

TABLEAU T.7 DÉFINITION DE LA SÉCHERESSE

Durée	Gravité
3 mois	90% des précipitations médianes 1979-2001 (-10%)
3 mois	75% des précipitations médianes 1979-2001 (-25%)
3 mois	50% des précipitations médianes 1979-2001 (-50%)
6 mois	90% des précipitations médianes 1979-2001 (-10%)
6 mois	75% des précipitations médianes 1979-2001 (-25%)
6 mois	50% des précipitations médianes 1979-2001 (-50%)

aux inondations devraient être fondées sur des modèles hydrologiques plutôt que d'être inférés par des données relatives aux inondations dans la base de données EMDAT.

T.4.5 Le cas des sécheresses

Identification des sécheresses

Les données utilisées dans cette analyse ont été tirées du maillage des précipitations mensuelles pour le monde, pour la période 1979-2001. Ce set de données a été fondé sur un mélange d'observations des stations au sol et d'estimations de précipitations enregistrés par des satellites. La première étape de l'évaluation de l'exposition pour la sécheresse météorologique a été de calculer, pour chaque mois de l'année, la précipitation médiane pour tous les points de grille, entre les latitudes 60S et 70N sur la période de base 1979-2001 (la période de 23 ans pour laquelle les données étaient disponibles). Puis, pour chaque point de grille, le pourcentage des précipitations médianes à long terme a été calculé pour chaque mois au cours de la période entre janvier 1980 et décembre 2000. Pour un mois donné, les points du maillage correspondant aux précipitations médianes à long terme de moins de 0,25 mm/jour ont été exclus de l'analyse. Des

précipitations médianes aussi basses peuvent intervenir soit au cours de la « saison sèche » dans un endroit déterminé, soit dans les régions désertiques. Dans les deux cas, notre définition de la sécheresse n'est pas applicable.

On considère l'existence d'un phénomène de sécheresse météorologique lorsque le pourcentage des précipitations médianes est égal ou inférieur à un seuil donné pour au moins trois mois consécutifs. Les différents seuils considérés ont été de 50%, 75% et 90% des précipitations médianes à long terme, le pourcentage le plus faible indiquant la sécheresse la plus grave, selon cette méthode. Le nombre total d'événements au cours de la période 1980-2000 a donc été déterminé pour chaque point du maillage, et les résultats furent ensuite utilisés pour le reste de l'analyse.

Calcul de l'exposition physique

Utilisant le groupe de données de l'IRI/Université de Columbia, l'exposition physique a été estimée en multipliant la fréquence des aléas par la population vivant dans une zone exposée. Les événements ont été identifiés en utilisant différentes mesures, fondées sur la gravité et la durée telle que décrite dans le Tableau T.7. Pour chacune des six définitions, la fréquence a alors été obtenue en divisant le nombre d'événements par 21 années, produisant ainsi une fréquence annuelle moyenne des événements.

T.5 Analyse statistique : Méthodes et résultats

T.5.1 Définition d'un modèle multiplicateur

L'analyse statistique est fondée sur deux hypothèses de base. Premièrement, que le risque peut être estimé par le nombre de victimes durant événements intervenus par le passé. Deuxièmement, que l'équation du risque suit un modèle multiplicateur, comme dans l'Équation 8.

Utilisant les propriétés logarithmiques, l'équation a été reformulée tel que suit (Équation 9).

Cette équation établit une relation linéaire entre des groupes de données logarithmiques. Cela permet de déterminer des paramètres socio-économiques significatifs V_i (avec des transformations lorsque cela est approprié) et des exposants α_i en utilisant des régressions linéaires.

