UNIVERSITÉ DU BURUNDI
GRENIER DU SAVOIR DU BURUNDI
Dépôt institutionnel officiel
https://repository.ub.edu.bi

Estimation de l’indice de clarté journalier en quelques sites du Burundi à l’aide de durées relatives d’ ensoleillement

Ndacayisaba, Gratien; Sous la direction du : Pr. Mathias BASHAHU

2023
UB-ENS, MASTER EN DIDACTIQUE DES SCIENCES
https://repository.ub.edu.bi/handle/123456789/1012

Document téléchargé depuis le **Grenier du Savoir du Burundi** — https://repository.ub.edu.bi

UNIVERSITE DU BURUNDI- ECOLE NORMALE SUPERIEURE

MASTER CONJOINT EN DIDACTIQUE DES SCIENCES

INSTITUT DE PEDAGOGIE APPLIQUEE

DEPARTEMENT DE PHYSIQUE-TECHOLOGIE

DEPARTEMENT DES SCIENCES NATURELLES

SECTION PHYSIQUE-TECHNOLOGIE

ESTIMATION DE L'INDICE DE CLARTE JOURNALIER EN QUELQUES SITES DU BURUNDI A L'AIDE DE DUREES RELATIVES D'ENSOLEILLEMENT

Par :

Ndacayisaba Gratien

Sous la direction du :

Pr. Mathias BASHAHU

Mémoire présenté et défendu publiquement en vue de l'obtention du grade de master en didactique des sciences physiques

Bujumbura, novembre 2023

MEMBRES DU JURY

Pr Déogratias Nurwaha : Président du jury

Pr Mathias Bashahu : Directeur

Dr Léonard Nduwayo : Secrétaire

Pr Alfred Vyabandi : Membre

DEDICACE

A Dieu tout puissant,

Au parent qui me reste, Bucumi Janvière,

A mes frères et sœurs,

A toute ma famille,

A tous ceux qui ont été à mes côtés dès le début de mes études,

A tous ceux qui me sont chers,

Je dédie ce mémoire.

REMERCIEMENTS

Un projet de mémoire nécessite pas mal d'individus pour être réalisé. C'est pourquoi je prends cette occasion pour lancer un remerciement intime, d'abord à Dieu tout puissant qui m'a fortifié pour réaliser le présent travail.

Mes vifs remerciements s'adressent ensuite à mes parents pour leur soutien tant affectif que matériel. Sans eux, je n'aurais pas pu atteindre ce niveau de mastère.

Ma gratitude s'adresse aussi au promoteur et directeur de ce mémoire, Professeur Bashahu Mathias, pour le suivi régulier qu'il a consacré à ce travail et les conseils qu'il m'a prodigués.

J'adresse aussi une profonde reconnaissance à tous mes éducateurs, depuis l'école primaire jusqu'à ce niveau, plus particulièrement à tous les enseignants de l'Ecole Normale Supérieure (E.N.S) et de l'Institut de Pédagogie Appliquée (I.P.A), spécialement ceux du Département de Physique-Technologie, pour la formation tant morale qu'intellectuelle qu'ils m'ont donnée.

Que tout celui qui a contribué tant matériellement que moralement à la réalisation de ce travail trouve ici le couronnement de son effort.

