffen worden, die „bereichsweise leicht bis deutliche erhöhte Gehalte der Schadstoffparameter Schwermetalle inkl. Arsen und/oder polyzyklische aromatische Kohlenwasserstoffe (PAK)“ aufwies. Im Rahmen einer mehrstufigen Untersuchung wurden auf einer Fläche von 67.000 m² insgesamt 91 Bohrungen abgeteuft und rund 200 Proben untersucht. Der Gutachter schätzte das Volumen der Auffüllung auf 362.257 m³ und die Entsorgungskosten auf 13,2 Mio. € Auf die Frage des Investors, wie sicher denn diese Schätzung sei, gibt der Gutachter ein Kostenrisiko von 30 % an. Auf die Nachfrage, ob denn dieser Betrag sicher nicht überschritten werden könne räumt er ein, dass dies nicht auszuschließen sei, legt sich aber nicht auf eine Obergrenze fest.

Wer etwas von Altlasten versteht, wird Verständnis für den Gutachter haben. Im Grunde hatte er sich mit den 30 % schon sehr weit aus dem Fenster gelehnt. Immerhin basierte die Schätzung auf einer Probe pro 1.800 m³ Auffüllung! Aber: dürfen wir uns damit zufrieden geben? Die für das Flächenrecycling benötigten (dringend gesuchten) Investoren brauchen – heute mehr denn je – Sicherheit. Und dort, wo Unsicherheiten bestehen, müssen wir diese transparent machen und nachvollziehbar beziffern!

Es gibt verschiedene Konzepte um diese Unsicherheiten im Zusammenhang mit der Grundstückswertermittlung zu quantifizieren. So bringt die Risikoabschätzung nach Kerth/Grient [9] realistische Spannen der Kostenunsicherheit, die sich im Lauf der Bearbeitung von anfangs +180 %/-90 % (noch keine Untersuchungen) auf +30 %/-15 % nach der Sanierungspla-

nung verringert. Eine derartige Abschätzung steht und fällt allerdings mit der Qualität der Erkundung. Wird der Sanierungsumfang aufgrund mangelhafter historischer Erkundung oder fehlerhafter Probenahmen unterschätzt, werden auch die Kostenrisiken zu gering angesetzt.

2 Einflussgrößen der Kostenschätzung

Um das Kostenrisiko einer Altlastensanierung durch Aushub im Einzelfall zu bewerten, müssen die Einflussgrößen einzeln betrachtet und bewertet werden. Diese sind

- Probenahme (PN)
- Analytik (LAB)
- Auswertung, z.B. Interpolation (A)
- technischen Toleranzen beim Aushub (Vermischung, Sicherheitszuschläge) (Bau)
- Schätzung des Preises (Preis)

Dabei bleibt offen, ob sich diese Fehler nach Gauss oder gemäß der Größtfehleraddition fortpflanzen. Im letzteren Fall kommt man selbst bei konservativer Annahme auf größtmögliche Abweichungen der geschätzten Kosten nach oben in der Größenordnung von 100 %.

Die größte Unsicherheit, sowohl von der mutmaßlichen Größenordnung des Fehlers, als auch von der Datengrundlage her, steckt im Probenahmefehler, der daher näher beleuchtet werden muss.

3 Grundlagen für die Fehlerabschätzung bei der Probenahme

Nach der DIN EN ISO/IEC 17025, die sich nicht nur auf Prüflaboratorien, sondern auch auf die probenehmenden Ing.-Büros bezieht, müssen diese Methoden zur Abschätzung der Messunsicherheit auf der Grundlage des gegenwärtigen Kenntnisstandes verfügen. Die akademische Diskussion, dass die Probenahme an sich keine Messung sei, hilft uns hier nicht weiter. Schließlich müssen wir als Gutachter dem Kunden gegenüber fundierte Aussagen zur Unsicherheit der gesamten Kette von der Probenahme bis hin zur Massen- und Kostenschätzung liefern und vertreten.

