

European and Mediterranean Plant Protection Organization
Organisation Européenne et Méditerranéenne pour la Protection des Plantes

Evaluation biologique des produits phytosanitaires

Mise en place et analyse des essais d'évaluation biologique

Champ d'application spécifique

Cette norme destinée à être utilisée en association avec les Normes OEPP de la série PP 1 (Normes pour l'évaluation biologique des produits phytosanitaires) fournit des conseils détaillés sur la mise en place et l'analyse des essais d'évaluation biologique.

Introduction

Cette norme a pour but de fournir des renseignements d'ordre général concernant la mise en place et l'analyse des essais pour l'évaluation biologique des produits phytosanitaires pour lesquels les Normes OEPP pour l'évaluation biologique des produits phytosanitaires donnent des instructions plus détaillées pour chaque combinaison hôte-ravageur. La mise en place d'un essai est étudiée en premier lieu (dispositif expérimental, disposition et taille des parcelles, rôle et implantation des témoins non traités). La nature des observations à réaliser est ensuite examinée (types de variables, modes d'observation). Enfin, des suggestions sont faites sur l'analyse statistique des résultats d'un essai et d'une série d'essais (estimation des effets, choix du test statistique, transformation des variables). Des exemples d'échelles utilisées dans les normes OEPP sont présentés à l'annexe I.

Les sections suivantes visent à esquisser la bonne pratique statistique dans l'analyse de données. Il ne s'agit pas, et il ne peut pas s'agir, d'une prescription valable pour toutes les analyses et toutes les situations. Les expérimentateurs ne doivent jamais sous-estimer l'utilité de l'avis d'un statisticien. Il est très important que les expérimentateurs comprennent l'avis qu'on leur donne et il est souvent préférable qu'ils effectuent une analyse simple qu'ils peuvent exposer et défendre avec confiance plutôt que d'accepter des conseils conduisant à une analyse qu'ils ne comprennent que partiellement. La bibliographie à la fin de cette norme peut être utile. Elle mentionne plusieurs textes qui présentent les principes de la bonne pratique statistique plutôt que de donner des recettes statistiques à appliquer aveuglément.

1. Dispositif expérimental

1.1 Champ d'application et objectifs expérimentaux

Avant de choisir un dispositif expérimental, le champ d'application et les objectifs de l'essai doivent être clairement définis car ils peuvent restreindre le choix des dispositifs envisageables. Dans la pratique, on utilise souvent un processus itératif : le champ d'application et les objectifs sont progressivement ajustés pour s'adapter aux ressources expérimentales

Approbation et amendements spécifiques

Approbation initiale en 1989-09.

Première révision approuvée en 1998-09.

Deuxième révision approuvée en 2006-09.

Troisième révision (principalement pour refléter l'évaluation zonale) approuvée en 2012-09.

disponibles. Il est fondamental que le champ d'application et les objectifs soient revus pour refléter les décisions prises au cours de ce processus.

Le champ d'application de l'essai donne l'ensemble des résultats pratiques qui peuvent être obtenus par l'essai et qui sont pertinents pour les objectifs fixés. Le champ d'application détermine, d'une part, la population qui sera échantillonnée par l'essai et d'autre part, l'ensemble des conditions environnementales, les cultures, les traitements chimiques, les méthodes d'application et les organismes cibles qui doivent être testés par l'essai. Le champ d'application définit le contexte dans lequel les unités expérimentales et les observations seront étudiées.

Les objectifs de l'essai doivent se présenter sous forme de questions sur les traitements auxquelles des réponses sont souhaitées. Les réponses typiques seront "oui" ou "non", un classement des traitements ou une estimation d'une valeur.

Le champ d'application et les objectifs doivent figurer dans le rapport de l'essai, comme décrit dans la Norme OEPP PP 1/181 *Conduite des essais d'évaluation biologique et présentation des rapports*. Les méthodes expérimentales, le dispositif et l'analyse prévus (voir ci-dessous) doivent également figurer dans le rapport.

1.2 Types de dispositifs

Les Normes OEPP pour l'évaluation biologique des produits phytosanitaires prévoient que les traitements expérimentaux concernent les "produit(s) à étudier, produit(s) de référence, et le témoin non traité; parcelles disposées selon un dispositif statistique adéquat". Il est également envisagé que les produits soient étudiés à des doses et/ou à des dates d'application différentes. Ceci concerne surtout l'utilisation d'une dose plus élevée dans les essais de sélectivité et d'études de dose en général.

Les dispositifs monofactoriels conviennent pour les essais si le(s) produit(s) à étudier, le(s) produit(s) de référence et le témoin non traité peuvent être considérés comme plusieurs valeurs d'un même facteur, et qu'aucun autre facteur ne doit être utilisé. Par contre, si, par exemple, l'effet de chaque produit dans un essai

d'évaluation biologique doit être étudié à différentes doses, un dispositif factoriel doit être utilisé, avec en général toutes les combinaisons de traitements possibles pour les deux facteurs représentés. Des interactions importantes entre les facteurs peuvent ainsi être mises en évidence et estimées.

Les principaux dispositifs aléatoires susceptibles d'être utilisés sont le dispositif complètement aléatoire et le dispositif en blocs aléatoires complets. Ces dispositifs sont illustrés ci-dessous à partir d'un exemple monofactoriel comprenant 8 traitements (5 produits à étudier, 2 produits de référence et un témoin non traité) et 4 répétitions.

1.2.1 Dispositif complètement aléatoire (randomisation totale)

Dans un dispositif complètement aléatoire (Fig. 1), les traitements sont attribués au hasard à l'unité expérimentale. Ce dispositif est potentiellement le plus puissant du point de vue statistique (probabilité maximale de détecter une différence significative si elle existe), parce qu'il permet de conserver le maximum de degrés de liberté pour la variance résiduelle. Il n'est cependant utilisable que si l'on est certain que le site de l'essai est très homogène. Si l'hétérogénéité entre les diverses parties du site d'essai devient importante, la variance résiduelle aura un niveau inacceptable. Il convient alors d'avoir recours à un dispositif qui tient compte de cela, tel qu'un dispositif en blocs aléatoires complets.

2	7	3	7	8	3	5	4
1	2	6	2	2	3	4	6
8	4	5	4	6	8	1	5
1	5	7	8	1	7	3	6

Fig. 1. Dispositif complètement aléatoire. Chaque traitement (numéroté de 1 à 8) est répété quatre fois; les traitements sont attribués complètement au hasard aux 32 parcelles

1.2.2 Dispositif en blocs aléatoires complets (blocs randomisés)

Un bloc est un groupe de parcelles dans lesquelles l'environnement à observer est homogène. Dans ce dispositif, les blocs sont créés délibérément pour que les parcelles qui les composent soient aussi uniformes que possible avant d'appliquer les traitements. En général, chaque traitement figure une et une seule fois dans chaque bloc. Les traitements sont distribués de manière aléatoire à l'intérieur des blocs, qui jouent le rôle de répétitions. L'attribution des traitements dans chaque bloc doit être aléatoire et être faite séparément pour chaque bloc. Dans les exemples ci-dessous (Figures 2-4), il y a 4 blocs et 8 traitements. La disposition des blocs a pour but de contrôler l'hétérogénéité du site (par ex. pente, direction du travail lors du semis ou de la plantation, exposition, degré d'infestation, etc.), des

végétaux (taille, âge, vigueur) ou des conditions lors de l'essai (application des traitements, notations). La disposition des blocs nécessite donc une connaissance préliminaire du site d'essai. La forme des parcelles peut influencer leur distribution à l'intérieur des blocs : les parcelles longues et étroites sont souvent placées côte-à-côte, alors que les parcelles carrées peuvent être disposées différemment. Les blocs, quant à eux, ne sont pas forcément adjacents. Une bonne connaissance préliminaire d'un champ peut être utilisée pour disperser les blocs dans l'ensemble du champ pour tenir compte d'une hétérogénéité observée précédemment (Figures 5-6). Dans une disposition aléatoire, il arrive que les traitements au sein d'une répétition apparaissent dans l'ordre du protocole, ceci est dans la mesure du possible à éviter pour ne pas introduire de biais dans les évaluations. On peut tenir compte d'une hétérogénéité complexe si l'on connaît bien le terrain et qu'on peut être confiant que les conditions vont sûrement se maintenir pour l'expérimentation à venir; il n'est alors même plus nécessaire que les parcelles d'un même bloc soient adjacentes. Par exemple, les blocs peuvent être éclatés pour tenir compte d'une infestation irrégulière de nématodes. Dans la figure 6, les parcelles du bloc 1 ont été placées délibérément à des points d'infestation visiblement faible, et les parcelles du bloc 2 à des points d'infestation visiblement forte.

Le choix du dispositif, et le cas échéant la taille et l'orientation des blocs utilisés, dépendent de l'hétérogénéité perçue dans le site d'essai (par ex. sol, pente, exposition, infestation par un organisme nuisible, cultivar, etc.). Ces paramètres ne sont jamais totalement uniformes, et un dispositif en blocs aléatoires complets dans une zone modérément hétérogène fournit généralement davantage d'informations utiles sur la performance d'un produit qu'un dispositif complètement aléatoire dans un site qu'on pense homogène mais qui par la suite se révèle hétérogène. La disposition des blocs dépend également de la dimension et de la forme des parcelles (Figures 5, 6). Les petits blocs sont généralement plus efficaces pour réduire l'hétérogénéité. Dans les essais comportant un nombre important de traitements, d'autres dispositifs doivent être considérés (par ex. dispositifs en treillis, dispositifs en blocs incomplets).

Des essais en blocs aléatoires conduits dans plusieurs régions à conditions agroclimatiques distinctes et/ou au cours de différentes années ou périodes de végétation peuvent parfois être considérés comme faisant partie d'une série d'essais. Dans l'analyse statistique, il est alors nécessaire de distinguer la variance entre sites de la variance entre blocs, et également d'évaluer l'interaction site-traitement qui pourra être particulièrement intéressante. Pour chaque nouvel essai, les traitements doivent de nouveau être attribués aléatoirement dans chaque bloc.

Bloc 1	3	8	7	2	5	4	6	1
Bloc 2	4	7	5	1	6	2	8	3
Bloc 3	5	6	7	2	8	3	1	4
Bloc 4	8	4	1	3	5	6	7	2

↓ gradient

Bloc 1	Bloc 2	Bloc 3	Bloc 4
5 7 1 2 8 4 3 6	4 6 1 5 3 8 2 7	3 8 2 5 4 7 6 1	2 3 1 8 5 6 7 4

Fig. 2. Disposition possible des blocs et des parcelles pour des blocs aléatoires dans un essai de plein champ. L'environnement présente un gradient le long du champ et on en tient compte, soit en disposant les blocs dans le sens du gradient, soit en les plaçant côte à côte. Dans chaque cas, les parcelles au sein de blocs disposés perpendiculairement au gradient sont influencées de la même manière par la variable environnement.