RESUME

Dans ce travail, la nécessité de la modélisation comme technique (alternative à la mesure) d'acquisition des données sur le R.S, a été motivée. On a ensuite cherché à établir, pour des sites choisis au Burundi, diverses formulations reliant, en termes de moyennes mensuelles des valeurs journalières, l'irradiation solaire globale sur une surface horizontale G (via l'indice de clarté, k_t), à la durée réelle d'ensoleillement S (via chacun des deux types suivants de durée relative d'ensoleillement : s_r et s'_r). A cet effet, des données des moyennes $\bar{S}$ et $\bar{G}$ (de S et G respectivement) ont été recueillies au sein de l'IGEBU. Elles se rapportent à une assez longue période de temps (20 ans ou légèrement plus) et aux six sites ci-après : Bujumbura-Aéroport, Gitega-Zege, Imbo-SEMS, Musasa, Ruvyironza et Rwegura. Puis, pour chacun de ces sites et des 365 jours d'une année normale, les valeurs des grandeurs ci-après ont été déterminées : durée théorique d'ensoleillement journalier (S_o), durée théorique d'ensoleillement journalier (S'_o) modifiée compte tenu de l'horizon naturel du site (au lieu de son horizon astronomique), irradiation solaire journalière hors atmosphère sur une surface horizontale (G_o). Les moyennes mensuelles $\bar{S}_o$,$\overline{S'_o}$ et $\bar{G}_o$ en ont été déduites. Ensuite, pour chacun des six sites, des échantillons de 12 valeurs mensuelles des deux couples de variables ($x = \bar{s}_r = \frac{\bar{S}}{\bar{S}_o}$, $y = \bar{k}_t = \frac{\bar{G}}{\bar{G}_o}$) et ($x' = \overline{s'}_r = \frac{\bar{S}}{\bar{S'}_o}$, $y = \bar{k}_t = \frac{\bar{G}}{\bar{G}_o}$) ont été constitués. Pour chacun de ces six sites également, les cinq couples de (ou les dix) formulations suivantes $y = y(x)$ et $y = y(x')$, de type modèle d'Ångström-Black (ou d'Ångström-Prescott) et modèles associés, ont été établies : (i) relations linéaires simples, (ii) fonctions logarithmiques, (iii) fonctions exponentielles, (iv) fonctions puissances, (v) polynômes du second degré. La performance de chacune de ces formulations (notée sous forme d'une régression linéaire, simple ou double selon le cas) a été évaluée à l'aide des cinq variables statistiques complémentaires ci-après : le coefficient de corrélation (R), celui de détermination (R^2), l'écart moyen (MBE) entre les valeurs calculées et celles mesurées de la variable y, leur écart quadratique moyen (RMSE) et la variable t de Student.

Les résultats de ce travail ont été commentés entre autres en termes : (i) des variations saisonnières (par site) des moyennes $\bar{S}$ et $\bar{G}$; (ii) du classement des six sites par ordre décroissant des moyennes annuelles de $\bar{S}$, $\bar{G}$, $\bar{s}_r$, $\overline{s'}_r$ et $\bar{k}_t$; (iii) du classement des six sites selon le degré de performance de l'ensemble des dix formulations ; (iv) du classement des dix types de formulations sur l'ensemble des six sites.

Comme faits saillants de ces résultats, il ressort essentiellement que : (i) les moyennes annuelles de $\bar{s}_r, \overline{s'}_r \; et \; \bar{k}_t$ caractérisent un ciel moyennement clair ou partiellement nuageux aux sites de cette étude ; (ii) vis-à-vis des valeurs de chacune des cinq variables statistiques (R, $R^2, MBE, RMSE \; et \; t$), toutes les 60 formulations obtenues ont des niveaux de performance très satisfaisants ; (iii) pour les sites et données de cette étude, l'utilisation de la variable indépendante $x' = \overline{s'}_r = \frac{\bar{s}}{\bar{s'}_o}$ présente un léger avantage (p/r à $x = \bar{s}_r = \frac{\bar{s}}{\bar{s}_o}$) dans l'estimation de la variable dépendante $y = \bar{k}_t = \frac{\bar{G}}{\bar{G}_o}$; (iv) les intervalles des valeurs de R, $R^2, MBE, RMSE \; et \; t$ de cette étude, cadrent bien avec les gammes des valeurs obtenues dans d'autres travaux sur ces variables et pour les mêmes types de formulations.