ÉQUATION 8 ESTIMATION DU NOMBRE DE DÉCÈS

$$\text{EQ 8} \quad T = C \cdot (\text{ExpPh})^{\alpha} \cdot V_1^{\alpha_1} \cdot V_2^{\alpha_2} \dots \cdot V_p^{\alpha_p}$$

Où

T est le nombre de personnes décédées du fait d'un certain type d'aléa

C est la constante multiplicative

ExpPh est l'exposition physique: la population vivant dans les zones exposées multipliée par la fréquence de l'intervention de l'aléa

V_i sont les paramètres socio-économiques

α_i est l'exposant de V_i , qui peut être négatif (pour le ratio)

ÉQUATION 9 PROPRIÉTÉS LOGARITHMIQUES

$$\text{EQ 9} \quad \ln(T) = \ln(C) + \alpha \ln(\text{ExpPh}) + \alpha_1 \ln(V_1) + \alpha_2 \ln(V_2) + \dots + \alpha_p \ln(V_p)$$

T.5.2 Processus détaillé

Données relatives aux victimes

Le nombre de personnes décédées a été dérivé de la base de données EMDAT, et calculé comme la moyenne de personnes décédées par année, au cours de la période 1980-2000.

Filtrage des données

Les modèles statistiques pour chaque type de catastrophe ont été fondés sur des sous-groupes de pays, desquels ont été exclus :

- Les pays ne montrant pas d'exposition physique, ou n'ayant subi aucune perte (valeurs égales à zéro ou nulles)
- Les pays dans lesquels il n'a pas été possible de confirmer les données relatives à l'exposition physique (par exemple le cas du Kazakhstan pour les inondations) ou sur les facteurs socio-économiques.
- Les pays connaissant une faible exposition physique (moins de 2% de la population totale) parce les variables socio-économiques ont été collectées à l'échelle nationale. La population exposée doit être d'une importance suffisante au niveau national en vue de refléter une relation dans le modèle.
- Les pays qui ne disposaient pas de toutes les variables socio-économiques.
- Les valeurs excentriques, lorsque des événements exceptionnels ou d'autres facteurs montrent des niveaux de victimes clairement anormaux, tels que l'Ouragan Mitch au Nicaragua et au Honduras, ou les sécheresses au Soudan et au Mozambique.

ÉQUATION 10 TRANSFORMATION DES VARIABLES ENTRE 0 ET 1

$$\text{EQ 10 } V_i' = \frac{V_i}{(1 - V_i)}$$

Où

V_i' est la variable transformée (entre $-\infty$ et $+\infty$)

V_i est la variable socio-économique (entre 0 et 1)

Transformation des variables socio-économiques

Pour l'analyse statistique, les variables socio-économiques testées ont dû être converties en moyennes sur 21 ans, puis transformées en une valeur logarithmique. Pour certaines des variables, le logarithme a été calculé directement. Pour celles exprimées sous forme de pourcentage, il a été appliqué une transformation afin que toutes les variables se situent entre $-\infty$ et $+\infty$. Pour les autres, aucune transformation logarithmique n'a été nécessaire. Par exemple, la « croissance de la population » se comporte déjà de manière cumulative, et a pu être utilisée directement dans les calculs.

Choix entre les variables

Une condition importante du calcul des régressions est que les variables incluses dans le modèle soient indépendantes, c'est-à-dire que la corrélation entre les deux groupes de

données soit faible. Ce n'est clairement pas le cas avec l'IDH et le taux d'écart de pauvreté dans le PNB par habitant (PNBh), qui ont une corrélation très élevée. Le PNBh a été plus utilisé que l'IDH parce que ce dernier faisait défaut pour de nombreux pays. Afin de maintenir l'échantillon aussi complet que possible, un choix a été opéré entre les variables disponibles, en retenant des variables de set de données aussi indépendants que possible les uns des autres. Ce choix a été opéré en utilisant tant la méthode matrix-plot que correlation-matrix (en utilisant une faible corrélation, et donc une valeur « p » faible, comme critère de sélection).

L'approche par paliers

Pour chaque type d'aléa, de nombreuses régressions logarithmique par paliers et ajustement successifs, ont été réalisés afin d'identifier les variables socio-économiques significatives. La validation du résultat de chaque régression a été menée en utilisant le R^2 , l'analyse de la variance et une analyse résiduelle détaillée.