Hinweise zur Umsetzung dieser Anforderung gibt z.B. die DACH¹ oder der DAR²: Die Methoden zur Fehlerabschätzung sollen zweckmäßig und nicht so genau wie möglich, sondern so

¹ Deutsche Akkreditierungsstelle Chemie GmbH

² Deutscher Akkreditierungsrat

genau wie nötig sein. Wo eine mathematisch statistisch strenge Vorgehensweise nicht möglich oder sinnvoll ist können auch Know-how und Erfahrung die Grundlage bilden.

4 Welche Fehler gibt es?

4.1 Grobe Fehler

Nach Fehlern bei der Altlastenerkundung, speziell bei der Probenahme gefragt, fallen jedem sofort Beispiele dafür ein, was man alles falsch machen kann. Dabei handelt es sich in aller Regel um „grobe Fehler“. Damit sind Fehler gemeint, die auf mangelnde Qualifikation, Inkompetenz, Fehlbedienung, Schlamperei u.ä. zurückzuführen sind ("geotechnische Kunstfehler"). Grobe Fehler können völlig unrealistische Ergebnisse liefern (z.B. Fremdkontamination einer Probe), aber auch systematisch sein (z.B. BTEX-Kontamination von Proben durch Lagerung neben Kraftstoffen). Die meisten groben Fehler lassen sich im nach hinein nicht mehr feststellen und bleiben damit unerkannt. Nur in seltenen Fällen dokumentieren „grob-Fehlergeneigte“ Probenehmer die Probenahme so genau und vollständig, dass sich Fehler aus dem Probenahmeprotokoll ableiten lassen. Das wäre zum Beispiel der Fall, wenn in einem Probenahmeprotokoll für Boden wird eine verchromten Handschaufel als Entnahmegesetz angegeben wird.

Grobe Fehler können auch weniger offensichtlich sein: Bei einer Grundwasserüberwachung traten unplausibel hohe Kupfer- und Zinkkonzentrationen auf. Der Grund: Bei der Membranfiltration der Probe tauchte die Messingspitze des Filtergerätes in die angesäuerte Probe wobei Kupfer und Zink in Lösung gingen

Nicht selten werden grobe Fehler aber auch aus Kostengründen bewusst in Kauf genommen. So zum Beispiel beim Einsatz von Rammkernsondierungen bei großen Korndurchmessern (insbesondere in Auffüllungen) und in zu großen Tiefen oder die Verwendung von Rammfilterpegeln für die Grundwasserbeprobung. In die gleiche Kategorie fallen auch Fehler bei der Probenahmeplanung, wie der Verzicht auf eine historische Erkundung oder eine erhebliche Reduzierung der Probenahmepunkte ohne die Verschlechterung der Erkennungswahrscheinlichkeit für eine unbekannt Kontamination zu berücksichtigen.

4.2 Systematische Fehler („systematische Messabweichungen“ nach DIN 1319,T1)

Ebenfalls auf Fehlverhalten zurückzuführen, aber unter Umständen besser kalkulierbar, sind systematische Fehler. Das sind alle Fehler, die reproduzierbar in der gleichen Größenordnung und im gleichen Sinn (mit gleichem Vorzeichen, d.h. Mehr- bzw. Minderbefund) auftreten. Systematische Fehler können Methoden bedingt sein (z.B. Minderbefunde bei Schichtmäch-

tigkeiten durch Stauchung bei RKS) oder auf einer fehlerhaften "Versuchsanordnung", also stets gleich falscher Ausführung einer Probenahmemethode, beruhen.

Ein Beispiel hierfür ist die Verwendung eines Lichtlotes mit überdehntem oder durch Reparatur verkürztem Maßband oder die Drift bei Vor-Ort-Messgeräten.

4.3 Zufällige Fehler („zufällige Messabweichungen“ nach DIN 1319,T1)

Zu den groben und systematischen Fehlern kommt der jeder Messung zuzuordnende zufällige Fehler. Darunter versteht man zufallsbedingte Abweichung einer Messung oder Prüfung, die stets auftreten, auch wenn alle groben und systematischen Fehler ausgeschlossen sind. Sie sind nach Betrag und Vorzeichen zufällig verteilt und lassen sich mit den Methoden der statistischen Mathematik behandeln.