Bloc 1	Bloc 2	Bloc 3	Bloc 4
3 1	8 1	8 2	3 7
6 4	2 6	6 5	1 6
8 5	7 5	3 1	5 8
7 2	3 4	7 4	4 2

Bloc 1	Bloc 2	Bloc 3	Bloc 4
1 3 2 4	2 7 8 5	4 1 6 3	7 3 5 1
8 5 6 7	4 6 1 3	5 7 8 2	4 8 2 6

Fig. 3. Disposition possible des blocs et des parcelles pour des blocs aléatoires dans un essai de plein champ. Alternative pour un dispositif en blocs aléatoires dans le cas où il n'existe pas de gradient environnemental clair mais la distance maximale entre les parcelles dans un même bloc est relativement grande et laisse supposer une hétérogénéité. Les 8 parcelles sont relativement proches dans un rectangle de 4 par 2, et les blocs sont côte à côte.

Bloc 1	2 7 3 1	8 5 2 7	Bloc 2
	8 5 4 6	6 3 4 1	
Bloc 3	3 6 8 7	6 3 5 2	Bloc 4
	1 4 5 2	7 4 8 1	

Fig. 4. Un autre exemple de disposition des blocs et des parcelles quand, comme dans la Fig. 3, une hétérogénéité est supposée mais qu'il n'existe pas clairement de gradient environnemental. De nouveau, les 8 parcelles sont relativement proches dans un rectangle de 4 par 2, mais les blocs sont disposés selon un rectangle de 2 par 2.

Bloc 1	5 3 8 2 4 7 1 6	Bloc 2	6 8 5 7 3 1 4 2
Bloc 3	4 7 1 6 2 8 3 5		
Bloc 4	3 4 6 8 7 5 1 2		

Fig. 5. Disposition possible des blocs et des parcelles pour des blocs aléatoires dans un essai de plein champ. Blocs dispersés dans l'ensemble du champ, selon une hétérogénéité observée précédemment.

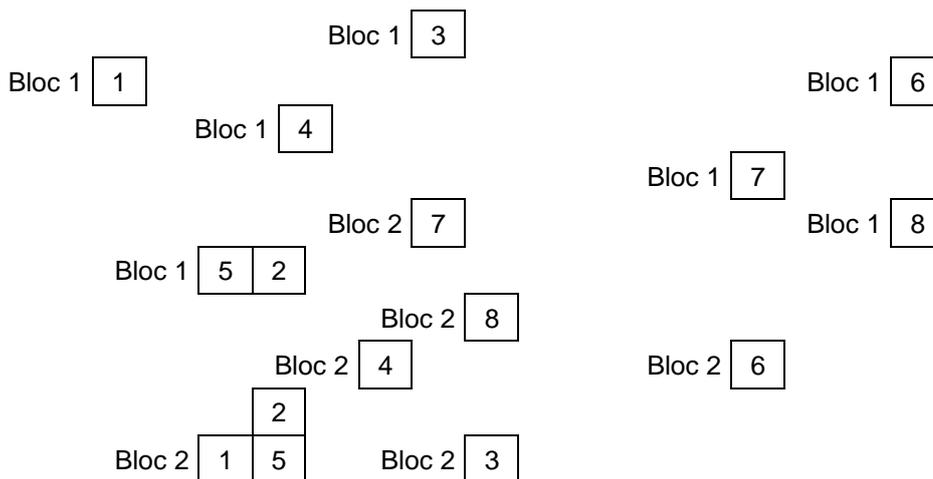


Fig. 6. Disposition possible des blocs et des parcelles pour des blocs aléatoires dans un essai de plein champ. Blocs disposés dans le champ en fonction d'une hétérogénéité complexe observée précédemment.

Bloc1					Bloc 2			
1A	2A	3A	4A	parcelle entière 1	2B	4B	3B	1B
3C	4C	1C	2C	parcelle entière 2	2A	3A	4A	1A
2B	3B	1B	4B	parcelle entière 3	1C	3C	2C	4C

Fig. 7. Exemple de dispositif en parcelles divisées. Les deux facteurs du traitement sont : le produit (1,2,3,4, randomisés dans les sous-parcelles des parcelles entières) et la méthode de culture (A,B,C, randomisés dans les parcelles entières de chacun des deux blocs).

1.2.3 Dispositif en parcelles divisées (split-plot)

Lorsqu'un essai multifactoriel est conduit, on utilise généralement un dispositif en blocs aléatoires complets dans lequel chaque combinaison de traitements apparaît une fois dans chaque bloc. Cependant, il arrive qu'on ne puisse pas randomiser totalement un des facteurs dans toutes les parcelles d'un bloc. Par exemple, supposons un essai à deux facteurs, le produit (avec quatre niveaux, de 1 à 4) et le matériel de culture (avec trois niveaux, A, B, C), et supposons que les parcelles sont relativement petites. La taille du matériel utilisé pour l'opération culturale en question peut ne pas permettre une randomisation totale sur les douze parcelles de chaque bloc. Un dispositif en parcelles divisées est alors recommandé : les sous-parcelles sont associées, dans chaque bloc, par groupes de quatre pour former trois parcelles entières par bloc. Le facteur "opération culturale" est randomisé dans ces trois parcelles, et le facteur "produit" est randomisé, séparément, dans les sous-parcelles de chaque parcelle entière (Fig. 7). Ce dispositif nécessite une analyse de variance légèrement plus complexe, à deux strates, chacune ayant un carré des écarts moyen que l'on peut utiliser pour tester l'effet des différents facteurs et leur interaction.

1.2.4 Dispositifs systématiques

Les dispositifs systématiques non randomisés ne conviennent jamais pour les essais d'évaluation biologique, sauf dans des cas très particuliers (par exemple essais variétaux sur la sélectivité des herbicides). En règle générale, ce type de dispositif n'est utilisable qu'à des fins de démonstration.

1.3 Puissance des tests

Lors de la mise en place d'essais, il est important de considérer la puissance requise pour tous les tests statistiques qui seront effectués. La puissance est la probabilité de détecter une différence entre les traitements si cette différence existe. La puissance dépend d'un certain nombre de paramètres, dont:

- la précision des résultats (variation résiduelle)
- le nombre de répétitions, y compris les répétitions sur les sites.

Il faut choisir un dispositif qui donne une bonne chance de détecter, avec une signification statistique, une différence importante dans la pratique pour le contraste étudié. On peut aussi souhaiter que les intervalles de confiance des estimations des traitements ne dépassent pas un écart préalablement défini. Avant d'entreprendre un essai, il faut choisir entre la performance d'un seul essai ou d'une série d'essais.

Selon la Norme OEPP PP 1/226 *Nombre d'essais recommandés*, la performance d'un produit phytosanitaire doit être démontrée en conduisant un certain nombre d'essais dans différents sites, régions, au cours de différentes années dans des conditions environnementales variées. Ainsi, pour évaluer la performance d'un produit phytosanitaire, une série d'essais doit être planifiée, conduite et analysée (voir aussi 3.4.1. pour une définition d'une série d'essais).

En général, des résultats d'essais précédents peuvent indiquer la variabilité probable des observations. Si ces données existent, il est possible de déterminer le dispositif et la dimension de l'essai nécessaires pour la puissance requise. Il est parfois possible de le faire à partir de considérations théoriques. Par exemple, pour des données suivant une loi binomiale, une limite supérieure peut être déterminée pour la variabilité des proportions. Différents systèmes informatisés ou graphiques sont disponibles pour aider à déterminer le nombre de répétitions nécessaires. Ils utilisent l'amplitude de la différence devant être estimée, ou le niveau de signification exigé pour cette différence, et la précision attendue. La section suivante donne des règles générales simples.

1.4 Nombre de degrés de liberté en fonction du nombre de traitements et de répétitions

Pour permettre une analyse statistique valide, le nombre de degrés de liberté (dl) doit être suffisamment élevé. Dans un essai avec 8 traitements et 4 répétitions dans un dispositif en blocs aléatoires, il y a 21 dl résiduels. Ce résultat est obtenu comme suit : dl totaux ($32 - 1 = 31$) moins dl traitements ($8 - 1 = 7$) moins dl blocs ($4 - 1 = 3$), c'est-à-dire $31 - 7 - 3 = 21$. Dans un essai avec 3 traitements et 4 répétitions répétés sur 4 sites, il y a 24 dl résiduels. Ce résultat est obtenu comme

suit : dl totaux ($48 - 1 = 47$) moins dl traitements ($3 - 1 = 2$) moins dl sites ($4 - 1 = 3$) moins dl interaction traitement-site ($(3-1)*(4-1)=6$) moins dl répétition sur les sites ($(4-1)*4=12$), c'est-à-dire $47-2-3-6-12=24$.

Le nombre de dl résiduels peut être augmenté en augmentant le nombre des répétitions, de traitements ou de sites. Le nombre de dl résiduels souhaitable dépend du degré de précision recherché (puissance); en cas de doute, obtenir l'avis d'un spécialiste. En général, l'expérience des essais/séries d'essais d'évaluation biologique montre qu'il est déconseillé de mettre en place un dispositif avec moins de 12 dl résiduels. Si pour toute raison valable, il est recommandé d'utiliser seulement 3 répétitions et 3 traitements, alors l'essai sera réalisé sur au moins 4 sites (donnant dl résiduels) en vue d'obtenir un minimum de 15 degrés de liberté résiduels requis pour une analyse statistique valide.

Le choix du dispositif expérimental influence également le nombre de dl résiduels. C'est le dispositif complètement aléatoire qui donne le maximum de dl. Le dispositif en blocs aléatoires utilise certains de ces dl pour tenir compte de l'hétérogénéité du milieu (par ex. selon un gradient). Le dispositif en parcelles divisées utilise des dl pour tenir compte de plusieurs composants de variation éventuels. L'expérimentateur doit essayer de conserver le nombre maximal de dl pour estimer la variation résiduelle, tout en choisissant un dispositif optimal permettant de minimiser cette variation en tenant compte de toutes les sources connues d'hétérogénéité (voir Norme OEPP PP 1/181).

La relation entre le nombre de répétitions et de degrés de liberté résiduels pour différents nombres de traitements et de sites peut être extraite du Tableau 1.