Une fois le modèle produit, les liens entre le nombre de victimes estimées selon le modèle, et le nombre de victimes observées dans la base EMDAT a été estimé à l'aide de graphiques et de calcul des coefficients de corrélation Pearson.

Si l'on peut comprendre de façon intuitive que l'exposition physique est sans conteste liée au nombre de victimes, et que le PNBh est inversement lié au nombre de victimes (plus le PNB est bas, plus le nombre de victimes est élevé), cela est moins évident pour d'autres variables telles que le pourcentage de terres arables. Cette méthode de régression logarithmique multiple permet l'évaluation des coefficients α_i . Leurs signes ont fourni des informations permettant de montrer si les variables figuraient en position de numérateur ou de dénominateur, et ainsi la relation inverse existant entre les variables et les décès modélisés.

Ce modèle a permis l'identification de paramètres conduisant à un risque plus élevé ou plus bas, mais ne devrait pas être utilisé comme modèle prévisionnel. De faibles différences dans l'échelle logarithmique peuvent aussi induire de grandes différences dans le nombre de décès modélisés.

Les résultats obtenus en suivant cette méthode ont été étonnamment élevés et pertinents, surtout si l'on considère l'indépendance des sources de données et l'imprécision de la résolution des données disponibles au niveau mondial.

T.5.3 Etablir des schémas de risque

Un choix a été opéré entre les différents indicateurs de risque (c'est-à-dire nombre de personnes décédées, nombre de décès par millions d'habitants, nombre de décès en proportion de la population exposée).

ÉQUATION 11 MODÈLE DE RÉGRESSION LOGARITHMIQUE MULTIPLE POUR LES TREMBLEMENTS DE TERRE

EQ 11 $\ln(T) = 1.26\ln(\text{ExpPh}) + 12,27 \cdot U_g - 16,22$

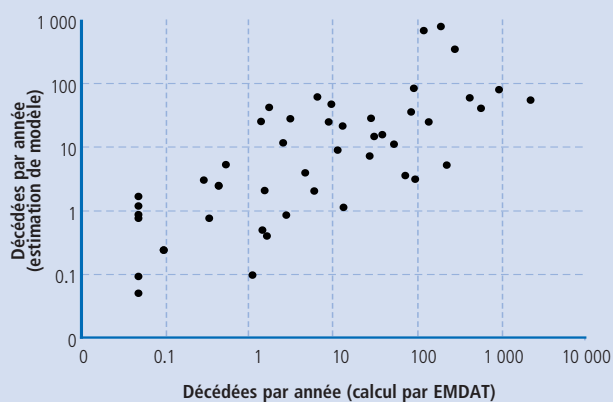
Où
 T est le nombre de personnes tuées du fait de tremblements de terre
 ExpPh est l'exposition physique aux tremblements de terre
 U_g est le taux de croissance urbaine (les pourcentages n'exigent pas de transformation dans la mesure où ce sont déjà des valeurs cumulatives)

TABLEAU T.8 EXPOSANT ET VALEUR DE P POUR LA RÉGRESSION MULTIPLE RELATIVE AUX TREMBLEMENTS DE TERRE

48 pays	B	p-value ^h
Intercepte	-16,22	0,000000
ExpPh	1,26	0,000000
U _g	12,27	0,047686

R = 0,75, R² = 0,56, ajusté R² = 0,54

SCHEMA T.6 NUAGE DE POINTS DE L'OBSERVATION DU NOMBRE DE PERSONNES DÉCÉDÉES DU FAIT DE TREMBLEMENTS DE TERRE (EMDAT) ET MODÈLE PRÉVISIONNEL



Source : EMDAT. Base de données internationale sur les catastrophes OFDA/CRED et PNUÉ/GRID - Genève

T.5.4 Tremblements de terre

Modèle statistique

La régression multiple a été basée sur 48 pays. La ligne de régression la plus adaptée a suivi l'Équation 11.