ABSTRACT

Sunshine duration (S) based empirical equations has been employed in this study to estimate the daily global solar radiation on a horizontal surface (G) for six meteorological stations in Burundi. Those equations include the Ångström-Prescott linear model and four amongst its derivatives, i.e. logarithmic, exponential, power and quadratic functions. Monthly mean values of daily global solar radiation and sunshine duration data for a period of 20 to 23 years, from the Geographical Institute of Burundi (IGEBU), have been used. For any of the six stations, ten single or double linear regressions have been developed, from the above-said five functions, to relate in terms of monthly mean values, the daily clearness index ($y = \bar{k}_t = \frac{\bar{G}}{\bar{G}_o}$) to each of the next two kinds of relative sunshine duration (RSD) : $x = \bar{s}_r = \frac{\bar{S}}{\bar{S}_O}$ and $x' = \overline{s'}_r = \frac{\bar{S}}{\bar{S'}_O}$. In those ratios, G_o, S_O and S'_o stand for the extraterrestrial daily solar radiation on a horizontal surface, the day length and the modified day length taking into account the natural site's horizon, respectively. According to the calculated mean values of the clearness index and the RSD, each station experiences a high number of fairly clear (or partially cloudy) days. Estimated values of the dependent variable (y) in each developed linear regression, have been compared to measured values in terms of the coefficients of correlation (R) and of determination (R^2), the mean bias error (MBE), the root mean square error (RMSE) and the t-statistics. Mean values of these statistical indicators have been used to rank, according to decreasing performance level, firstly the ten developed equations per station on account of the overall six stations, secondly the six stations on account of the overall ten equations. Nevertheless, the obtained values of those indicators lay in the next ranges for all developed sixty equations: $R \in [0.8751; 0.9494]$; $R^2 \in [0.7657; 0.9013]$;

$MBE \in [-0.000817; +0.001700]$; $RMSE \in [0.011872; +0.113109]$;

$t \in [0.000000; 0.130589]$, with $t_c = t(n-1=11, \gamma = 99.5\%) = 3.106$. These results indicate that any of the sixty developed linear regressions (and thus equations $\bar{k}_t$ in terms of $\bar{s}_r$ and $\overline{s'}_r$), fits very adequately measured data, and should be used to estimate monthly average daily global solar radiation with sunshine duration for the relevant station. It is also found that using $\overline{s'}_r$ as RSD, is slightly more advantageous than using $\bar{s}_r$ for estimating the monthly average daily clearness index, $\bar{k}_t$. Moreover, values of statistical indicators of this study match adequately with data from other works on the same kinds of empirical equations.

TABLE DES MATIERES

MEMBRES DU JURY ..... **i**
DEDICACE ..... **ii**
REMERCIEMENTS ..... **iii**
RESUME ..... **iv**
ABSTRACT ..... **vi**
TABLE DES MATIERES ..... **vii**
LISTE DES TABLEAUX ..... **x**
LISTE DES FIGURES ..... **xiii**
SIGLES ET ABREVIATIONS ..... **xiv**
AVANT-PROPOS ..... **xvi**
CHAPITRE I. CADRE, DESCRIPTION DES CONCEPTS, PROBLEMATIQUE, FORMULATIONS CHOISIES ET OBJECTIFS **1**
I.1. Cadre général du travail 1
I.2. Description des concepts……..2
I.2.1. Durées d'ensoleillent 2
I.2.1.1. Durée réelle 2
I.2.1.2. Durée théorique 2
I.2.1.3. Durée relative 4
I.2.2. Irradiances et irradiations solaires, puis indices de clarté 5
I.2.2.1. Irradiances et irradiations solaires hors atmosphère 5
I.2.2.2. Rayonnement solaire au sol après traversée d'une atmosphère sans nuages 10
I.2.2.3. Rayonnement solaire au sol après traversée d'une atmosphère réelle 14
I.3. Problématique, cadre spécifique et motivation du sujet du travail 16
I.3.1. Problématique 16
I.3.2. Alternative : cadre spécifique du présent travail 17
I.3.2.1. Modèles d'estimation de G en fonction de S 17
I.3.2.2. Modèles d'estimation de G en fonction d'autres variables ou groupes de variables climatologiques 18
I.3.3. Motivation de ce travail 19
I.4. Description du sujet de recherche choisi et des objectifs 20
I.4.1. Formulations choisies 20