Tableau 1. Nombre de degrés de liberté résiduels en fonction du nombre de sites, de traitements et de répétitions dans un site

Sites	1						4						6					
	3	4	5	6	7	8	3	4	5	6	7	8	3	4	5	6	7	8
Répétitions	3	4	5	6	7	8	3	4	5	6	7	8	3	4	5	6	7	8
Traitements																		
3	4	6	8	10	12	14	16	24	32	40	48	56	24	36	48	60	72	84
4	6	9	12	15	18	21	24	36	48	60	72	84	36	54	72	90	108	126
5	8	12	16	20	24	28	32	48	64	80	96	112	48	72	96	120	144	168
6	10	15	20	25	30	35	40	60	80	100	120	140	60	90	120	150	180	210
7	12	18	24	30	36	42	48	72	96	120	144	168	72	108	144	180	216	252
8	14	21	28	35	42	49	56	84	112	140	168	196	84	126	168	210	252	294

1.5 Unités/parcelles expérimentales : taille, forme, nécessité de bordures

L'unité expérimentale est cette partie du matériel expérimental à laquelle un seul traitement est appliqué

et sur laquelle des observations sont réalisées. Un nombre suffisant d'unités est nécessaire pour les traitements et répétitions planifiés. Dans la pratique, le matériel expérimental étant limité, il est souvent nécessaire de faire des compromis. Comme exemples

d'unités expérimentales, on peut citer : une surface de culture (parcelle), un pot contenant un ou plusieurs végétaux, un seul végétal, une partie d'un végétal (par ex. feuille, tige, branche) et un point d'appâtage dans un champ. Les unités expérimentales doivent être représentatives de la population testée par l'essai et être aussi uniformes que possible. Un manque d'uniformité peut parfois être atténué par des répétitions.

En règle générale, les parcelles seront rectangulaires et de même taille dans un essai individuel et de taille similaire dans une série d'essais. La précision augmente avec la dimension de la parcelle, mais seulement dans une certaine limite, car la variabilité des conditions du sol et d'infestation tend également à augmenter. Des parcelles rectangulaires longues et étroites conviennent bien pour la récolte mécanisée. Des parcelles presque carrées diminuent le risque d'interférence entre les parcelles. Pour les observations d'organismes nuisibles agrégés dans l'espace, comme certaines adventices ou maladies transmises par le sol, mieux vaut beaucoup de petites parcelles que peu de grandes parcelles.

Dans les normes spécifiques de l'OEPP, les dimensions des parcelles sont préconisées pour les combinaisons particulières d'organisme nuisible/plante-hôte. Dans le cas où les interférences entre parcelles sont probables, les parcelles auront une dimension plus grande (parcelle brute) et les observations se feront seulement sur une partie centrale plus restreinte (parcelle nette). La différence entre la parcelle nette et la parcelle brute est appelée zone de garde. En règle générale, les normes de l'OEPP suggèrent des dimensions de parcelle nette, la dimension de parcelle brute étant alors laissée à l'appréciation de l'expérimentateur qui doit déterminer quelles sont les zones de garde nécessaires en considérant toutes les sources potentielles d'interférences entre les parcelles pour chaque essai ou série d'essais. Une source courante d'interférences est la dispersion du produit phytosanitaire (par ex. dérive des produits pulvérisés ou des vapeurs, mouvements latéraux sur/dans le sol) en dehors de la parcelle, ce qui contamine les parcelles adjacentes. Ce phénomène est particulièrement important pour les pulvérisations appliquées sur de grandes plantes. Cependant, avec des zones de garde plus importantes, l'erreur expérimentale peut souvent être réduite au minimum.

Une autre source courante d'interférences est la dispersion des organismes nuisibles (par ex. champignons portés par le vent ou insectes très mobiles) à partir des parcelles non traitées ou des parcelles où la lutte est peu efficace. Cette dispersion peut à la fois augmenter la population de l'organisme nuisible dans les parcelles des traitements les plus efficaces et la diminuer dans les parcelles les moins efficaces. De même, si un produit phytosanitaire est étudié sur une culture dans laquelle la lutte intégrée est pratiquée, les effets négatifs sur les prédateurs et les parasites peuvent être masqués par leur migration entre les parcelles.

La compétition pour la lumière et les nutriments représente une source d'interférences supplémentaire. C'est particulièrement le cas lorsqu'il s'agit de mesurer le rendement. Si les zones de garde entre les parcelles sont différentes des parcelles elles-mêmes (par ex. allées nues, culture différente), il faut faire preuve de prudence lors de la sélection du site à évaluer.

Selon le matériel utilisé pour l'application ou la récolte, il peut être nécessaire d'augmenter la dimension de la parcelle nette au-delà de la dimension nécessaire pour effectuer les observations.

Les parcelles peuvent être disposées perpendiculairement ou parallèlement à la direction de travail (semis ou plantation). Le dispositif perpendiculaire (Fig. 8) présente l'avantage d'exposer uniformément toutes les parcelles d'un bloc à une erreur éventuelle pendant le travail (façons culturales, semis, etc.). Par contre, le traitement et la récolte sont plus difficiles. Le dispositif parallèle présente des avantages pratiques pour l'application des traitements et la récolte, mais il présente un plus grand risque d'hétérogénéité le long de blocs très longs. Le dispositif hybride peut éventuellement faire office de compromis.

1.6 Rôle et implantation des témoins non traités

1.6.1 Utilité du témoin non traité

La principale caractéristique des "témoins non traités" est qu'ils ne reçoivent aucun des traitements phytosanitaires en cours d'étude. Les témoins reçoivent toutes les interventions qui sont appliquées de façon uniforme à l'ensemble de l'essai, en particulier les façons culturales et les traitements visant les organismes nuisibles non étudiés. Même si le témoin non traité ne reçoit normalement aucun traitement contre l'organisme à étudier, dans certaines circonstances, il peut être utile d'inclure dans le témoin non traité certaines opérations appliquées aux autres traitements. Par exemple, lorsque les autres traitements reçoivent les produits dans une solution aqueuse distribuée par un pulvérisateur dans la parcelle, le témoin non traité peut inclure le passage d'un pulvérisateur contenant uniquement de l'eau. Il s'agit d'imiter autant que possible les opérations réalisées dans les autres traitements, à l'exception de l'application du produit.

L'objectif principal du témoin non traité est de démontrer la présence d'un niveau d'infestation suffisant. Par exemple, si le témoin non traité n'a pas confirmé la présence d'un niveau d'infestation suffisant, l'efficacité ne peut pas être démontrée et les résultats n'ont alors pas de sens. Cette confirmation peut être qualitative (présence d'espèces dominantes, type de flore, adventices...) ou quantitative (respect de seuils maximaux ou minimaux, répartition dans l'espace). Des circonstances exceptionnelles peuvent empêcher la mise en place d'un témoin non traité (par ex. organismes de quarantaine).

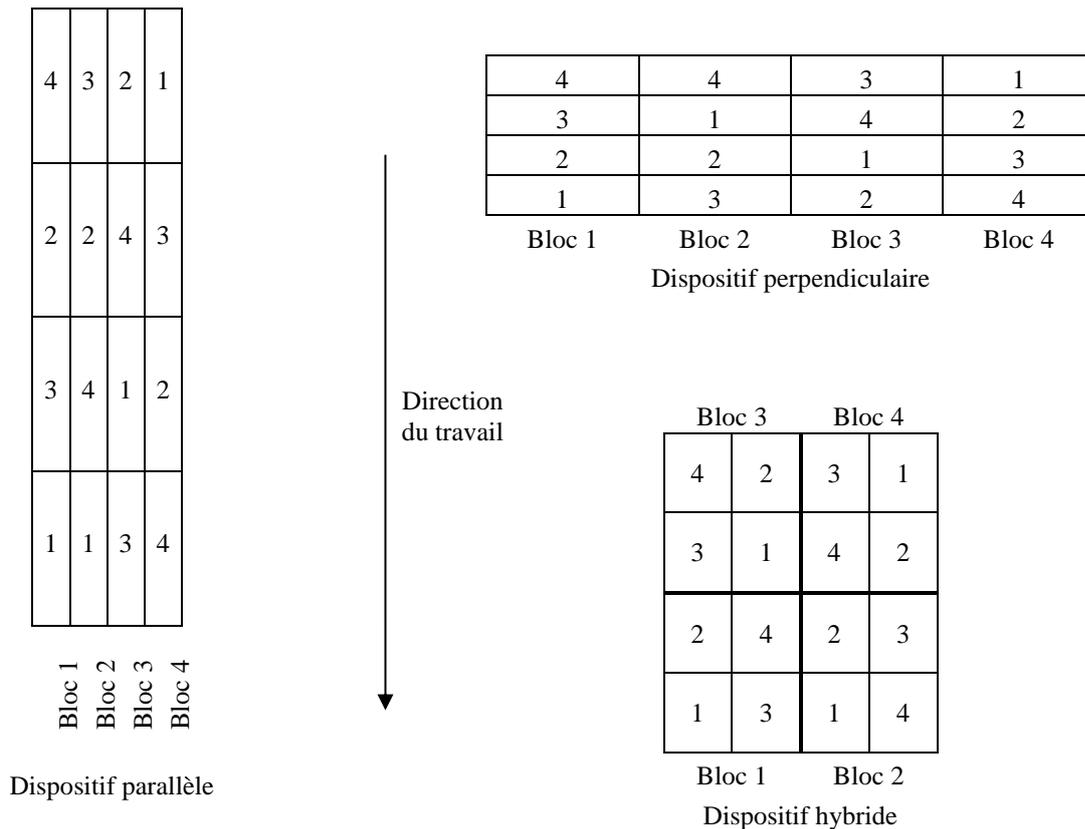


Fig. 8. Dispositifs en blocs aléatoires similaires, mais avec une disposition des parcelles différente par rapport à la direction du travail.

Selon l'objectif et le type d'essai, les témoins non traités jouent un rôle utile, et parfois plusieurs rôles à la fois, dont les suivants :

- montrer l'efficacité d'un nouveau produit et du produit de référence. La preuve primaire de l'efficacité d'un nouveau produit ou d'un produit de référence est toujours obtenue à partir d'une comparaison avec le témoin non traité.
- une aide pour les observations. Une estimation visuelle des dégâts ou de l'infestation peut parfois être effectuée en termes relatifs par rapport à un témoin.
- l'utilisation de la technique du "témoin adjacent" pour mesurer et tenir compte de la répartition des parcelles dans l'espace.
- l'observation du développement de l'organisme nuisible (émergence, envol, libération de spores, etc.), en particulier pour déterminer les dates d'application ou d'observation.
- une réserve d'inoculum pour assurer que le niveau d'inoculum ne tombe pas trop bas ou ne devient pas trop hétérogène (dans les cas extrêmes, ceci peut équivaloir à une inoculation artificielle).
- une assistance pour interpréter les résultats des essais. Par exemple, une différence significative entre deux traitements peut ne pas avoir la même importance selon le niveau d'infestation.
- l'amélioration de la compréhension des résultats de l'analyse en les exprimant sous une autre forme, ou en permettant leur représentation graphique (par ex. transformation de la mortalité en taux d'efficacité).