ÉQUATION 12 MODELO DE REGRESIÓN LOGARÍTMICA MÚLTIPLE PARA LOS CICLONES TROPICALES

EQ 12 $\ln(T) = 0,63\ln(\text{ExpPh}) + 0,66\ln(\overline{\text{Pa}}) - 2,03\ln(\overline{\text{IDH}}) - 15,86$

Où T est le nombre de victimes décédées du fait des cyclones
 ExpPh est l'exposition physique aux cyclones
 $\overline{\text{Pa}}$ est la valeur transformée du pourcentage de terres arables
 $\overline{\text{IDH}}$ est la valeur transformée de l'Indicateur du développement humain

Les variables conservées par la régression comprennent l'exposition physique et le taux de croissance urbaine. La variance qui en ressort est plus faible que pour les inondations ou les cyclones (R²=0.544), toutefois, si l'on considère le délai réduit pris en compte (21 années, comparé à la longue période de retour des tremblements de terre), l'analyse trace une relation raisonnablement correcte. La pertinence de l'exposition physique est semblable à celle des valeurs de p antérieures. La croissance urbaine connaît aussi une forte croissance négative avec le PNB et l'IDH. Ainsi, une corrélation similaire (mais légèrement inférieure) aurait pu être dérivée en utilisant l'IDH et le PNB.

T.5.5 Cyclones tropicaux

Modèle statistique

La régression multiple est fondée sur 32 pays, et la ligne de régression la plus adaptée a suivi l'Équation 12.

Le graphique montre une claire répartition des données comme le montre le Schéma T.7.

Les paramètres soulignés montrent que l'exposition physique, l'IDH et le pourcentage de terres arables sont associés au risque de cyclone.

Le PNB_h est fortement lié à l'IDH, ou négativement lié au pourcentage de croissance urbaine. Dans la plupart des cas, la variable PNB_h doit être remplacée par l'IDH, comme il a été expliqué ci-dessus. Toutefois, ces résultats montrent de manière convaincante que les pays pauvres et les pays qui ont un faible indice de développement humain sont plus vulnérables aux cyclones.

Si l'on considère qu'une grande partie de la variance est expliquée par la régression (R² = 0.863), et qu'un fort degré de confiance peut être attribué aux variables choisies (valeur de p très faible) pour un échantillon de 32 pays, le modèle réalisé est solide.

Dans le modèle, les conséquences de l'Ouragan Mitch pourraient facilement être exposées. En effet, le Honduras et le Nicaragua se situent loin de la ligne de régression (manifestement sous-estimés). Cela s'explique par le fort impact de Mitch en comparaison des autres ouragans. Les valeurs extrêmes données par cet événement ont même conduit à rejeter le Honduras et le Nicaragua du modèle.

h. D'une façon générale, une valeur de p inférieure à 0.05 montre l'importance de l'indicateur choisi. Toutefois, il ne faut pas en user aveuglément.

TABLEAU T.9 SEQ EXPOSANT ET VALEUR DE P POUR LA RÉGRESSION MULTIPLE RELATIVE AUX CYCLONES

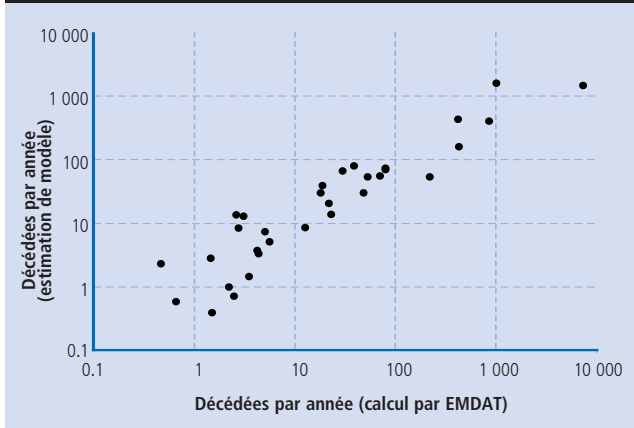
21 pays	B	valeur de p ⁱ
Intercepte	-15,86	0,00000
ln(ExpPh)	0,63	0,00000
ln(Pal)	0,66	0,00013
ln(IDH)	-2,03	0,00095