I.4.2. Objectifs 21
I.4.2.1. Objectif principal 21
I.4.2.2. Objectifs spécifiques 21
CHAPITRE II : DEMARCHE METHODOLOGIQUE 22
II.1. Données de base 22
II.1.1. Moyennes mensuelles de S et G 22
II.1.1.1. Sites et périodes d'études choisis 22
II.1.1.2. Valeurs de S et G 23
II.1.2. Moyennes mensuelles de So , $S'o$ et Go , puis valeurs de x, x′et y 25
II.1.2.1. Démarche pour leur détermination 25
II.2. Détermination des coefficients des régressions choisies et des variables statistiques de test des performances de ces régressions 35
II.2.1. Objectifs visés 35
II.2.2. Formulations réductibles à des régressions linéaires simples 35
II.2.2.1. Formulations concernées et leur réduction 35
II.2.2.2. Expressions des coefficients à déterminer 36
II.2.2.3. Autres variables statistiques utilisées pour le test des performances des formulations recherchées 37
II.2.3. Formulations réductibles à des régressions linéaires doubles 39
II.2.3.1. Formulations en question et leur réduction 39
II.2.3.2. Expressions des coefficients à déterminer 39
II.2.3.3. Autres variables statistiques à déterminer 41
CHAPITRE III. RESULTATS OBTENUS ET LEUR INTERPRETATION 42
III.1. Exemples des calculs et résultats sur un site pour chacun des couples des formulations recherchées 42
III.1.1. Relations linéaires simples k_t $vs\ s_r$ et k_t $vs\ s'_r$ 42
III.1.2. Fonctions logarithmiques k_t vs s_t et k_t vs s'_r 44
III.1.3. Fonctions exponentielles $k_t\ vs\ s_r$ et $k_t\ vs\ s'_r$ 46
III.1.4. Fonctions puissances k_t vs s_r et k_t vs s'_r 49
III.1.5. Polynômes du second degré k_t vs s_r et k_t vs s'_r 52
III.2. Résultats des calculs sur tous les sites 57
III.3. Interprétation des résultats 67

III.3.1. Appréciation des formulations obtenues vis-à-vis des valeurs de R et R^2, puis selon les sites 67
III.3.1.1. Classement de ces formulations, site par site, par ordre décroissant des valeurs 67
III.3.1.2. Classement de ces formulations sur l'ensemble des six sites 69
III.3.2. Appréciation des formulations obtenues vis -à-vis des valeurs des autres variables statistiques 71
III.3.2.1. Vis-à-vis des valeurs de MBE 71
III.3.2.2. Vis-à-vis des valeurs de RMSE 74
III.3.2.3. Vis-à-vis des valeurs de t 77
III.3.3. Compléments d'appréciation des formulations obtenues sur base des valeurs des variables statistiques considérées prises globalement 80
III.3.3.1. Tableaux synthétiques 80
III.3.3.2. Premiers compléments d'appréciation 81
III.3.3.3. Autres compléments d'appréciation 83
CHAPITRE IV : SYNTHESE, CONCLUSION ET PERSPECTIVES 85
IV.1. Synthèse des étapes du travail 85
IV.2. Synthèse des principaux faits dans les résultats 87
IV.3. Conclusion et perspectives 88
REFERENCES 90