4.4 Variabilität

Nicht verantwortlich ist der Probenehmer/Gutachter für einen weiteren Faktor, der erheblich zur Unsicherheit einer Aussage bei der Altlastenerkundung beiträgt: Die Variabilität der untersuchten Medien. Die quantitative Erfassung der Variabilität einer Kontamination ist Aufgabe einer Detailuntersuchung. Dabei ist die Anzahl der zu untersuchenden Proben abhängig von der Streubreite und der angestrebten Aussagesicherheit.

Unter dieses Stichwort fallen die meisten „Überraschungen“ bei der Altlastensanierung wie kleinräumige Schadstoffnester, verborgene Fundamente, Bombentrichter usw.. Allerdings lassen sich auch diese Inhomogenitäten statistisch beschreiben. So lässt sich mit Methoden der Statistik berechnen, welche Anzahl von Bohrpunkten in einem Raster erforderlich wäre, um z.B. jeden Bombentrichter mit mindestens 10 m Durchmesser zu erfassen, oder anders ausgedrückt: wie groß die Wahrscheinlichkeit ist, jedes Schadenszentrum einer bestimmten Dimension erfasst zu haben.

Geht man zum Beispiel bei einer Fläche von 10.000 m² von der Hypothese aus, dass rund 10 % der Fläche, mit Schadstoffgehalten über dem Hilfswert 2 belastet ist, und diese Belastung auf 10 einzelne sogenannte Hot-Spots verteilt ist, so ergibt sich bei 15 in einem Dreiecksraster angeordneten Kleinrammbohrungen eine Wahrscheinlichkeit von rund 85 % mindestens einen dieser Hot-Spots aufzufinden. Um die Wahrscheinlichkeit auf 99 % zu erhöhen, muss man die Anzahl der Bohrungen mit 31 mehr als verdoppeln. Um nicht nur eines, sondern alle 10 vermuteten Schadstoffnester zu entdecken (oder auszuschließen), benötigt man schon 100 Bohrungen.

5 In welcher Größenordnung liegen die Fehler bei der Altlastenerkundung?

Am problematischsten für die Fehlerabschätzung sind die **groben Fehler**, die erheblich größer als 100 % sein können. Anders ausgedrückt: Ergebnisse von Untersuchungen, in deren Verlauf grobe Fehler gemacht wurden, sind praktisch wertlos! Leider werden grobe Fehler meist nicht eindeutig erkannt bzw. benannt und werden daher nur in Ausnahmefällen wirklich aussortiert. Die Folge wäre, dass man die Erkundung neu aufziehen oder zumindest fundiert qualifiziert überprüfen müsste. Da man diesen Aufwand meist scheut, werden Planungen dann auf zweifelhaften Daten aufgebaut –mit fatalen Folgen. Es kann aber auch passieren, dass es nicht gelingt, fehlerhafte Mehrbefunde unter den Tisch fallen zu lassen, weil die Fach- und Genehmigungsbehörden diese zumindest als „Hinweis“ werten. Ein Wert, sei er noch so falsch, ist nicht mehr aus der Welt zu schaffen und kann auf Dauer mehr Kosten verursachen, als ggf. bei der Untersuchung eingespart wurde.

Für die Quantifizierung **systematischer** und **zufälliger Fehler** liegen kaum Erfahrungswerte vor. Einen Eindruck von der Dimension des systematischen Fehlers bei der Ermittlung von Schichtmächtigkeiten vermittelt die vom ITVA-Fachausschuss F2 durchgeführte Vergleichs-Probenahme, an der 8 Ingenieurbüros bzw. Untersuchungsstellen teilgenommen haben. Bei diesem Versuch wurden Kleinrammbohrungen mit 50 mm Durchmesser in ein bekanntes (künstliches) Bodenprofil aus Fein- bis Mittelsanden mit bekannter MKW-Kontamination abgeteuft. Dabei ergaben sich bei Schichtmächtigkeiten zwischen 0,2 und 0,6 m Abweichungen von – 50 % bis + 90 %. Die MKW-Gehalte der von den verschiedenen Büros entnommenen Proben wichen zwischen –30 % und +40 % vom bekannten wahren Wert der mit MKW dotierten Schichten ab. Bei diesen auf den ersten Blick großen Abweichungen ist zu bedenken, dass aufgrund der Versuchsanordnung grobe Fehler ausgeschlossen werden können. Da alle Proben im gleichen Labor (VEGAS) untersucht wurden, sind auch die Analyseergebnisse miteinander auch vergleichbar.

