

分けたことで説明変数を加えたような効果が起こり、それが良い方向に働いたと考えられる。(4)重回帰モデルと数量化モデルを比較すると、全てにおいて数量化の方が高い。特に底質では著しい向上が見られる。これは、数量化は非線形性を表現できるためであると考えられる。

次に、両モデルの検証を全物質・全地点（水質）を用いて行った。方法としては、全サンプルを無作為に半分に分け、片方をモデル構築に、片方をモデル検証に用いた。なお解析は10回繰り返して行った。その結果の平均値を表3に示す。数量化モデルの方が多少SEのばらつきが大きい、両モデルともモデル構築と検証は同程度のSEを示している。よって、両モデルともモデル構築と同程度の精度（予測値が実測値から±1オーダーの範囲に入る）で予測が行えることが分かった。

重回帰モデルを用いて、物質別・地点別で解析を行った。その結果を図2に示す。図2より、予測値は実測値から±1オーダーの範囲に入っている。それぞれの物質（地点）について調べると、全物質・全地点での解析に比べてモデルへの適合度が良いことが分かった。

最後に、物質別・地点別（水質）についてモデルの検証を行った。方法は全物質・全地点の場合と同様である（表4）。これより、モデル検証の方が構築よりかなりSEが大きい。また、SEのばらつきも大きい。これは、10個程度のサンプルでモデルを構築したため、10個のサンプルに回帰式を合わせた形になり、モデルへの適合度は良いが予測精度が下がったと考えられる。そこで、安定した濃度予測を行うのに必要なサンプル数を調べるために、隅田川河口（水質）において物質数を増やしてモデルの検証を行った。追加した物質は過去に隅田川河口で検出され、その生産が現在も続いている8物質である。検証方法は、全サンプル29個のうちで構築用のサンプル（n）を10、15、20、25と増やしていき、残り（n'）を検証に用いた。なお、解析は10回繰り返して行った。その結果を表5に示す。nが10のとき、モデル構築のSEは最も小さいにも関わらず、検証では最も大きい。また、SEのばらつきも最も大きい。それ以外では、モデル構築、検証共にn（n'）が増える（減る）に従ってSEが小さくなっている。また、徐々にモデル構築と検証のSEが近づいている。よって、物質別・地点別モデルにおいてもサンプル数が15以上ある場合にはモデルの安定性が高くなり、精度の良い濃度予測を行えると考えられる。

4. 結論

重回帰モデル、数量化モデルのようなモデル構築の容易な統計モデルを用いて、物性の異なる化学物質の環境中濃度を±1オーダーの範囲で予測できることが分かった。また、両モデルの使用方法として、重回帰モデルは、多くの物質・地点について解析を行う場合にも、ある一地点でいくつかの物質について解析を行う場合にも使用できる。数量化モデルは、多くの物質・地点について、特に非線形性を含むと考えられる場合に有効である。しかし、数量化モデルは予測を行う際に、モデル構築に用いた説明変数のデータ範囲に入らないデータの外挿ができないという短所を持っている。

表3 重回帰モデル、数量化モデルの検証（全物質・全地点）

		n	R ²	SE	SEの幅	
重回帰モデル	モデル構築	158.1	0.496	0.579±0.019	+0.029	-0.036
	モデル検証	153.9		0.602±0.021	+0.033	-0.025
数量化モデル	モデル構築	156.6	0.587	0.523±0.037	+0.070	-0.061
	モデル検証	155.4		0.583±0.042	+0.061	-0.080

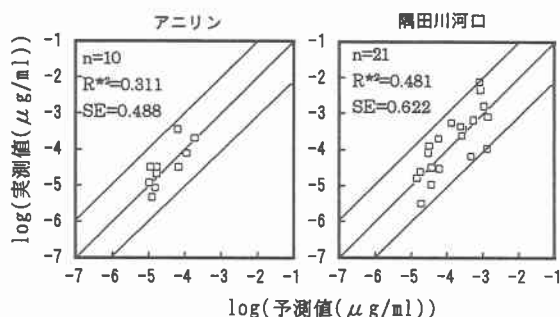


図2 物質別・地点別（水質）の予測値－実測値

表4 重回帰モデルの検証（物質別・地点別）

	物質・地点名	n	R ²	SE	SEの幅	
モデル構築	アニリン	7.5	0.226	0.427±0.155	+0.268	-0.290
	隅田川河口	10.5	0.600	0.478±0.157	+0.263	-0.276
モデル検証	アニリン	6.5		0.769±0.313	+0.602	-0.315
	隅田川河口	10.5		0.749±0.306	+0.733	-0.327

表5 サンプル数増加に伴う予測精度

	n(n')	R ²	SE	SEの幅	
構築	10	0.253	0.645±0.242	+0.415	-0.377
	15	0.085	0.771±0.106	+0.174	-0.232
	20	0.202	0.749±0.063	+0.107	-0.099
	25	0.208	0.728±0.031	+0.047	-0.045
検証	19		1.451±0.403	+0.741	-0.411
	14		0.811±0.132	+0.205	-0.221
	9		0.776±0.188	+0.444	-0.344
	4		0.707±0.213	+0.236	-0.379