CHAPTER 19
Possibilities and Limitations of Short-term Tests for Ecotoxicologic Effects: Terrestrial Approaches

JAMES W. GILLETI

19.1 INTRODUCTION
This chapter attempts to set forth the problems of evaluation and assessment of impacts of organic chemicals in the terrestrial environment. It is very difficult to distinguish between and separate terrestrial and aquatic systems over concerns about chemical pollution. Land use patterns determine much of water quality. At the interfaces between terrestrial and aquatic environments (wetlands, stream- and lakeside, beaches), there are many cross-inputs. Some species (invertebrates and amphibians) spend parts of their lives in one and then another habitat. Some, apparently terrestrial, species may associate almost exclusively with the aquatic environment. The osprey, for example, only touches land in the form of its nest and perch, feeding exclusively on fish. Is it an aquatic or terrestrial species?

19.2 PAST AND PRESENT APPROACHES TO TERRESTRIAL ECOTOXICOLOGY

19.2.1 Animals
19.2.1.1 General considerations
Terrestrial toxicologic investigations have mainly focused on health effects in mammalian species as models of potential impact on human health. However, the responses of white rats to acute oral, chronic dietary, and lifetime dietary exposures provide basic mammalian toxicological data not only on mortality and pathology, but also on growth, reproduction, gross organ toxicity, and functional impairment (e.g. neurotoxicity). Since these same concerns pertain to wild terrestrial mammals as well as humans, then rat toxicity studies (and surrogate in vitro tests) also serve terrestrial ecotoxicology.
Short-term Toxicity Tests for Non-genotoxic Effects

Most single species toxicity testing has been conducted with insects (to screen the effectiveness of products against injurious (pest) or beneficial insects) and rodents (mainly rats, mice and guinea-pigs and undertaken for human health and safety reasons). Other species have received much less attention. Interest in impacts on avifauna, spurred by Rachel Carson’s *Silent Spring* (1962), led to extensive testing of chemicals in birds (Tucker and Crabtree, 1970). However, the ecotoxicologic problems that became evident (Pimentel, 1971) involved sparse species at the top of the food chain, and the extensive testing which took place on birds such as chickens, quail and pheasant (all gallinaceous species and grain eaters) did not result in the same effects that were being observed in predators such as eagles, hawks, owls, ospreys and pelagic sea birds (which suffered from egg-shell thinning and reproductive failure).

Substantial efforts by the US Fish and Wildlife Service and other research organizations around the world have developed laboratory-reared species for such tests (e.g. the sparrowhawk (kestrel) and barn owl). Routine testing of pesticides, however, continues to involve bobwhite quail, pheasants and mallard ducks, selected to represent indigenous species at risk of exposure. Protocols in use for the evaluation of pesticides call for acute, subchronic, and field pen testing for mortality and effects on reproduction and behaviour. However, factors such as impact of migration and cross-exposure to the same or other pesticides in global transits are not readily tested.

Acute tests (of species) are usually performed as range-finding tests to ascertain levels to use in chronic exposures. Chronic or iterative exposures in turn are used not only for determining mortality but also to obtain information about toxic effects on specific physiological functions, some examples of which are shown in Table 19.1. These examples demonstrate effects that have been found to be useful in analysing or predicting outcomes of the use of pesticides and other toxic substances in natural environments. They can be critical for one or more species, but may be irrelevant in terms of human toxicity. On the other hand, alternative animal models can provide better prediction of adverse human impacts on specific physiological functions than many traditional laboratory species (such as the rat). Examples of these alternative animal models are given in Table 19.2.