Den Versuchsaufbau und die einzelnen Ergebnisse werden wir demnächst im altlasten spektrum veröffentlichen.

Die Größenordnung der **Variabilität** bei einer Altlastenerkundung hängt von den Eigenschaften der Schadstoffe und denen des Untergrundes ab. Am größten ist sie beim Auftreten lipophiler organischer Phasen in klüftigem Festgestein, am geringsten bei löslichen Salzen in gleichförmigen Lockersedimenten. Die Variabilität ist naturgemäß in jeden Einzelfall anders und muss auch in jedem Einzelfall neu ermittelt werden. Die Spannbreite der möglichen Konzentrationen kann dabei von Gehalten unter der Bestimmungsgrenze bis in den höheren Pro-

zentbereich reichen. Typisch sind dabei ausgeprägt rechtsschiefe Verteilungen (Median << Mittelwert) und Standardabweichungen, die den statistischen Mittelwert der Schadstoffkonzentrationen um den Faktor 2 und mehr übersteigen.

6 Folgerungen und Ausblick

Die bei der Probenahme auch ohne grobe Fehler auftretenden erheblichen Abweichungen sind zwar im Prinzip jedem Experten geläufig, werden aber in der Praxis nicht berücksichtigt. Ein Grund dafür ist, dass Zahlen zur Bemessung der Unsicherheiten fehlen. Die vom Fachausschuss F2 durchgeführte Vergleichsprobenahme vermittelt einen Eindruck, in welcher Größenordnung diese Unsicherheiten liegen können. Bei dem dort verwendeten künstlich kontaminierten Boden müssten die MKW-Gehalte der einzelnen Büros mit einer Fehlerbreite von rund 1/3 angegeben werden, also z.B. 1.200 mg/kg (- 360 mg/kg / + 480 mg/kg). Betrachtet man die Erfahrungen aus Laborringversuchen, dürfte der Gesamtfehler wohl noch größer sein.

Aus fachlicher Sicht kann man demnach bei einem MKW-Gehalt von 1.200 mg/kg nicht von einer Überschreitung des Hilfswertes 2 des Bayer. LfW-Merkblattes 3.8/1[1] bzw. der Z2-Wertes[10] sprechen. Dennoch wird bei der orientierenden Erkundung aufgrund des Besorgnisgrundsatzes feststellen müssen, dass sich der Verdacht auf das Vorliegen einer schädlichen Bodenveränderung in diesem Fall bestätigt hat.

Den auf die hohe Variabilität zurückzuführenden Unsicherheiten bei Massen- und Kostenschätzungen kann man durch eine höhere Anzahl von Proben und Analysen begegnen. Diese Konsequenz zieht auch eine Studie der Forschungsvereinigung Recycling u. Wertstoffverwertung im Bauwesen e.V., die für eine Erhöhung der Probenzahlen je Charge plädieren[16].

Gegen grobe Fehler allerdings hilft nur eines: Qualitätssicherung! Vor diesem Hintergrund ist der bayerische Weg der Zulassung von Untersuchungsstellen gem. §18 BBodSchG für die Probenahme zu begrüßen.

Doch auch bei der Ausschaltung aller groben und systematischen Fehler und der genauen Erfassung der Variabilität wird der Fehler von Kostenschätzung bei der Altlastensanierung nie die Genauigkeit anderer Bau-Gewerke erreichen. Unvorhersehbare Kostenexplosionen bei der Altlastensanierung um ein Vielfaches der Schätzkosten sind allerdings kein Naturgesetz. Durch qualifizierte Probenahme, von der Aufstellung des Probenahmeplans bis zur Auswertung, werden auch die Kosten der Altlastensanierung kalkulierbar.