R = 0,93, R² = 0,86, ajusté R² = 0,85

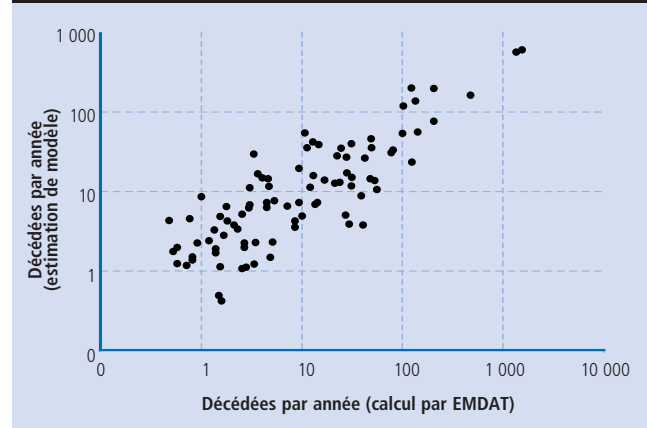
TABLEAU T.10 EXPOSANT ET VALEUR DE P POUR LES INDICATEURS RELATIFS AUX INONDATIONS

90 pays	B	valeur de p ⁱ
Intercepte	-5,22	0,00000
ln(ExpPh)	0,78	0,00000
ln(PIBh)	-0,45	0,00002
ln(Densité)	-0,15	0,00321

R = 0,84, R² = 0,70, ajusté R² = 0,69

SCHÉMA T.7 NUAGE DE POINT DU NOMBRE DE VICTIMES DÉCÉDÉES DU FAIT DE CYCLONES TROPICAUX (CONSTATÉ DANS EMDAT) ET PRÉVISIONS DU MODÈLE

Source : EMDAT. Base de données internationale sur les catastrophes OFDA/CRED et PNUE/GRID - Genève

SCHÉMA T.8 NUAGE DE POINTS SUR LE NOMBRE DE PERSONNES DÉCÉDÉES DU FAIT DES SÉCHERESSES (EMDAT), ET LES PRÉVISIONS DU MODÈLE

Source : EMDAT. Base de données internationale sur les catastrophes OFDA/CRED et PNUE/GRID - Genève

T.5.6 Inondations

Modèle statistique

La régression multiple s'est fondée sur 90 pays. La ligne de régression de la plus adaptée a suivi l'Équation 13.

En raison de contraintes de place, des pays sélectionnés ont été inclus dans le nuage de points ci-dessus. Une liste globale des pays affectés par les inondations est fournie ci-dessous : Albanie, Algérie, Angola, Argentine, Australie, Azerbaïdjan, Bangladesh, Bénin, Bhoutan, Bolivie, Botswana, Brésil, Burkina Faso, Burundi, Cambodge, Cameroun, Canada, Tchad, Chili, Chine, Colombie, Costa Rica, Côte d'Ivoire,

République tchèque, République dominicaine, Equateur, Egypte, El Salvador, Ethiopie, Fidji, France, Gambie, Georgie, Allemagne, Ghana, Grèce, Guatemala, Haïti, Honduras, Inde, Indonésie, Iran (République islamique d'), Israël, Italie, Jamaïque, Japon, Jordanie, Kenya, République populaire démocratie Lao, Malawi, Malaisie, Mali, Mexique, République du Maroc, Mozambique, Népal, Nicaragua, Niger, Nigeria, Pakistan, Panama, Papouasie Nouvelle Guinée, Paraguay, Pérou, Philippines, Pologne, Portugal, République de Corée, République de Moldavie, Roumanie, Fédération de Russie, Rwanda, Arabie saoudite, Sierra Leone, Slovaquie, Afrique du Sud, Espagne, Sri Lanka, Thaïlande, Tunisie, Turquie, Ouganda, Ukraine, Royaume Uni de Grande