Single species toxicity tests serve only as a measure of impact on individual members of an ecosystem. Of greater importance are measures of irremedial or irreversible loss of species or function due to chronic or iterative exposure to a toxicant (Cairns *et al.*, 1981). Neuhold and Ruggerio (1976) noted at least five types of serious adverse effects for which ecotoxicologic concerns are high:

(i) loss of primary productivity;
(ii) loss of secondary productivity, growth and reproduction;
(iii) disturbance in material and nutrient cycling;
(iv) altered ecosystem structure, diversity and complexity; and
(v) loss of endangered species or their habitats.
Table 19.1 Tests with end-points other than mortality for the measurement of toxicity in terrestrial environments

<table>
<thead>
<tr>
<th>Species</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avian</td>
<td>Reproduction—egg-shell thinning</td>
</tr>
<tr>
<td>Chicken, quail, pheasant, grouse, sparrowhawk, mallard, barn owl</td>
<td>Reproductive behaviour</td>
</tr>
<tr>
<td>Herring gull</td>
<td>Salt gland</td>
</tr>
<tr>
<td>Doves, finches</td>
<td>Reproductive behaviour</td>
</tr>
<tr>
<td>Amphibian</td>
<td>Morphogenesis</td>
</tr>
<tr>
<td>Frogs, toads</td>
<td>Morphogenesis</td>
</tr>
<tr>
<td>Arthropods</td>
<td></td>
</tr>
<tr>
<td>Honeybees</td>
<td>Communication, other behaviours</td>
</tr>
<tr>
<td>Crickets</td>
<td>Calling behaviour</td>
</tr>
<tr>
<td>Various spp.</td>
<td>Metamorphosis</td>
</tr>
<tr>
<td>Spiders</td>
<td>Web-spinning</td>
</tr>
<tr>
<td>Protista</td>
<td></td>
</tr>
<tr>
<td>Bacteria</td>
<td>Nitrogen-fixation, ammonification, sulphate reduction, cellulytic decomposition</td>
</tr>
<tr>
<td>Fungi</td>
<td>Decomposition process</td>
</tr>
</tbody>
</table>

Table 19.2 Animal models used to predict toxic effects on human health

<table>
<thead>
<tr>
<th>Species</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicken, cat</td>
<td>Delayed type neurotoxicity</td>
</tr>
<tr>
<td>Armadillo</td>
<td>Genotoxicity (quadruplicate birth)</td>
</tr>
<tr>
<td>Various rodents, poultry</td>
<td>Nutritional effects of toxicant</td>
</tr>
<tr>
<td>Rats, mice</td>
<td>Genotoxicity (cancer, reproduction)</td>
</tr>
<tr>
<td>Monkeys</td>
<td>Behavioural toxicology</td>
</tr>
<tr>
<td>Swine</td>
<td>Cardiovascular and gastrointestinal effects</td>
</tr>
</tbody>
</table>

19.2.1.2 Bioaccumulation

Studies with birds and field studies were among the first activities that revealed the significance of food chain uptake and magnification of chemicals, both of which continue to be a strong focus of ecotoxicologic testing. Uptake by all routes (bioaccumulation) and the transfer of that accumulated chemical between species (biomagnification) are often confused. In terrestrial species, bioaccumulation is dependent on many factors, including:
Short-term Toxicity Tests for Non-genotoxic Effects

(1) exposure rate or concentration (Kenaga, 1972),
(2) properties of the chemical determining its stability, ability to pass through membranes, and extent of retention in tissues (Hansch, 1980); and
(3) pharmacodynamics of the chemical in the organism, in turn governed by physiological and biochemical factors such as blood flow and tissue- or organ-specific degradation or binding (Lindstrom et al., 1974; Bungay et al., 1980).

The physicochemical properties of a chemical are adequate to indicate potential bioaccumulation. The ability to predict bioaccumulation (Kenaga, 1980) partly depends on the octanol–water partition coefficient (Kow or P), which may be measured, or may be predicted from the chemical structure of the toxicant (Leo et al., 1971). The extent of bioaccumulation is proportional to P for values of P between $10^3$ and $10^6$. However, if the value of log P exceeds 6 (i.e. P exceeds $10^6$), bioaccumulation may be substantially less than predicted. Nevertheless, suspected bioaccumulation is readily diagnosed from structure, providing that covalent reaction in biota is not in question.