LISTE DES TABLEAUX

Tableau 1 : Périodes des données de $\boldsymbol{S}$ et $\boldsymbol{G}$, puis coordonnées géographiques des sites de cette étude 23
Tableau 2 : Valeurs de $\boldsymbol{S}$ (en heures) 24
Tableau 3 : Valeurs de $\boldsymbol{G}$ (J/cm^2/jour) 24
Tableau 4 : Valeurs de S_0 27
Tableau 5 : Valeurs de $\boldsymbol{x = s_r}$ 28
Tableau 6 : Valeurs de S'_0 (en heures) 30
Tableau 7 : Valeurs de $\boldsymbol{x' = s'r}$ 31
Tableau 8 : Valeurs de $\boldsymbol{G_0}$ (en J/cm^2/jour) 33
Tableau 9 : Valeurs de y =k_t 34
Tableau 10 : Données nécessaires au calcul des coefficients de la régression linéaire (2.2.1) et des coefficients R et R^2 pour le site de Musasa 42
Tableau 11 : Données nécessaires au calcul des autres variables statistiques : relation (2.2.1), site de Musasa 43
Tableau 12 : Données nécessaires au calcul des coefficients de la régression linéaire (2.2.2) et des coefficients R et R^2 pour le site de Musasa 45
Tableau 13 : Données nécessaires au calcul des autres variables statistiques ; relation (2.2.2), site de Musasa 46
Tableau 14 : Données nécessaires au calcul des coefficients de la régression linéaire (2.2.3) et des coefficients R et R^2 pour le site de Musasa 47
Tableau 15 : Données nécessaires au calcul d'autres variables statistiques ; relation (2.2. 3), site de Musasa 48
Tableau 16 : Données nécessaires au calcul des coefficients de la régression linéaire (2.2.4) et des coefficients R et R^2 pour le site de Musasa 50
Tableau 17 : Données nécessaires au calcul d'autres variables statistiques ; relation (2.2.4), site de Musasa 51
Tableau 18 : Données nécessaires au calcul des coefficients de la régression linéaire double (2.5.1) et des coefficients $\mathbf{R = R_{y.12}}$ **et** R^2 53
Tableau 19 : Données nécessaires au calcul d'autres variables statistiques ; relation 2.2.4, site de Musasa 56

Tableau 20 : Synthèse, pour les six sites, des formulations obtenues et des valeurs correspondantes des variables statistiques du test des performances de ces formulations 58
Tableaux 21. Classement des 10 types de formulations dans l'ordre décroissant des valeurs de R ou R^2 67
Tableau 21.1 : Bujumbura-Aéroport 67
Tableau 21.2 : Gitega-Zege 67
Tableau 21.3 : Imbo-SEMS 68
Tableau 21.4. : Musasa 68
Tableau 21.5 : Ruvyironza 68
Tableau 21.6 : Rwegura 68
Tableau 21.7 : Classement des six sites par ordre décroissant de R_{moy} sur les dix formulations 69
Tableau 21.8 : Fréquences d'apparition des 10 formulations dans diverses positions de leur classement, sur l'ensemble des six sites, par ordre décroissant des valeurs de R ou (de R^2) 70
Tableaux 22 : Classement des 10 types de formulations par ordre croissant des valeurs absolues de MBE 71
Tableau 22.1 : Bujumbura-Aéroport 71
Tableau 22.2 : Gitega-Zege 71
Tableau 22.3 : Imbo-SEMS 71
Tableau 22.4 : Musasa 72
Tableau 22.5 : Ruvyironza 72
Tableau 22.6 : Rwegura 72
Tableau 22.7 : Classement des six sites par ordre décroissant des performances, sur l'ensemble des 10 formulations, selon les valeurs de $\boldsymbol{MBE}$ 73
Tableau 22.8 : Fréquences d'apparition des 10 formulations dans diverses positions de leur classement, sur l'ensemble des six sites, par ordre croissant des valeurs absolues de MBE 73
Tableaux 23 : Classement des 10 types de formulations par ordre croissant des valeurs de RMSE 74
Tableau 23.1 : Bujumbura-Aéroport 74
Tableau 23.2 : Gitega-Zege 74