- la possibilité d'observations supplémentaires, en particulier le rendement quantitatif et qualitatif, qu'il peut être intéressant de relier aux autres résultats de l'essai.
- enfin, et exceptionnellement, formation d'un terme de comparaison pour les traitements à l'étude si aucun produit de référence n'est disponible. Ceci peut être le cas, par exemple, lorsque le type de produit ou son usage sont nouveaux ou lorsque l'utilisation de tous les produits de référence potentiels a été interdite. Exceptionnellement, ce rôle est alors similaire au rôle du produit de référence, bien que son interprétation soit très différente. Les témoins peuvent alors être comparés aux différents traitements à l'aide de tests statistiques, de la même manière que pour le produit de référence dans les essais habituels.

1.6.2 Types d'implantation des témoins non traités

Il est possible d'envisager quatre types d'implantation des témoins.

Témoins inclus : les témoins sont considérés comme les autres traitements. Les parcelles témoins sont de la même forme et de la même dimension que les autres parcelles, et elles sont randomisées dans l'essai. Le témoin inclus est la façon la plus courante de conduire des essais, toutes les autres versions ne sont qu'exceptionnellement utilisées (principalement dans les tests d'herbicides).

Témoins imbriqués : les parcelles témoins sont disposées systématiquement dans l'essai. Leur dimension et leur forme ne sont pas forcément les mêmes que pour les autres parcelles de l'essai. Les observations réalisées dans ces parcelles sont d'une autre nature et ne doivent pas être prises en compte dans l'analyse statistique. Le but du dispositif est de garantir une distribution plus homogène de l'effet d'une zone adjacente non traitée que ce qu'il est possible d'obtenir avec un témoin inclus. Différents dispositifs peuvent être utilisés; les parcelles peuvent être placées entre les blocs ou entre les parcelles traitées à l'intérieur des blocs (Fig. 9).

Témoins exclus : Les parcelles témoins sont placées à l'extérieur de la zone de l'essai (et n'en sont pas adjacentes), dans une zone où les conditions sont très proches de celles de la zone d'essai. Les répétitions ne sont pas indispensables mais peuvent être utiles si la zone n'est pas homogène. Les observations réalisées dans ces parcelles ne doivent pas être prises en compte dans l'analyse statistique.

Témoins adjacents : Chaque parcelle est divisée en deux sous-parcelles égales, dont une (au hasard) n'est pas traitée. Les observations sont faites de la même manière dans les deux sous-parcelles. Les observations réalisées dans ces parcelles ne doivent pas être prises en compte dans l'analyse statistique sauf si on tient compte du fait que le dispositif est une forme de split-plot. Dans un dispositif split-plot, la variabilité à l'intérieur des parcelles peut être différente de la variabilité entre les parcelles; par conséquent, l'analyse de variance doit

comporter deux strates d'erreur. L'avis d'un spécialiste peut être nécessaire pour interpréter les résultats.

1.7 Sélection de la taille de l'échantillon dans une parcelle

En prenant plusieurs échantillons dans une parcelle, on peut réduire la variabilité de la moyenne estimée de la parcelle à un niveau approprié pour la variable à évaluer. La taille de l'échantillon doit être suffisamment importante pour y parvenir. La taille de l'échantillon requise dépend beaucoup de la nature de l'observation et de la variabilité au sein de la parcelle. Les Normes OEPP sur l'évaluation de ravageurs, adventices et maladies spécifiques donnent des recommandations sur les tailles d'échantillon. Dans la pratique, les tailles d'échantillon comprises entre 10 et 50 éléments suffisent généralement à obtenir une estimation correcte de la valeur moyenne d'une parcelle, en fonction de sa variabilité intrinsèque. A noter que, si les traitements sont appliqués aux parcelles, l'augmentation de la taille de l'échantillon à l'intérieur des parcelles ne donnera qu'un gain d'efficacité strictement limité, les comparaisons entre traitements devant être effectuées à l'échelle inter-parcelles.

L'échantillonnage doit toujours être aléatoire et couvrir de manière adéquate la surface de la parcelle et le matériel expérimental. Pour des raisons pratiques, il peut s'avérer nécessaire de sous-échantillonner. Une liste des méthodes d'échantillonnage et des références bibliographiques peuvent être consultées dans Perry (1994).

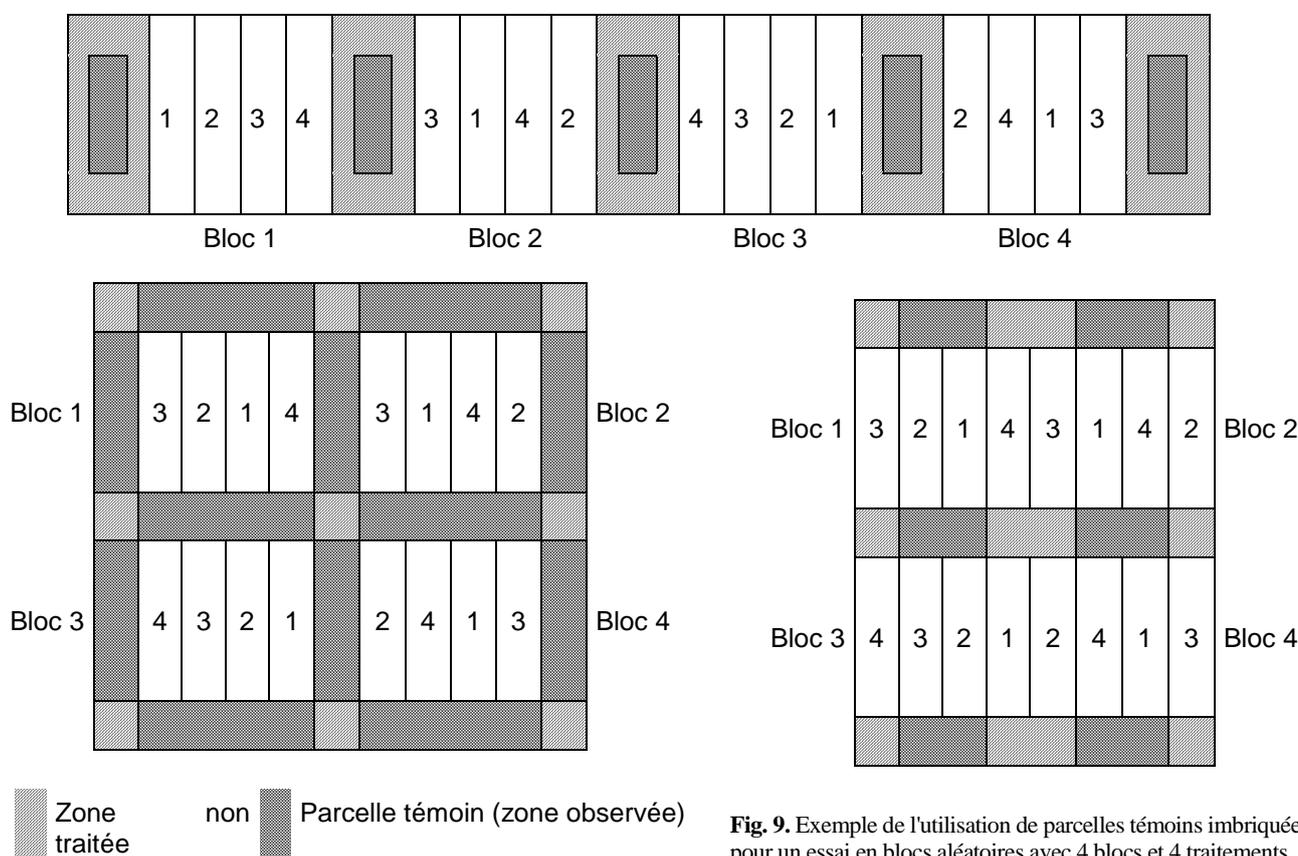


Fig. 9. Exemple de l'utilisation de parcelles témoins imbriquées pour un essai en blocs aléatoires avec 4 blocs et 4 traitements.

2. Principes de l'évaluation des effets des produits phytosanitaires

Lors de la notation des effets des produits phytosanitaires au cours d'un essai d'évaluation biologique, des variables sont notées par des modes d'observation.

2.1 Variables

La nature de la variable est importante car elle influence en général la méthode statistique choisie pour interpréter les résultats. Plusieurs grandes catégories de variables peuvent être distinguées :

Variables binaires (par ex. oui/non, présence/absence) : elles peuvent donner des variables binomiales, qui représentent le nombre de fois qu'un état a été observé pour un nombre donné d'observations (par ex. le nombre de plantes infestées sur 20 choisies par sondage dans une parcelle).

Variables nominales : variables ayant des valeurs d'importance équivalente ne pouvant être classées dans un ordre particulier (non ordinales), par ex. espèces d'organismes nuisibles, couleurs. Les valeurs de ces variables sont généralement indiquées par des mots.

Variables ordinales : variables correspondant à des classes que l'on peut mettre dans un certain ordre mais qui ne sont pas mesurées. Elles sont généralement qualitatives, et les classes ne peuvent alors être ordonnées que les unes par rapport aux autres (par exemple: mauvais, modéré, bon; grades descriptifs des dégâts sur les feuilles). Dans d'autres cas, les valeurs peuvent être exprimées numériquement (elles pourraient être mesurées précisément, mais ne le sont pas pour des raisons pratiques), par ex. catégories de couverture par les adventices, ou catégories d'infestation par les pucerons.

Variables quantitatives : variables pouvant être mesurées, et qui le sont en pratique, par ex. rendement, hauteur des végétaux, nombre de larves, pourcentage de la surface foliaire infectée par une maladie. Elles peuvent être discrètes si elles sont exprimées seulement par des entiers (par ex. résultats des comptages) ou continues (par ex. poids, tailles). Les variables quantitatives peuvent également résulter d'opérations mathématiques. Des différences ou des sommes peuvent être obtenues (par ex. la différence entre une valeur avant et après traitement). Des valeurs relatives peuvent être calculées, sous forme de proportions ou de ratios. Une proportion représente le quotient "quantité partielle/quantité totale" et est comprise entre 0 et 1 (il s'agit d'une variable quantitative continue bornée). En pratique, une proportion est souvent une fréquence relative: "fréquence dans une classe/fréquence totale", c'est à dire une variable binomiale ou multinomiale. Elle est souvent exprimée sous forme de pourcentage (c'est à dire une valeur comprise entre 0 et 100). Les ratios n'ont pas de limite supérieure, par ex. (valeur initiale – valeur finale)/valeur initiale. Ils peuvent

également être exprimés en % (auquel cas il est possible d'obtenir une valeur supérieure à 100). Une valeur exprimée en pourcentage peut être une proportion, ou un ratio, ou même une variable binomiale, il est donc important de distinguer ces cas pour les analyses statistiques.