This latter exception is demonstrated with the example of methyl mercury. As a fungicide, methyl mercury had been tested in various formulations without notable problems being apparent. Based on its relatively high water solubility and low log P, it had not been suspected of bioaccumulation. Shortly thereafter, it was recognized that mobilization of mercury from sediments (in the form of methyl mercury) was the cause of Minimata disease, and environmental surveys revealed heavy residues of mercury in many game birds suspected of having fed on fungicide-dressed seeds. Further investigation showed that the methyl mercury, accumulated by covalent reaction with protein sulphydryl groups, was chronically neurotoxic, although not especially lethal in acute doses. Thus, animals are capable of developing heavier total mercury loadings from methyl mercury than inorganic forms (Gillett, unpublished observations).

With some notable exceptions, biodegradability may also decrease with increasing log P. One such exception are the easily photolysed and hydrolysed pyrethroid chemicals which are not bioaccumulated in spite of high values of Log P. This covariant relationship is also affected by species-specific enzyme nature and amount, so that full predictability for all species is unattainable.

19.2.2 Plants

Most of the concern for toxicity to plants has been for indirect or non-target phytotoxic effects of chemicals such as insecticides, and concern is generally low unless crop productivity is significantly decreased. Testing of plant responses has focused largely on germination and early seedling growth, by means of field observation of effects. Various efforts to improve the seed germination-early seedling growth assays have been successful, but practically no effort has been made to standardize the several methods in use.
Relatively little testing of lower plants, other than algae, has taken place, since these tend to be of neither economic importance nor pest species. Because of emphasis on crop and pest species (largely rooted, higher plants) in agriculture, relatively little information on non-agricultural species and on later life stages, interspecies interactions, succession and community structure is available.

One exception is the body of information developed around the sensitivity of lichens (fungal–algal associations) to air pollutants (Duffus, 1980). The distribution of lichens between and within communities has served as a convenient index of air pollution. However, laboratory testing has not been widely employed, in part due to technical difficulties.

Much of the phytotoxicology centres around interference with photosynthesis, both because of the primary energetic considerations and because of the uniqueness of this system in plants. Many herbicides (Audus, 1964) depend on inhibition of one or more reactions of the photosynthetic cycle for their effectiveness. Tschan et al. (1975) developed an especially sensitive test for photosynthetic inhibitors, using light emission by a marine bacterium activated by the oxygen generated by photosynthesis. Any toxicant affecting reduction of water to oxygen or decreasing algal photosynthetic and oxidative efficiency leads to lower light emissions. Although theoretically operable for general metabolic toxicants (interfering with any oxygen-producing or oxygen-utilizing reactions), the Tschan test appears most sensitive to photosynthetic inhibitors. Improvements and modifications have been made, but the test still requires elaborate and expensive equipment without offering any better insight into ecotoxicologic problems.

Seed germination tests only examine a brief part of a plant’s life cycle. Usually, the test provides an opportunity to examine early effects on morphology, susceptibility to pathogens and growth of root and hypocotyl. However, these observations are not always made or evaluated. Plant sensitivity to the action of chemical agents at later stages of its life cycle can be determined through field observation or greenhouse tests. The reproductive process (flowering, pollination, fruit formation and development), maturation, and senescence may also be affected by synthetic organic chemicals.

Attempts to establish a plant life cycle test using Arabidopsis spp. are underway (Tingey, personal communication). This small member of the wort family has a seed-to-seed time of around 30 days permitting it to be studied in botany classes. Its biology is thus well known, but its sensitivity to toxicants and the degree to which it may represent other, longer-lived species are in question. Furthermore, the mechanical difficulties in handling the very small seeds present a particular problem.

19.2.3 Microcosms

Given the status of single species testing, it is hardly surprising that multi-species tests of ecological effects, although considered most important, are just being developed and have only a small database to support their interpretation and use
(Cairns et al., 1981). The laboratory model ecosystem or microcosm represents the highest order of testing outside of the field. A variety of terrestrial, aquatic and mixed media systems have been developed and applied to the evaluation of various synthetic chemicals, mostly pesticides. The technology has been reviewed extensively (Gillett and Witt, 1979; Giesy, 1980; Hammonds, 1981; Van Voris et al., 1983a).