ÉQUATION 13 MODÈLE DE RÉGRESSION LOGARITHMIQUE MULTIPLE RELATIF À LA SÉCHERESSE

$$\text{EQ 13} \quad \ln(T) = 0,78\ln(\text{ExpPh}) - 0,45\ln(\text{PNBh}) - 0,15\ln(D) - 5,22$$

Où

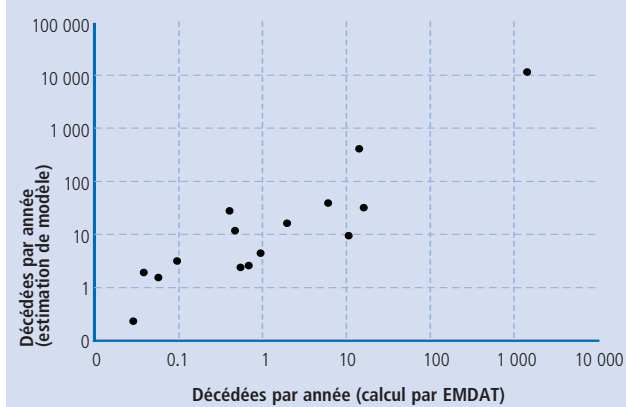
T	est le nombre de victimes décédées du fait des inondations
ExpPh	est l'exposition physique aux inondations
PNBh	est le Produit national brut par habitant (taux d'écart de pauvreté)
D	est la densité de la population locale (c'est-à-dire la population affectée divisée par la zone touchée)

i. D'une façon générale, une valeur de p inférieure à 0.05 montre l'importance de l'indicateur choisi. Toutefois, il ne faut pas en user aveuglément.

TABLEAU T.11 EXPOSANT ET VALEUR DE P POUR LA RÉGRESSION MULTIPLE RELATIVE AUX SÉCHERESSES

Variable indépendante	Coef	SE Coef	T	valeur de p ^j
Constante	14,390	3,411	4,22	0,001
ExpPh3_5	1,2622	0,2268	5,57	0,000
EAU _{TOT} ⁽¹ⁿ⁾	-7,578	1,077	-7,03	0,000

S = 1,345, R-Sq = 0,812, R-Sq (ajusté) = 0,78

SCHÉMA T.9 NUAGE DE POINTS SUR LE NOMBRE DE PERSONNES DÉCÉDÉES DU FAIT DES SÉCHERESSES (EMDAT), ET LES PRÉVISIONS DU MODÈLE


Source : EMDAT. Base de données internationale sur les catastrophes OFDA/CRED et PNUE/GRID - Genève

Bretagne et d'Irlande du Nord, République-Unie de Tanzanie, Etats-Unis d'Amérique, Viet Nam, Yémen et Zimbabwe.

Les variables choisies lors de l'analyse statistique sont l'exposition physique, le PNBh et la densité de la population locale. Le PNBh étant directement lié à l'IDH, ce dernier aurait pu aussi être choisi. Le PNBh a été choisi en raison d'une meilleure corrélation entre le modèle et le nombre de décès constaté, ainsi qu'à cause d'une valeur de p plus faible. L'analyse régressive suppose l'introduction de paramètres n'ayant pas de corrélation, ce qui empêche l'utilisation de toutes ces variables.

La portion de la variance expliquée ($R^2 = 0.70$), associée à une valeur de p significative (entre 10^{-23} et $2 \cdot 10^{-3}$) pour 90 pays

confirme une solide confiance dans le choix des variables (voir Tableau T. 10).

T.5.7 Sécheresse

Modèle statistique

L'analyse régressive a été effectuée en utilisant les différents groupes de données relatifs à l'exposition dérivés des cartes de l'IRI sur les sécheresses. En général, les modèles ont été fondés sur des seuils de trois mois afin d'obtenir de meilleurs résultats. Les sets de données fondés sur un seuil de sécheresse de trois mois, et recevant moins de 50% des précipitations médianes entre 1979-2001 ont finalement été retenus comme données relatives à l'exposition.