2.2 Mode d'observation des variables

Les variables doivent être mesurées aussi précisément que possible.

Dans la pratique, prendre les précautions suivantes lors de la notation d'un essai :

- établir l'échelle, la clé ou la méthode de mesure à l'avance. La méthode choisie doit être utilisée dans tous les essais d'une même série.
- faire la notation sans connaissance anticipée du plan de traitement.
- travailler bloc par bloc.
- utiliser la même méthode dans toutes les unités expérimentales, par ex. observer toutes les parcelles dans la même direction pour éviter des différences d'éclairage.

Afin de noter les variables, il est possible d'avoir recours à quatre modes d'observation : mesure, estimation visuelle, classement et notation.

2.2.1 Mesure

Une mesure détermine la valeur de manière objective. Le résultat d'une mesure peut être une variable continue (poids, taille) ou discrète (comptage). Pour les essais au champ, lorsque la mesure ne concerne pas l'ensemble de la parcelle, elle doit porter sur un échantillon dont la taille et le mode d'échantillonnage ont été déterminés au préalable en fonction de la précision requise.

2.2.2 Estimation visuelle

L'estimation visuelle détermine les valeurs de manière subjective, mais avec les mêmes échelles et la même gamme de valeurs que les mesures. Elle concerne généralement des variables continues (par ex. surface couverte par les adventices) mais peut également concerner les comptages dans le cas de grands nombres (par ex. nombre de lésions foliaires). En général, une estimation est plus facile à faire par rapport à une référence (ou un témoin non traité) qu'en valeur absolue. Il faut souligner qu'on obtient le même type de variable qu'avec des mesures. Les résultats sont donc exprimés dans les mêmes unités. Les valeurs obtenues, transformées si nécessaire, peuvent être considérées comme des estimations réelles d'une variable continue et être analysées à l'aide de procédures statistiques normales.

L'observateur doit avoir bénéficié d'un bon apprentissage pour faire les estimations et il doit avoir recours à un étalonnage préalable. Si ces conditions

sont remplies, l'estimation peut être très exacte. La précision peut toutefois dépendre de la valeur de la variable observée. Par exemple, si on estime le pourcentage de surface foliaire atteinte, les valeurs extrêmes seront mieux déterminées que les valeurs intermédiaires. Ceci a été pris en compte lors de la conception des diverses aides à l'estimation et échelles (voir ci-dessous). Une transformation statistique des données estimées peut également être nécessaire.

2.2.3 Classement

Le classement situe chaque individu par rapport aux autres (lui donne un rang). Le résultat d'un classement est une variable ordinale qualitative. Si le nombre d'échantillons à comparer est peu élevé, leur classement pour une variable donnée peut être effectué au champ. S'il n'y a pas d'alternative aux observations classées, on peut utiliser des méthodes non paramétriques comme l'analyse de variance des classements. Cependant, ce type de méthode non paramétrique est généralement moins puissant que les procédures paramétriques. Le classement n'est donc pas une approche idéale et doit être évité en particulier si le nombre de répétitions est peu élevé.

2.2.4 Notation

Effectuer une notation consiste à affecter un objet à une classe définie sans ambiguïté. L'ensemble de ces classes est généralement appelé "échelle", surtout lorsque la variable sous-jacente est ordinale, ce qui est souvent le cas. La notation est également utilisée pour les variables binaires et nominales. Elle est par définition subjective. Elle s'applique à de nombreux types de variables: ordonnées ou non, continues ou discrètes. Elle peut être utilisée pour les variables qualitatives ainsi que pour les variables quantitatives trop onéreuses à mesurer. Son principal intérêt réside dans le fait qu'elle est rapide et non destructrice. Elle peut également être utilisée pour caractériser une parcelle entière par une seule notation. Le nombre de paliers sur une échelle détermine sa sensibilité. Cette

dernière ne doit être ni trop faible (sinon les résultats n'apportent qu'une information limitée), ni trop élevée (l'échelle devient alors peu pratique à utiliser).

Les échelles sont adaptées à des usages spécifiques et en général il n'existe pas d'échelle unique et universelle pour un type de variable donné. Les normes OEPP donnent de nombreux exemples d'échelles dont l'usage est recommandé (voir annexe I) et qui permettent la notation de diverses combinaisons organisme nuisible/culture. Des règles simples ont généralement été suivies lors de l'élaboration de ces échelles, surtout pour définir leurs extrêmes. Le point le plus bas sur une échelle ordinale (pas d'effet) doit recevoir la note 1 (et non 0 qui est réservé aux observations manquantes dans de nombreux systèmes de notation) et le point le plus élevé doit correspondre à la valeur la plus forte de l'effet, avec des valeurs intermédiaires par ordre croissant.

Le Tableau 2 résume les différents modes d'observation et les différents types de variables obtenues.

2.3 Utilisation des échelles pour l'estimation visuelle et la notation

L'estimation visuelle et la notation sont souvent confondues. Ces opérations sont superficiellement semblables, mais leurs résultats diffèrent: l'estimation visuelle produit une série de valeurs estimées pour une variable quantitative discrète ou continue. La notation donne les nombres par classes. Les classes d'une échelle de notation sont souvent représentées par des nombres séquentiels (par ex. 1 à 9), mais cela ne veut pas dire que les intervalles entre les valeurs de l'échelle sont identiques. Si les intervalles varient, il n'est pas conseillé d'entreprendre une analyse statistique sans l'avis d'un spécialiste, ou d'estimer les paramètres statistiques sans réflexion préalable. Les résultats calculés doivent être interprétés avec beaucoup de précaution. Les valeurs des échelles pourraient aussi bien être représentées par les lettres de l'alphabet, ce qui montre bien leur caractère de variable ordinale et les dangers d'une approche trop simpliste.

Tableau 2. Différents modes d'observation et types de variables

	Mesure	Estimation visuelle	Classement	Notation
Binaire				X
Nominale				X
Ordinale			X	X
Discrète	X	X		
Continue limitée	X	X		
Continue non limitée	X	X		

Par contre, les échelles peuvent aussi être utilisées comme aides à l'estimation visuelle ("variable ordinale avec intervalles"). Si les valeurs des échelles correspondent aux valeurs réelles d'une variable quantitative (par ex. sur une échelle visuelle du % de surface foliaire atteinte), l'observateur attribue des valeurs de l'échelle ou interpole des valeurs intermédiaires selon son jugement. Les valeurs obtenues, transformées si nécessaires, peuvent alors être considérées comme des estimations d'une variable continue et être analysées à l'aide des procédures statistiques habituelles. Il est important de souligner que si l'observateur dispose des ressources nécessaires (temps, main-d'œuvre, expérience) pour faire des estimations plus précises, ou même des mesures, les données pourront être analysées avec une meilleure précision et une plus grande puissance. En revanche, rien ne sert d'effectuer une estimation relativement précise (par ex. % de surface foliaire atteinte) pour ensuite lui substituer la valeur correspondante (bien moins précise) d'une échelle. Les avantages de la notation par échelle (vitesse et simplicité) n'existent que si l'observateur note directement dans une classe d'échelle (il doit y être entraîné) sans essayer d'estimer de manière plus précise.

2.4 Qualité d'un mode d'observation

Les modes d'observation peuvent être distingués par un certain nombre de qualités:

"précision"	-absence de biais, dans le contexte statistique;
"fiabilité"	-variabilité (ou variance) faible;
"exactitude"	-combinaison de la précision et de la fiabilité.
"sensibilité"	-réaction du mode d'observation à une modification faible de la valeur de l'unité expérimentale;
"répétitivité"	-un même observateur attribue la même valeur (ou des valeurs très proches) à des unités expérimentales identiques;
"reproductibilité"	-des observateurs différents attribuent la même valeur (ou des valeurs très proches) à la même unité expérimentale.

Ces qualités importantes détermineront le choix des modes d'observation selon les objectifs, en particulier dans les séries d'essais.

3. Analyse statistique des résultats de l'essai

La décision sur la nécessité de conduire une analyse statistique des résultats d'un essai ou d'une série d'essais dépend des résultats démontrés et de l'objectif de l'essai. L'analyse statistique n'est pas essentielle pour tous les essais d'homologation. Mais elle est

particulièrement utile, par exemple, pour comparer les effets des traitements à différentes doses, l'efficacité de différentes formulations d'un même produit, ou l'effet sur le rendement par rapport à un autre traitement.

3.1 Principes

Les sections suivantes visent à esquisser la bonne pratique statistique dans l'analyse de données. Il ne s'agit pas, et il ne peut pas s'agir, d'une prescription valable pour toutes les analyses et toutes les situations.

Les expérimentateurs ne doivent jamais sous-estimer l'utilité de l'avis d'un statisticien. Il est très important que les expérimentateurs comprennent l'avis qu'on leur donne et il est souvent préférable qu'ils effectuent une analyse simple qu'ils peuvent exposer et défendre avec confiance plutôt que d'accepter des conseils les conduisant à une analyse qu'ils ne comprennent que partiellement. La bibliographie à la fin de cette norme peut être utile. Elle mentionne plusieurs textes qui présentent les principes de la bonne pratique statistique plutôt que de donner des recettes statistiques à appliquer aveuglément.

3.2 Analyse statistique d'un essai individuel

3.2.1 Structure de base et séquence de l'analyse

Les Normes OEPP pour l'évaluation biologique des produits phytosanitaires mentionnent que "les résultats seront en principe soumis à une analyse statistique, par des méthodes qui doivent être précisées". La procédure à suivre peut être illustrée par un essai caractéristique dans lequel plusieurs produits à étudier sont appliqués à une dose unique et comparés avec un produit de référence en présence d'un témoin non traité. L'efficacité du produit est évaluée par une variable quantitative mesurée. L'objectif de l'essai est de comparer les produits à étudier au produit de référence, et en particulier d'identifier les plus efficaces. La séquence de l'analyse, pour un essai conduit conformément à la norme OEPP pertinente, est la suivante :

a) L'essai est-il réaliste, c'est-à-dire capable de fournir des données utiles ? Cela est le cas seulement si l'infestation par l'organisme nuisible est suffisamment forte et pas trop variable.

b) Les résultats sont-ils cohérents ? Le produit de référence donne-t-il les résultats attendus par comparaison avec le témoin non traité ?

c) Si ces deux conditions sont satisfaites, on peut comparer les produits à étudier au produit de référence, et éventuellement comparer les produits entre eux. L'analyse a pour objectif premier d'estimer l'amplitude des différences ou ratios entre le produit à étudier et le produit de référence et de fournir une estimation de la variabilité des estimations, sous la forme d'une erreur-type, d'intervalles de confiance ou d'un indice similaire. Des systèmes identiques peuvent être mis au point pour d'autres essais d'évaluation biologique, dont en

particulier le cas spécial de la sélectivité des herbicides, et le cas exceptionnel où il n'existe pas de produit de référence et où les traitements doivent être comparés au témoin non traité (voir section 1.6.1).