The microcosm is inherently safer and more easily manipulated than field sites, provides detail unattainable in the field, and is not subject to the vicissitudes of weather and geologic cataclysm. It permits rigorous testing of hypotheses developed from the laboratory (chemical or single species toxicity tests). While combining good features of both laboratory and field tests, microcosms have a variety of shortcomings including: never fully representing all ecological processes; being difficult to make self-sustaining with adequate complexity; and being fraught with methodologic difficulties. In a number of cases, the microcosm has been demonstrated as useful in assessing effects on primary productivity (Cole et al., 1976; Van Voris et al., 1983b), growth and reproduction (Gillett et al., 1983a), nutrient cycling (Van Voris et al., 1980), and interspecies interaction (Gillett et al., 1983a). Because these are meaningful ecotoxicologic end-points, they suggest that further development will be worthwhile.

### 19.3 FUTURE TRENDS FOR SHORT-TERM TESTING IN TERRESTRIAL ECOTOXICOLOGY

#### 19.3.1 General considerations

The trend towards the increasing use of short-term tests (especially *in vitro* tests) could result in a situation where the results of tests on whole animals (now used to support ecotoxicologic assessments as well as to assess the safety for human beings), might not be available. There is a significant implication in this since much as the single species test fails to provide integration of effects for the ecosystem (Cairns et al., 1981; Levin and Kimball, 1983), so the short-term *in vitro* test fails to provide integration of effects at the organism level. For example, *in vitro* test systems such as perfused rat liver or hepatocytes may reveal much about the effects of a chemical on liver but other important effects (e.g. behaviour, feeding efficiency, etc.) may not be evidenced.

The challenge, then, is to establish a testing system which minimizes testing resources, adverse environmental impact (caused by the test itself), and excessive use of laboratory test species while gaining better information capable of ecological integration. The sheer complexity of this task suggests that a variety of approaches will have to be combined.
19.3.2 Structure–activity relationships

For new chemicals (and inadequately tested older chemicals), much can be gained by analysis of known structure–activity relationships (SAR). Even when there are few or no data available from toxicity tests, it is possible to express the nature and extent of concerns for a chemical from simple relationships (Gillett, 1983) in order to develop testable hypotheses which must then be investigated by means of additional testing. Many physicochemical characteristics of a specific chemical can be predicted from consideration of its structure or from other measured physicochemical properties (Lyman et al., 1982). These estimated values can be employed in mathematical models (Neely, 1980) to estimate potential exposures. As part of a screening system, the estimation of physicochemical properties has already contributed much to simplifying assessment and testing. Unfortunately, as with all simplifications, much has been done without adequate background documentation.

Presently, SAR are based on statistical analysis of empirical data. For example, extensive investigation of bioaccumulation has provided substantial support for the use of SAR in predicting bioaccumulation potential. For interpolations, these SAR are quite accurate (error factors of only 2 or 10); however, extrapolation is less certain.

Acute toxicity for a given chemical in the laboratory rat can be predicted with some confidence if the acute toxicities of other members of the same class of compound are already known. However, chronic toxicity (where lethality is not the end-point) is much harder to predict. Numerous databases contain much information which would help bolster SAR efforts, but no suitable means of bringing together these data from around the world has been developed. Furthermore, there are questions about the quality of such data which have usually been generated over a decade or more.

The USEPA has instigated comprehensive acute and chronic toxicity databases for wildlife (TERRATOX) and plants (PHYTOX) to accompany the aquatic database AQUIRE and the chemical properties databases CHEMFATE, DATALOG and BIOLOG in the master program SPHERE (Miles et al., 1983). Based on material in peer-reviewed, open literature and further screened for quality, these systems will be heavily employed in evaluation of new chemicals under the Toxic Substances Control Act and other legislation. Use of sophisticated computerized structural connectivity indices (UNICORN) will further enhance data accessibility. The LOG P DATABASE™ (Technical Database Services, Inc.) contains measured and evaluated partitioning data on over 5000 chemicals. Other commercial endeavours are largely oriented toward human health.