La régression multiple a été fondée sur 15 pays. La ligne de régression la plus adaptée a suivi l'Équation 14.

Pays rejetés : le Swaziland et la Somalie (valeur EAU_{TOT} inexistante), la Corée du Nord (la valeur de 100% pour EAU_{TOT} est fortement douteuse), le Soudan et le Mozambique (valeurs excentriques, qui suggèrent d'autres explications aux décès).

Les faibles valeurs de p observées laissent entendre qu'une sélection pertinente des indicateurs a été opérée parmi la liste des groupes de données disponibles. Il faut noter que le fort coefficient pour EAU_{TOT} (-7.578) dénote une forte sensibilité dans la qualité des données. Cela signifie qu'une modification d'1% seulement dans l'accès total à l'eau entraînerait des modifications significatives dans les résultats. Ce serait le cas spécialement pour les faibles valeurs, pour lesquelles de faibles amplitudes ont une plus grande influence en proportion.

Ce modèle ne pourrait pas servir à des fins prévisionnelles. Des incobérences ont été trouvées dans les données, qui appellent des vérifications supplémentaires.

Les variables associées au risque de catastrophe par le biais de l'analyse statistique ont été l'exposition physique et le pourcentage de la population qui dispose d'un accès amélioré aux réserves en eau. Une forte corrélation a été établie ($R^2 = 0.81$), indiquant la solidité de la méthode ainsi que la fiabilité de ces sets de données pour une analyse de cette ampleur.

ÉQUATION 14 MODÈLE DE RÉGRESSION LOGARITHMIQUE MULTIPLE RELATIF À LA SÉCHERESSE

$$\text{EQ 14} \quad \ln(T) = 1,26\ln(\text{ExpPh3_50}) - 7,58\ln(\text{EAU}_{\text{TOT}}) + 14,4$$

Où

T est le nombre de personnes décédées du fait de la sécheresse

ExpPh3_50 est le nombre de personnes exposées aux sécheresses chaque année. Une sécheresse est définie comme une période d'au moins trois mois, ou égale à 50% du niveau moyen de précipitations (IRI, CIESIN/IFPRI/WRI)

EAU_{TOT} est le pourcentage de la population qui dispose d'un accès amélioré aux ressources en eau (OMS/UNICEF)

j. D'une façon générale, une valeur de p inférieure à 0.05 montre l'importance de l'indicateur choisi. Toutefois, il ne faut pas en user aveuglément.

Le tableau T.9 montre la distribution (en échelle logarithmique) des décès modélisés par rapport au décès observé suite à une sécheresse. Une régression claire peut être visualisée, mais il faut noter que si l'on excluait l'Éthiopie, la corrélation serait plus faible ($R^2 = 0.6$). Toutefois, la pente de la droite de régression et l'ordonnée à l'origine ne changent pas de manière significative, ce qui atteste de la solidité du modèle.

Puisque 1.26 est proche d'1, le nombre de personnes décédées croît proportionnellement à l'exposition physique. De même, le nombre de personnes décédées décroît en proportion de la population lorsque croît l'accès amélioré aux réserves en eau. Cette dernière variable devrait être considérée comme un indicateur du niveau de développement du pays, dans la mesure où il entre en corrélation avec les autres variables du développement, tels que le taux de mortalité des moins de cinq ans (Corrélation de Pearson $r = -0.64$) et Indice du développement humain ($r = 0.65$).

Certains pays marqués par une forte exposition physique n'ont connu aucun décès liés à la sécheresse (États-Unis d'Amérique, Viet Nam, Nigeria, Mexique, Bangladesh, Iran, Iraq, Colombie, Thaïlande, Sri Lanka, Jordanie, Équateur). Diverses raisons peuvent l'expliquer. Soit la vulnérabilité est nulle ou extrêmement basse, par exemple aux États-Unis et en Australie, soit le nombre de décès relatifs à l'insécurité alimentaire est placé sous la rubrique conflits dans EMDAT, par ex. en Iraq et en Angola. Pour d'autres pays, une étude pourrait être nécessaire.