Si deux (ou plus) produits de référence sont inclus (voir exemple au 1.2.2), le mode d'analyse doit être défini à l'avance. Une comparaison séparée de chaque produit de référence avec le nouveau produit sans ajustement ou correction est recommandée. Si une comparaison entre le produit à étudier et une combinaison des produits de références est jugée opportune, alors un test d'homogénéité entre les références doit être conduit au préalable.

3.2.2 Choix de la méthode d'analyse

Globalement, le type de variable détermine la méthode d'analyse. Si la variable est quantitative (binaire, binomiale, discrète ou continue), une méthode statistique paramétrique est utilisée, généralement basée sur des Modèles Linéaires Généralisés (MLG), par exemple : analyse de variance, régression linéaire, régression logistique. Si la variable est qualitative, des méthodes non paramétriques conviennent. L'analyse de variance repose sur trois hypothèses : l'additivité des effets, l'homogénéité de la variance et la normalité de l'erreur. L'utilisation de méthodes non paramétriques est recommandée si ces hypothèses ne sont pas satisfaites. Cependant, la non-additivité et la non-normalité peuvent très souvent être améliorées et ne justifient pas forcément l'utilisation de méthodes non paramétriques qui manquent généralement de puissance.

3.2.3 Additivité des effets

Il est important de considérer si les effets sont susceptibles de s'additionner sur l'échelle d'analyse de la variable de réponse. Par exemple, si la variable est la densité de population d'un insecte, les effets du traitement tels qu'un insecticide ou une attaque d'un champignon seront probablement multiplicatifs, et toucheront une proportion de la population. Alternativement, si la variable est la proportion d'adventices tuées par un herbicide, alors les effets seront probablement additifs, non pas sur l'échelle naturelle mais, par ex., sur une échelle probit ou logit. Deux méthodes sont couramment utilisées pour transformer l'échelle naturelle en une échelle plus réaliste: l'utilisation de transformations et l'utilisation de modèles linéaires généralisés.

Les modèles linéaires généralisés sont une forme de régression qui généralise l'analyse de variance. Ils sont supérieurs aux transformations dans la mesure où ils traitent séparément et simultanément le problème de l'additivité des effets et de l'égalité des variances (non normalité). Ils permettent de spécifier directement la distribution de la variable. Par exemple, pour les comptages d'insectes, un modèle de ce type peut spécifier une fonction logarithmique (pour tenir compte

des effets multiplicatifs) et une distribution de Poisson pour les comptages (pour tenir compte directement du problème d'égalité de variances et de non normalité). Alternativement, une variable binomiale peut être analysée en utilisant une fonction logit (pour permettre l'additivité) et en précisant une distribution binomiale (correspondant aux données, qui peut se présenter sous la forme de r plantes malades sur n traitées). Il existe de nombreux parallèles entre l'analyse de la déviation qui résulte de l'utilisation d'un modèle linéaire généralisé, et l'analyse traditionnelle de la variance. En particulier, les concepts de sommes des carrés, de degrés de liberté, de contrastes orthogonaux, de test du χ^2 et de test F , et des moyennes estimées avec des erreurs-types ont tous des parallèles spécifiques dans les modèles linéaires généralisés et ceux-ci peuvent être fournis pour étude.

3.2.4 Homogénéité de la variance

Les transformations peuvent résoudre le problème de l'additivité des effets mais elles ne garantissent généralement pas l'homogénéité des variances. Celle-ci doit être vérifiée indépendamment, même si l'additivité est normalement plus importante. Pour les données discrètes sous forme de comptages, une transformation logarithmique garantit généralement à la fois l'additivité et l'égalité des variances. Pour les données binaires, les données binomiales ou les données sous forme de proportions, une transformation logit, probit ou une transformation complémentaire log-log garantira en général l'additivité, mais pas forcément l'égalité des variances.

3.2.5 Normalité et indépendance de l'erreur

La distribution des erreurs doit être normale. Les tests et représentations graphiques standards sont là pour le vérifier. Dans la pratique, l'analyse de variance est souvent robuste par rapport aux écarts à la moyenne. Dans la mesure du possible, il faut confirmer que les erreurs sont indépendantes des facteurs de traitement.

3.3 Analyse de variance

3.3.1 Tableau des moyennes

Que ce soit dans un modèle linéaire, un modèle linéaire généralisé ou suite à une transformation, une analyse de variance est recommandée. Elle prendra d'abord la forme d'un tableau de la moyenne de chaque traitement, avec une estimation de la variabilité des moyennes, normalement sous forme d'une erreur-type ou d'un intervalle de confiance. Ce tableau souligne l'amplitude des effets et est recommandé pour surmonter le problème bien connu que la pertinence biologique n'est pas égale à la signification statistique; l'amplitude et l'importance des effets peuvent être fortes mais ne pas être significatives en raison de la puissance insuffisante de l'analyse ou du test. L'analyse peut aussi utiliser un modèle linéaire généralisé, dont

l'analyse de variance est un cas spécial, ou toute autre méthode appropriée.

Dans le tableau d'analyse de variance, il faut faire attention à attribuer les unités aux strates correctement, avec une structure des traitements et des blocs convenant au dispositif choisi. En particulier, il faut se prémunir contre le problème bien connu de la pseudo-répétition, qui survient lorsqu'on ne tient pas compte du fait que les traitements n'ont pas été totalement randomisés dans toutes les unités d'échantillonnage mais ont été appliqués à des groupes d'unités.

3.3.2 Tests F et contrastes orthogonaux

En plus de la présentation des tableaux des moyennes et des erreurs-types, des tests statistiques, normalement des tests F , peuvent être appliqués à l'ensemble des données. Un test d'ensemble sur tous les traitements ne constitue pas une preuve de l'efficacité, sauf dans les cas les plus simples, car il est en général influencé par les données des témoins non traités et des traitements à faible efficacité. Au contraire, il peut être utile que la somme des carrés des traitements soit divisée entre les composants d'intérêt biologique en définissant des contrastes orthogonaux (indépendants).

Par exemple, dans le premier exemple où les 8 traitements ont été comparés, il y avait 5 produits à étudier, 2 produits de référence et un témoin non traité. Les huit traitements donnent 7 dl dans les sommes des carrés des traitements. Des contrastes pertinents sont: le témoin non traité par rapport à la moyenne des sept autres traitements (1 dl); le premier produit de référence par rapport au second produit de référence (1 dl); la moyenne des produits de référence par rapport à la moyenne des produits à étudier (1 dl); les différences entre les moyennes des produits à étudier (4 dl). Parmi ces contrastes, les deux premiers ont pour but d'éliminer les variations relativement peu intéressantes du point de vue biologique, et les deux dernières aident à mettre en évidence les objectifs réels de l'essai. Chaque contraste fournit une valeur distincte de F , qui peut être utilisée pour tester formellement des hypothèses intéressantes, qui, dans cet exemple, pourraient être qu'en moyenne, les produits à étudier ne sont pas meilleurs que les produits de référence, et qu'il n'y a pas de différence entre les produits à étudier. L'interprétation de la première de ces hypothèses peut être influencée par la différence entre les produits de référence, surtout si elle est importante. Si des contrastes non orthogonaux doivent être testés, par exemple les cinq contrastes séparés (1 dl chacun) entre la moyenne de chaque produit à étudier et un produit de référence spécifique, alors un test F (ou un test t le cas échéant) doit également être appliqué, en utilisant le carré des écarts résiduel moyen de l'analyse de variance.

Les contrastes et les hypothèses doivent de préférence être précisés par avance, au stade de la conception du dispositif, et utilisés avec modération. Les tests ne doivent pas être effectués simplement parce qu'une

analyse préliminaire post-hoc a mis en évidence des différences qui semblent importantes et pourraient être significatives en cas de test. La régularité est souvent un meilleur indicateur de la présence d'un effet réel que des tests de signification isolés, en particulier si la puissance du test est faible. Par exemple, si un produit à étudier se révèle plus efficace qu'un produit de référence à chacun de onze sites distincts, mais pas de manière significative dans aucun, le bon sens permet de dire que les résultats sont réguliers (et d'ailleurs, un test binomial démontre que la probabilité d'obtenir un résultat aussi régulier s'il y avait en fait aucune différence réelle entre les traitements est inférieure à 0,001).

3.3.3 Procédures de tests multiples

A des fins d'homologation, toutes les comparaisons par paires ne sont pas pertinentes et tous les contrastes orthogonaux ne peuvent pas être pris en considération. De toutes les $(k(k-1))/2$ comparaisons par paires possibles, seules certaines permettent de démontrer l'efficacité d'un produit à étudier. Par exemple, prenons un essai où 7 traitements sont comparés, avec 5 produits à étudier, un témoin non traité et un produit de référence. Selon les règles décrites au paragraphe 3.2.1, plusieurs tests spécifiques doivent être réalisés. Premièrement, la pertinence de l'essai doit être démontrée en évaluant le niveau d'infection, du témoin non traité par rapport à un niveau d'infection prédéfini. Deuxièmement, la différence entre le produit de référence et le témoin non traité doit être évaluée pour démontrer la cohérence de l'essai. Si ces conditions sont remplies, alors la troisième procédure consiste à comparer chaque produit à étudier avec le produit de référence pour tenter de démontrer au moins une égalité des effets par rapport au produit de référence. Pour réaliser ce dernier test, de nombreuses procédures paramétriques et non paramétriques sont acceptées et disponibles dans la littérature scientifique (Hothorn et Bleiholder, 2006).

Dans un essai factoriel (par ex. test à doses multiples), il est généralement inutile de réaliser toutes les comparaisons par paires sur les combinaisons factorielles (Perry, 1986). Il est plus pertinent d'analyser les données en fonction de la structure du protocole. Selon le résultat d'une analyse de variance à deux facteurs, il convient généralement de comparer les moyennes marginales ou les moyennes simples d'un facteur avec des valeurs distinctes de l'autre facteur, et vice-versa.