Literatur:

- [1] **BAYER. LANDESAMT FÜR WASSERWIRTSCHAFT (2001):** Untersuchung und Bewertung von Altlasten, schädlichen Bodenveränderungen und Gewässerverunreinigungen - Wirkungspfad Boden-Gewässer -, Slg LfW Merkblatt 3.8/1, Stand 31.10.01.- München
- [2] **DIN 4021**, Ausgabe:1990-10
Baugrund; Aufschluss durch Schürfe und Bohrungen sowie Entnahme von Proben
- [3] **DIN ISO 10381-1**, Ausgabe:2003-08, Bodenbeschaffenheit - Probenahme - Teil 1: Anleitung zur Aufstellung von Probenahmeprogrammen (ISO 10381-1:2002),
- [4] **DIN ISO 10381-2**, Ausgabe:2003-08, Bodenbeschaffenheit - Probenahme - Teil 2: Anleitung für Probenahmeverfahren (ISO 10381-2:2002)
- [5] **DIN ISO/IEC 17025:** Allgemeine Anforderungen an die Kompetenz von Prüf- und Kalibrierlaboratorien: 2004-02
- [6] **DIN 38402-41**, Ausgabe:1984-05
Deutsche Einheitsverfahren zur Wasser-, Abwasser- und Schlammuntersuchung; Allgemeine Angaben (Gruppe A); Ringversuche, Planung und Organisation (A 41)
- [7] **IMMLER, L.G., ZAHN, M.T.:** Die flächenhafte Variabilität bodenphysikalischer Parameter und des Corg-Gehaltes in den Pflugsohlen je eines Ton-, Sand- und Lößstandortes, Z. Pflanzenernähr. Bodenk., 157, 251-257 (1994)
- [8] **ITVA-ARBEITSHILFE F2-2/95:** Aufschlussverfahren zur Feststoffprobengewinnung für die Untersuchung von Verdachtsflächen und Altlasten
- [9] **KERTH, M., VAN DE GRIENDT, B.:** Monetäre Bewertung des „Altlastenrisikos“, HdA, 22. Erg.-Lfg. 2. Aufl., Oktober 2000
- [10] **LÄNDERARBEITSGEMEINSCHAFT ABFALL 'LAGA' (NOV. 1998):** Mitteilung 20: Anforderungen an die stoffliche Verwertung von mineralischen Reststoffen/Abfällen - Technische Regeln. 4. erweiterte Auflage
- [11] **LIDL, R. UND TEUTSCH, G:** Statistische Absicherung der Probenahme bei der Gefahrenherdenerkundung, Abschlussbericht LAG 98-01/0460 zum LfU-Werkvertrag Nr. 126760/54, Tübingen, 30.06.2002
- [12] **RAMSEY, M.H, SQUIRE, S., GARDNER, M.J.:** Synthetic reference sampling target for the estimation of measurement uncertainty, Analyst, 1999, 124, 1701-1706
- [13] **SQUIRE, S., RAMSAY, M.H., GARDNER, M.J.:** Collaborative trial in sampling for the spatial delineation of contamination and the estimation of uncertainty, Analyst, 2000, 125, 139-145
- [14] **SQUIRE, S., RAMSAY, M.H., GARDNER, M.J., LISTER D.:** Sampling proficiency test for the estimation of uncertainty in the spatial delineation of contamination, Analyst, 2000, 125, 2026-2031
- [15] **SQUIRE, S., RAMSAY, M.H.:** Inter-organisational sampling trials for the uncertainty estimation of landfill gas measurements, J. Environ. Monit. 2001, 3, 288-294
- [16] **WEFERS, H. ET. AL.:** LAGA-Kriterien für Baureststoffe: Evaluation der Analytik und Entwicklung praxisnaher Umsetzungsmöglichkeiten, Forschungsprojekt der For-

schungsvereinigung Recycling u. Wertstoffverwertung im Bauwesen e.V. (RWB),
AIF-Vorhaben FV-NR. 12350 N