Subsequently, it may be possible to estimate unmeasured values through a number of multivariant analyses and other statistical techniques. Hudson et al. (1979) showed that acute (single dose) and subchronic (5-day dietary exposure plus 3-day feed-off) toxicity of chemicals to birds could be estimated from rat data within one order of magnitude. The degree to which such computerized SAR
devices will substitute for actual testing will depend on the acceptability of this degree of error.

Although there are many relationships identified for SAR at the enzymatic level and a number at the species level, no such relationships have been developed for functions at levels of biological organization above the organism. The Pre-Biologic Screen (Gillett, 1983) uses a combination of log \( P \) (octanol–water partition coefficient), log (Henry’s law constant), and log (biodegradation half-life) to pose testable hypotheses about a chemical in regard to ecotoxicologic concerns (bioaccumulation and chronic action; multi-species/multi-media involvement; chronic action in the water column, including leaching and plant uptake; and indirect effects due to atmospheric action). Concerns are ranked as high, moderate, and low (or negligible). While providing a useful means of ranking chemicals in relation to potential adverse effects, these methods do not constitute a guarantee of actual effects or serve as a surrogate for actual data.

It must be emphasized that testing is only one of the links in a complete assessment (Goss and Wyzga, 1982). Unnecessary testing has costs to the regulatory agency, the public and, of course, the manufacturer. Devising unambiguous criteria is an important objective so that much thinking has gone into the interpretation of schemes employing various SAR screens. Even so, for non-mammalian species, we are still very far away from predicting end-points other than death or genotoxic responses.

19.3.3 Generalized toxicity tests

The initial hope that these tests might serve as surrogates for predicting effects on higher organisms has not been borne out in several investigations. In the past two decades, several microbial tests have been proposed as means of estimating toxicity of individual chemicals, complex effluents and other mixtures. The Microtox™ test (Beckman Instruments, Inc., 1980) employs a marine bioluminescent bacteria which releases light through normal metabolism; toxic chemicals reduce luminescence (Dutka and Kwan, 1981). The Tschan test (described earlier) assesses phytotoxic chemicals acting through inhibition of the Hill reaction in photosynthesis. Respiratory inhibition (Liu, 1981; Bauer et al., 1981) is more readily measured at a lower cost and without elaborate equipment (Gillett et al., 1983b). Analyses of respiratory inhibition have shown (Dutka and Kwan, 1981; Gillett et al., 1983b) that measurements were primarily affected by the size of the active microbial population rather than by inhibition of enzymes. Mixed population sources of organisms cultured under equivalent conditions were not statistically different, but selected species (pure cultures) might be more or less sensitive (Bauer et al., 1981).

Now the question changes from, ‘Are such generalized assays applicable to higher species, levels of biological organization, etc.? ’ to ‘Are such \textit{in vitro} tests applicable to even the tested species in the field?’ Various \textit{in vitro} studies have not yielded responses relevant to field results. The \textit{in vitro} tests were either too sensitive
or exposure was radically different in the field. In a microcosm study of 3,4-
dichlorophenol (to which the bacteria of soil and sewage sludge were expected to be
'naive'), the \textit{in vitro} EC50 was 10 to 20 mg/litre for pure and mixed cultures from a
variety of sources (Gillett \textit{et al.}, 1983b). However, soil respiration was uninhibited
by 1000 p.p.m. (mg/kg dry soil) (Gillett \textit{et al.}, unpublished results). The test chemi-
cal appeared to be both adsorbed and tightly bound to the soil in an unextractable
form. The unextractable portion increased in proportion and extent as the proportion
of organic matter increased between soil series, but also declined with time, releas-
ing free (extractable) 3,4-DCP.

Thus, our inability to describe the chemodynamics of chemicals in sufficient
detail, as to ascertain the specifics of exposure, limits laboratory-to-field extrapola-
tion and the applicability of generalized toxicity assays. Much more work is needed
to validate the capacity of all these toxicity tests to predict potential effects under
field conditions.