Les procédures "types" de comparaison multiple selon Tukey (1953) ou le très courant test de Duncan (Duncan, 1955) ou le test de Newman-Keuls (Keuls, 1952) comparent toutes les combinaisons par paires, qui sont par nature bilatérales. Des procédures bien moins conservatrices avec des comparaisons adéquates sont possibles quand elles sont formulées comme des tests unilatéraux. Les tests unilatéraux et les intervalles de confiance sont adaptés sur le plan biologique parce que, par exemple, c'est généralement la réduction de

l'infection qui nous intéresse et non son augmentation. Le test de Duncan, fréquemment employé, et le test de Newman-Keuls ne contrôlent pas la taille globale du test (valeur α), mais seulement sa taille locale (valeur α). Si le test est basé sur une valeur α prédéfinie de 0,05, ceci ne pourra donc être vrai qu'en comparant deux moyennes de traitement ; avec un nombre croissant de moyennes comparées simultanément, la valeur α augmente de façon exponentielle. Lors de l'utilisation de procédures de tests multiples, il est recommandé de sélectionner uniquement les procédures dont on sait qu'elles contrôlent la valeur α locale et globale simultanément.

Comme les essais d'homologation pour démontrer l'efficacité d'un nouveau produit à étudier sont réalisés en phase finale du développement dudit produit, la direction attendue de chaque différence devrait être claire en fonction du contexte. L'usage des tests unilatéraux et des limites de confiance unilatérales est donc recommandé pour garantir un certain niveau de puissance, pour le nombre normal de répétitions généralement utilisées dans les essais au champ. Cependant, ceci n'exclut pas l'usage des autres tests statistiques mentionnés ci-dessus.

3.3.4 Modèles d'effets aléatoires

Cette Norme a considéré les traitements comme des effets fixes. Certains expérimentateurs, dans certains essais (en particulier les essais d'uniformité) peuvent souhaiter considérer les effets du traitement comme un échantillon aléatoire d'une population inconnue plus large; cette méthode est appelée modélisation des effets aléatoires. Les essais peuvent aussi contenir des effets fixes et aléatoires, appelés modèles mixtes. Pour ces modèles, la technique moderne du REML (estimation résiduelle par la vraisemblance maximale) peut être recommandée. Le REML peut également être utilisé lorsqu'on veut comparer plusieurs laboratoires ou sites, pour estimer les composantes de la variance, ou lorsqu'un dispositif ne peut pas être analysé par analyse de variance car il comporte tellement de valeurs manquantes qu'il est déséquilibré. Il existe là aussi des similitudes entre les concepts et les valeurs du REML et ceux de l'analyse de variance; cependant l'avis d'un expert est généralement nécessaire.

3.3.5 Données ordinales

Les méthodes modernes d'analyse des données catégorielles sont décrites dans les livres d'Agresti (1984) et de Brunner et Munzel (2002), mais un avis statistique est nécessaire pour les utiliser à bon escient. Par ailleurs, il est parfois nécessaire de traiter des variables entières comme des variables ordinales, si leur gamme de variation n'est pas suffisamment importante pour qu'on puisse les considérer comme continues et que l'essai reste néanmoins valide.

3.3.6 Données qualitatives et méthodes non paramétriques

Des méthodes non paramétriques peuvent être une procédure statistique utile pour l'analyse de données qui sont réellement qualitatives (par exemple les données nominales) ainsi que pour certaines données classées, ou pour les données qui ne suivent pas une distribution paramétrique bien connue (comme les distributions normale, binomiale, beta, gamma ou de Poisson). Les méthodes non paramétriques sont moins puissantes que les méthodes paramétriques, et doivent donc être utilisées avec une extrême prudence si le nombre d'échantillons est peu élevé. Cependant, la quantité d'informations que ce type d'analyse peut communiquer est pour l'objectif de cette directive suffisamment importante pour obtenir des résultats utiles lors de tests d'efficacité d'un produit. Une très bonne description des tests traditionnels est donnée par Siegel (1956) et Brunner & Munzel (2002), et explique clairement quel test convient pour quelle série de données. Des approches plus modernes utilisent des techniques informatisées comme les tests de randomisation. Les méthodes de randomisation peuvent être très utiles lorsqu'on doit se méfier des approches paramétriques, par ex. si les données sont très éloignées de données normales, ou si elles comportent une forte proportion de zéros (dans le cas où l'essai est néanmoins considéré valide). D'autres méthodes informatisées non paramétriques sont recommandées pour améliorer les estimations, ou pour mieux calculer la variabilité d'une estimation. Elles incluent les méthodes de "bootstrapping" et de "jackknifing", et nécessitent souvent l'avis d'un spécialiste.

3.4 Analyse statistique d'une série d'essais

La cohérence des effets des traitements, par ex. de la comparaison du nouveau produit par rapport au produit de référence, dans différents environnements (régions, sites) est un critère pertinent et important pour l'homologation. Les séries d'essais seront donc préférées aux essais individuels.

3.4.1 Définition

Pour les besoins de cette norme, on définit une série d'essais comme un ensemble de traitements testé sous différentes conditions environnementales sur une ou plusieurs années. Tous les traitements d'une même série doivent être analysés ensemble à l'aide d'un modèle statistique identique.

3.4.2 Planification

Lors de la planification d'une série d'essais, les expérimentateurs doivent définir la question de l'essai et tous les paramètres (c'est-à-dire la liste des traitements principaux, le dispositif de l'essai et répétitions, le nombre de sites, les méthodes de test,

etc.) requis pour appliquer le modèle biométrique qui sera utilisé pour l'analyse de la série d'essais.

3.4.3 Objectifs

Les objectifs de l'analyse sont les suivants :

- Estimer les effets des traitements en plusieurs endroits et sur plusieurs années
- Tester les interactions entre les traitements, les sites et les années. Les différences environnementales ou autres entre les sites et les années peuvent créer des confusions entre ces facteurs
- Si possible tester également si les différences entre les traitements et les références sont significatives.

3.4.4 Structure de base et séquence de l'analyse

Avant d'entamer l'analyse statistique des résultats d'une série d'essais, les données de chaque essai doivent être validées. Cette validation s'applique aux trois points suivants :

- Validation méthodologique : la conduite de tous les essais doit être conforme au protocole original.
- Validation agronomique et biologique : les essais ne doivent pas être perturbés par des facteurs externes ou exceptionnels. Ils doivent être représentatifs de la région et de l'année. Les produits de référence doivent avoir une efficacité normale dans tous les essais. L'infection doit être suffisante (niveau significatif pour les essais d'efficacité, niveau faible pour les essais de sélectivité).
- Validation statistique : les essais doivent être précis et montrer une erreur-type (ou coefficient de variation) normale.

L'analyse d'une série d'essais porte sur l'efficacité ainsi que sur l'interaction traitement-environnement ou traitement-facteur. L'analyse des interactions a pour objectif de démontrer qu'il n'existe pas d'interactions importantes pour tous ou presque tous les environnements et les facteurs et ainsi d'identifier toutes les situations où l'efficacité pourrait être affectée de façon importante (d'autres informations sur les soumissions et l'évaluation par zone se trouvent dans la Norme OEPP *Principes de la production et de l'évaluation des données zonales*. Ceci ne peut pas être démontré de façon appropriée par la simple présence d'un Test F global non significatif. Il convient plutôt d'examiner toutes les composantes d'interaction au moyen des contrastes, afin de démontrer la similarité des effets des traitements sur tous ou au moins la majorité des environnements. Les interactions qualitatives peuvent donc en être exclues, les interactions quantitatives ne devraient pas être tolérées au-delà d'un niveau acceptable dans la pratique. Les sites ne montrant pas d'interaction traitement-environnement peuvent alors être mis en commun pour analyse. Les sites montrant un niveau d'interaction trop important doivent être analysés et discutés séparément.

3.4.5 Choix de la méthode statistique

Comme pour un essai individuel, les méthodes statistiques sont déterminées par le type de variable à analyser. Ces méthodes sont les mêmes ou proches de celles utilisées pour un essai individuel (par ex. analyse de variance, méthodes non paramétriques). L'objectif principal de l'analyse d'une série d'essais est de mesurer et tester les interactions entre les produits à étudier et l'environnement ou le site, c'est-à-dire de montrer que les différences entre les produits sont 'égales' sur chaque site. Les essais peuvent être regroupés avant l'analyse selon des critères pertinents (par ex. type de sol, niveau d'infestation), ou a posteriori à l'aide de méthodes analytiques et des résultats des interactions.

Remerciements

L'OEPP remercie vivement le docteur H. Bleiholder et le professeur L.A. Hothorn pour la révision de cette norme en 2006.

Bibliographie

- Agresti A (1984) *Analysis of Ordinal Categorical Data*. Wiley, New York (US).
- Bauer P, Röhm J, Maurer W & Hothorn LA (1998) Testing strategies in multi-dose experiments including active control. *Statistics in Medicine* **17**, 2133-46.
- BBA (1980) *Richtlinie für Planung, Durchführung und Auswertung von Versuchen mit Pflanzenbehandlungsmitteln: 1. Versuchsplanung; 2. Versuchsdurchführung*. Biologische Bundesanstalt für Land- und Forstwirtschaft, Braunschweig (DE).
- BBA (1982) *Richtlinie für Planung, Durchführung und Auswertung von Versuchen mit Pflanzenbehandlungsmitteln: 3. Auswertung des Einzelversuches; 4. Sachregister, Tabellen*. Biologische Bundesanstalt für Land- und Forstwirtschaft, Braunschweig (DE).
- Brunner E & Munzel U (2002) *Nichtparametrische Datenanalyse*. Springer, Berlin (DE).
- CEB (1981) Rôle et implantation des témoins sans traitement dans les essais de produits phytosanitaires. ANPP-DT4. J Arnoux
- CEB (1983) Principes d'appréciation des effets des produits phytosanitaires dans les essais de plein-champ. ANPP-DT5. Y Ribrioux
- CEB (1986) Utilisation des tests statistiques dans l'interprétation des essais de produits phytosanitaires. ANPP-DT6. J Arnoux - JP Gouet
- CEB (1990) Les réseaux d'essais. ANPP-DT9. JP Gouet
- CEB (1990) Les unités expérimentales. ANPP-DT10. JP Gouet
- Cochran WG & Cox GM (1957) *Experimental Design* (2nd edn). Wiley, New York (US).