T.6 Intégration du risque multiple

Jusqu'à présent la précision et la qualité des données, ainsi que la sensibilité du modèle ne permettent pas d'établir un classement par rang des pays pour le risque de catastrophe, par contre des catégories peuvent être identifiées.

T.6.1 Méthodes

Comment comparer les pays et les catastrophes

Un modèle de risque multiple a été établi en additionnant les décès modélisés pour chaque type d'aléa et pour chaque pays. Afin de réduire le nombre de pays pour lesquels aucune donnée n'existe, ce qui aurait conduit à leur exclusion du modèle, pour les pays qui ne sont pas affectés par l'un des aléas (ou marginalement), la valeur zéro a été substituée en cas d'absence de donnée. Éviter de ce fait d'évincer des pays qui aurait des absences de données relatives à un aléa pour lequel l'exposition est insignifiante. Les pays ont été considérés comme non affectés, si les conditions suivantes étaient remplies : une exposition physique de moins de 2% de la population nationale ET une population affectée de moins de 1000 personnes par année.

Sur 249, seuls 39 pays ont été exclus de l'analyse. Malgré cela, on sait que chacun d'entre eux a été exposé à un certain niveau d'aléa, et que 37 d'entre eux ont des décès enregistrés dans EMDAT. La liste des pays identifie les lieux pour lesquels une amélioration de la collecte des données (surtout socio-économiques) est nécessaire afin de permettre leur intégration dans les futurs travaux. Parmi les raisons pour lesquels des pays individuels ont été exclus on peut indiquer : pays marginalement affectés par un aléa spécifique, pays affectés mais absence de données ; et pays où la répartition du risque ne peut être expliquée par le modèle (par exemple, pour la sécheresse au Soudan, où l'insécurité alimentaire et la famine correspondent davantage à des conflits armés qu'à la sécheresse météorologique telle qu'elle est définie dans le modèle).

Une fois identifiés les pays à inclure dans le modèle, un calcul Booléen a été effectué afin d'allouer l'une des cinq catégories de risque multiple déterminé statistiquement, pour chaque pays. Le Schéma T.10 (cf. page suivante) illustre les différentes mesures prises afin d'incorporer les valeurs dans un indice de risque multiple. Ce processus a donné trois produits :

- Un tableau des valeurs pour les pays comprenant les données pour les aléas pertinents ou pour les pays sans données mais affectés de façon marginale (210 pays).
- Une liste des pays pour lesquels les données manquent (pays qui dans lesquels il y a eu des victimes, mais sans les données appropriées)
- Une liste des pays où le modèle n'a pas pu être appliqué (les indicateurs n'étant pas en mesure de rendre compte de la situation dans ces pays, cas des pays dont la situation ne peut être expliquée par le modèle, ou qui ont rejetés au cours de l'analyse, parce que les indicateurs ne sont pas pertinents au regard de la situation).

Calcul du risque multiple

Le risque multiple a été calculé en utilisant la succession des formules telles que décrites dans l'Équation 15.

Entre chaque addition, il a fallu suivre tout le processus décrit dans le Schéma T.10, afin d'identifier les pays pour lesquels il était nécessaire de remplacer une valeur correspondant à zéro par une valeur calculée à partir d'un modèle d'aléa donné. A défaut, le pays a été placé dans la catégorie « non pertinent », ou « absence de données » (voir ci-dessous).

Afin d'étudier l'adaptation entre le risque dérivé du modèle à risque multiple et les décès enregistrés, les données des deux sources ont été catégorisées en cinq classes de risque pays. Une analyse groupée minimisant la distance intra-classe, et maximisant les inter-classes (en utilisant une analyse par Cluster méthode « *K-mean clustering* ») a été effectuée. Cela veut dire qu'un processus purement statistique (plus objectif) a été utilisé pour identifier la gravité du risque au regard du