Use of other species in generalized assays has not produced acceptable methods
either. \textit{Daphnia} spp., houseflies, earthworms (especially \textit{Eisenia foetida}), honeybees
and other invertebrates can be assayed accurately, reproducibly, and sensitively for
numerous chemicals. On the other hand, each species or group is insensitive to
certain classes of chemicals. Therefore, the effects noted in one species are not
necessarily applicable or able to be extrapolated to other species. Lack of detailed
research at higher levels of biological organization over a sufficient range of chemi-
cal classes, exposure scenarios, etc., precludes the general use of single species tests
as indicators of ecotoxicologic concerns.

19.3.4 Microcosm studies

Microcosm tests are the principal alternative to single species tests for ecotoxi-
ologic effects. These are expected to gain in value and significance as they are
widely applied to environmental problems (Gillett and Witt, 1979). Initially, this
type of study suffered from relatively high capital and operating costs and from
heavy requirements for professional expertise. Recent efforts to develop standard-
ized protocols have resulted in evidence that simple microcosms are cost-effective
means of examining complex situations. Improvements in multi-seasonal operation
of terrestrial systems indicate the ability to shorten the testing period. However,
these same studies have not provided support for the earlier idea of microcosms as
broad screening tools.

Further development of microcosm screening assays is likely to occur. Very small
soil core microcosms (10 cm in diameter and 10 cm deep) (Draggan, 1976) are too
highly variable (Gile \textit{et al.}, 1979) for much use in terrestrial tests, but the soil litter
microcosm (Lighthart \textit{et al.}, 1982) of approximately 100 g of soil in a glass canning
jar, and larger soil cores (with or without intact plant communities) of 15 cm
diameter and 60 cm depth (Van Voris \textit{et al.}, 1984) have suitable sensitivity, low
variability and cost for some screening and confirmation uses.
Microcosms for agroecosystems can be readily and rationally constructed and operated under conditions that provide for substantial management decision-making power in assessments; however, there is a lack of understanding about the sensitivity and operation of microcosms representing non-agricultural systems. Although many mathematical models of ecosystem function have been constructed, few analyses of sensitivity of processes within systems or between systems have been performed relative to organic chemical insults. In part, this is due to deficiencies in knowledge of ecosystem science as applied to environmental impacts, but also to operational difficulties in developing and maintaining microcosms representing certain system (wetlands, forest systems) or general lack of interest in impacts on certain ecosystems (desert biomes, for example).

19.4 CONCLUSIONS

Terrestrial ecotoxicologic assessments are not likely to be enhanced by short-term in vitro toxicity test methods that depend on effects at a lower order of biological integration than currently practised in toxicologic studies. Present methods are already under criticism (Levin and Kimball, 1983) for inadequacies regarding representativeness, completeness, meaningful sensitivity, statistical validity, and quality assurance.

Terrestrial ecotoxicologic assessments currently utilize much information generated in the course of human health assessments, especially mammalian toxicity data. Two other approaches seem useful in enhancing currently available techniques and tests:

1. The development and improvement of SAR methods, which can shorten and focus testing requirements, obviating all but those needed to confirm or discriminate particular problems. Development to extend those methods from the sub-organismal and organismal levels to higher levels of biological organization is required. As computer-accessible databases are developed and organized, attention to quality assurance of data and other details are suspected to be important.

2. Microcosm technology, which needs a broader database with more chemical classes and types of observations of higher level functions within the model ecosystems. Validation of the applicability of this technology in particular assessments is needed, as well as further system improvements and broader representation of ecotypes.

Because of the potential cost-effectiveness and incisiveness of microcosm studies, particularly when used in conjunction with SAR techniques to establish testable hypotheses of adverse action, this technology may lead to reduced resource needs and costs, the basic objective of short-term testing.
REFERENCES

Short-term Toxicity Tests for Non-genotoxic Effects