- Cox DR (1958) *Planning of Experiments*. Wiley, New York (US).
- Crawley MJ (1993) *GLIM for Ecologists*. Blackwell Scientific, Oxford (GB).
- Crowder MJ & Hand DJ (1990) *Analysis of Repeated Measures*. Chapman & Hall, London (GB).
- Cullis BR & Gleeson AC (1991) Spatial analysis of field experiments – an extension to two dimensions. *Biometrics* **47**, 1449-1460.
- Dagnelie P (1969) *Théorie et Méthodes Statistiques*, Vol. 2. Duculot, Gembloux (BE).
- Denis JB & Vincourt P (1982) Panorama des méthodes statistiques pour l'étude de l'interaction génotype x milieu. *Agronomie* **2**, 219-30.
- Denis JB (1980) Analyse de régression factorielle. *Biométrie-Praximétrie* **19**, 15-34.
- Denis JB, Gouet JP & Tranchefort J (1980) Méthodes d'étude de la structure de l'interaction génotype x milieu et de recherche d'un modèle explicatif à effets fixes: application à l'analyse des résultats d'un réseau d'essais de variété de blé tendre. In: *Biométrie et Génétique*, pp. 98-109. Société Française de Biométrie, Paris (FR).
- Dobson AJ (2002) *An introduction to generalized linear models* (2nd edn). Chapman & Hall, CRC/ Boca Raton (US).
- Duncan DB (1955) Multiple range and multiple F tests. *Biometrics* **11**, 1-42.
- Dyke GV (1988) *Comparative Experiments with Field Crops*. Griffin, London (GB).
- Finney DJ (1971) *Probit Analysis* (3rd edn). Cambridge University Press, Cambridge (GB).
- Finney DJ (1978) *Statistical Method in Biological Assay* (3rd edn). Griffin, London (GB).
- Finney DJ (1980) *Statistics for Biologists*. Chapman & Hall, London (GB).
- Gouet JP & Philippeau G (1992) *Comment Interpréter les Résultats d'une Analyse de Variance?* ITCF, Paris (FR).
- Gouet JP (1974) *Les Comparaisons de Moyennes et de Variances. Application à l'Agronomie*. ITCF, Paris (FR).
- Hollander M & Wolfe DA (1973) *Non-parametric Statistical Methods*. Wiley, London (GB).
- Horn M & Vollandt R (1995) *Multiple Tests und Auswahlverfahren*. Gustav Fischer Verlag, Stuttgart (DE).
- Hothorn LA & Bleiholder H (2006) Statistical aspects of efficacy evaluation of plant protection products in field trials – a comment to the EPPO PP1/152(2) guideline. *Bulletin OEPP/EPPO Bulletin* **36**, 31-45.
- Hughes G & Madden LV (1992) Aggregation and incidence of disease. *Plant Pathology* **41**, 657-660.
- Hurlbert SH (1984) Pseudoreplication and the design of ecological field experiments. *Ecological Monographs* **54**, 187-211.
- Keuls M (1952) The use of studentized range in connection with an analysis of variance. *Euphytica* **1**, 112-122.
- Little TM (1978) If Galileo published in HortScience. *HortScience* **13**, 504-506.
- McCullagh P & Nelder JA (1983) *Generalized Linear Models*. Chapman & Hall, London (GB).
- Mead R & Curnow RN (1983) *Statistical Methods in Agriculture and Experimental Biology*. Chapman & Hall, London (GB).
- Mead R (1988) *The Design of Experiments: Statistical Principles for Practical Applications*. Cambridge University Press, Cambridge (GB).
- Nelder JA (1971) Contribution to the discussion of the paper by O'Neill and Wetherill. *Journal of the Royal Statistical Society series B* **36**, 218-250.
- Parker SR, Whelan MJ & Royle DJ (1995) Reliable measurement of disease severity. *Aspects of Applied Biology* **43**, *Field experiment techniques*, pp. 205-214.
- Patterson HD & Williams ER (1976) A new class of resolvable incomplete block designs. *Biometrika* **63**, pp. 83-92.
- Pearce SC, Clarke GM, Dyke GV & Kempson RE (1988) *Manual of Crop Experimentation*. Griffin, London (GB).
- Perry JN (1986) Multiple-comparison procedures: a dissenting view. *Journal of Economic Entomology* **79**, 1149-1155.
- Perry JN (1989) Review: population variation in entomology: 1935-1950. I. Sampling. *The Entomologist* **108**, 184-198.
- Perry JN (1994) Sampling and applied statistics for pests and diseases. *Aspects of Applied Biology* **37**, 1-14.
- Perry JN (1997) Statistical aspects of field experiments. In *Methods in Ecological and Agricultural Entomology* (ed. Dent DR & Walton MP), pp. 171-201. CAB International, Wallingford (GB).
- Plackett RL (1981) *The Analysis of Categorical Data* (2nd edn). Griffin, London (GB).
- Preece DA (1982) The design and analysis of experiments: what has gone wrong? *Utilitas Mathematica* **21A**, 201-244.
- Rasch D, Herrendörfer G, Bock J, Victor N & Guiard V (1996) *Verfahrensbibliothek, Versuchsplanung und -auswertung*, Band I. R. Oldenbourg Verlag, München (DE).
- Rasch D, Herrendörfer G, Bock J, Victor N & Guiard V (1998) *Verfahrensbibliothek, Versuchsplanung und -auswertung*, Band II. R. Oldenbourg Verlag, München (DE).
- Siegel S (1956) *Non-parametric Statistics for the Behavioral Sciences*. McGraw Hill, New York (US).
- Tukey JW (1953) *The Problem of Multiple Comparisons*. Mimeographed monograph. Princeton University, Princeton NY (US).

Annexe 1 Exemples d'échelles utilisées dans les normes OEPP pour l'évaluation biologique des produits phytosanitaires

1. Nominale

Décoloration des feuilles de pomme de terre (Norme OEPP PP 1/135 *Evaluation de la phytotoxicité*) :

- chlorose ;
- nervures jaunes ;
- taches jaunes ;
- albinisme.

2. Ordinale sans intervalles définis quantitativement

Evaluation des racines du chou pour *Plasmodiophora brassicae* (PLADBR) (Norme OEPP PP 1/39 *Evaluation biologique de fongicides contre Plasmodiophora brassicae*) :

- (1) aucune hernie ;
- (2) légère hernie, pour la plupart sur les racines latérales ;
- (3) hernie moyenne, sur la racine principale et/ou les racines latérales ;
- (4) forte hernie de la racine principale et/ou des racines latérales.

Evaluation des laitues pour *Botryotinia fuckeliana* (BOTRCI) (Norme OEPP PP 1/54 *Evaluation biologique de fongicides contre Botrytis spp. et Sclerotinia spp. sur cultures légumières*) :

- (1) pas d'attaque ;
- (2) attaque légère, infection des pétioles de la base seulement ;
- (3) attaque modérée, lésions de la tige ne l'entourant pas complètement ;
- (4) attaque importante, lésions de la tige l'entourant complètement, ou feuilles supérieures infectées, laitues ne pouvant pas être commercialisées (y compris les plants totalement détruits par *B. fuckeliana* au cours de l'essai).

3. Echelles ordinales avec intervalles définis basés sur des nombres

Attaque par *Venturia inaequalis* (VENTIN) sur pomme (Norme OEPP PP 1/5 *Evaluation biologique des fongicides contre Venturia inaequalis et V. pirina*) :

- (1) pas d'attaque ;
- (2) 1-3 taches d'infection par fruit ;
- (3) > 3 taches d'infection par fruit.

Nombre de lésions sur betteraves dues par ex. à *Scutigerella immaculata* (SCUTIM) (Norme OEPP PP 1/45 *Evaluation biologique des insecticides contre le complexe des ravageurs du sol de la betterave*) :

- (1) pas de lésion ;
- (2) 1-2 lésions ;
- (3) 3-5 lésions ;
- (4) > 5 lésions.

Certaines échelles sont basées en partie sur un nombre et en partie sur une surface, par ex. :

- (1) feuille saine ;
- (2) 1-2 taches par feuille ;
- (3) > 2 taches par feuille ;
- (4) > 1/3 de la surface foliaire infectée.

4. Echelles ordinales avec intervalles définis basés sur des variables continues

Evaluation de l'infection des céréales par *Oculimacula (Tapesia) yallundae* (PSDCHE) and *Oculimacula (Tapesia) acuformis* (PSDCHA) provoquant le piétin-verse (Norme OEPP PP 1/28 *Evaluation biologique des fongicides contre le piétin-verse des céréales*) :

- (1) pas de symptôme ;
- (2) moins de 50 % de la circonférence de la talle attaquée au niveau où l'infection est la plus forte ;
- (3) plus de 50 % de la circonférence de la talle attaquée au niveau où l'infection est la plus forte, mais tissu toujours ferme ;
- (4) 100 % de la circonférence de la talle attaquée avec tissu décomposé (ramollissement).

Généralement, une telle échelle sera (au moins en partie) logarithmique.

Surface foliaire du tabac attaquée par *Peronospora hyoscyami* (PEROTA) (Norme OEPP PP 1/68 *Evaluation biologique des fongicides contre Peronospora hyoscyami*) :

- (1) pas d'infection ;
- (2) 0-5 % de la surface foliaire attaquée ;
- (3) 5-25 % de la surface foliaire attaquée ;
- (4) 25-50 % de la surface foliaire attaquée ;
- (5) 50-100 % de la surface foliaire attaquée.

Bien que ces échelles soient en apparence logarithmiques, il n'y a en fait pratiquement jamais de progression logarithmique constante entre la valeur centrale d'une classe et la valeur centrale suivante. Ainsi, bien qu'en théorie il soit possible d'analyser les notes d'une échelle comme des variables continues correspondant à la simple transformation logarithmique des variables originales, ce cas ne se présente presque jamais car les échelles ne sont pas réellement logarithmiques. D'autre part, l'attribution de la valeur 1 à la classe zéro n'est pas homogène avec le reste de l'échelle.

Un autre point à souligner est que les classes sont définies par les intervalles d'une variable continue. Dans le cas des feuilles de tabac atteintes par *P. hyoscyami* (ci-dessus) l'observateur estime simplement la catégorie d'infection (par ex. 4 ou 5), sans essayer de différencier entre 50 ou 51% (ce qui est manifestement impossible).

Dans de rares cas, des catégories descriptives sont mélangées avec des intervalles définis par exemple l'évaluation des feuilles de pommiers pour *Podosphaera leucotricha* (PODOLE) (Norme OEPP PP 1/69 *Evaluation biologique de fongicides contre Podosphaera leucotricha*) :

- (1) pas d'oïdium ;
- (2) attaque légère (taches éparses) ;
- (3) attaque moyenne à importante (taches couvrant jusqu'à 50 % de la surface foliaire) ;

- (4) attaque très importante (plus de 50 % de la surface foliaire atteinte ; bords des feuilles commençant à s'enrouler et à se dessécher).

5. Echelles ordinales avec classes définies par leur valeur centrale

Ce type d'échelle est à considérer comme une aide à l'estimation. Les échelles visuelles sont les plus fréquentes (exemple : pour *Cercospora beticola*, (CERCBE) PP1/1, *Phytophthora infestans* (PHYTIN) PP1/2, etc.). Ces échelles servent généralement à estimer le % de surface foliaire atteinte, et ont fait l'objet d'un calibrage soigné. Les degrés sont généralement choisis de manière à couvrir la gamme d'attaque prévue (exemple : 1, 5, 10, 25, 50) et de manière à permettre l'interpolation, plutôt que dans une séquence régulière pratiquement logarithmique (qui serait préférable pour faire des notations).