Tableau 23.3 : Imbo-SEMS 74
Tableau 23.4 : Musasa 75
Tableau 23.5 : Ruvyironza 75
Tableau 23.6 : Rwegura 75
Tableau 23.7 : Classement des six sites par ordre décroissant des performances de l'ensemble des 10 types de formulations, selon les valeurs croissantes de $\boldsymbol{RMSE}$ 76
Tableau 23. 8 : Fréquences d'apparition des 10 formulations dans diverses positions de leur classement, sur l'ensemble des six sites, par ordre croissant des valeurs de RMSE 76
Tableaux 24 : Classement des 10 types de formulations par ordre croissant des valeurs de t pour les six sites 77
Tableau 24.1 : Bujumbura-Aéroport 77
Tableau 24.2 : Gitega-Zege 77
Tableau 24.3 : Imbo-SEMS 77
Tableau 24.4 : Musasa 78
Tableau 24.6 : Rwegura 78
Tableau 24.7 : Classement des six sites par ordre décroissant des performances, sur l'ensemble des dix formulations, selon les valeurs moyennes $\boldsymbol{t}$ de la variable t 79
Tableau 24.8 : Fréquences d'apparition des 10 formulations dans diverses positions de leur classement, sur l'ensemble des six sites, par ordre croissant des valeurs de la variable t 79
Tableau 25 : Classement, sur l'ensemble des 10 types de formulations, des 6 sites par ordre des moyennes des 4 variables statistiques considérées 80
Tableau 26 : Classement, sur l'ensemble des 6 sites, des 10 types de formulations par ordre décroissant des performances, selon les valeurs des 4 types de variables statistiques statistiques 81
Tableau 27 : Valeurs de R^2, MBE, RMSE, et t obtenues dans quelques autres travaux sur les mêmes types de modèles $\boldsymbol{k_t}$ vs $\boldsymbol{s_r}$ qu'ici 84

LISTE DES FIGURES

Figure 1 : Schéma représentant les coordonnées horizontales et équatoriales du soleil, plus la latitude du site 3
Figure 2 : Spectre de l'irradiance solaire normale de référence hors atmosphère : données du WRC et de la NASA. 7
Figure 3 : Comparaison du spectre solaire normale de référence avec celui du rayonnement d'un corps noir à la température de 5777 K 7
Figure 4 : Représentation des six sites de cette étude sur une carte du Burundi 23
Figure 5. : Variations de S_0 avec J, exemple du site de Musasa. 26
Figure 6 : Variations de S'_0 avec J, exemple du site de Musasa 29
Figure 7 : Variations de G_0 avec J, exemple du site de Musasa 32

SIGLES ET ABREVIATIONS

c-à-d	: c'est-à-dire
CIMO	: Commission pour les Instruments et Méthodes d'Observation
C.N	: cops noir
ddl	: degré(s) de liberté
é.m(s)	: électromagnétique (s)
éq(s)	: équation (s)
ERs	: énergie(s) renouvelable(s)
etc	: etcetera (ainsi de suite)
ex.	: exemple (s)
fig (s)	: figure (s)
IGEBU	: Institut Géographique du Burundi
l.d.o	: longueur(s) d'onde
LIDAR	: light detection and ranging
NASA	: National Aeronautics and Space Administration
N.B	: nota bene (notez bien)
ndlm	: niveau de la mer
NIP	: normal incidence pyrheliometer
n^o, n^{os}	: numéro, numéros
o.d.g	: ordre de grandeur
O.M.M	: Organisation Mondiale de la Météorologie
p/r	: par rapport
rad.é.m	: radiation(s) électromagnétique(s)
R-D	: recherche-développement
r.é.m	: rayonnement électromagnétique
réf (s)	: référence(s)
RMSE	: root mean square error
R.S	: rayonnement solaire
RSD	: relative sunshine duration
T.M.Y	: typical meteorological year
u.h.s	: unité hors système
U.S.A	: United States of America

u.S.I : unité(s) du Système International

vs : versus (en fonction de)

WRC : World Radiation Center

1^{er}, 1^{ers} : premier, premiers

$\sim$: Proportionnel à

AVANT-PROPOS

Ce travail de recherche s'intitule : « Estimation de l'indice de clarté journalier en quelques sites du Burundi à l'aide de durées relatives d'ensoleillement » s'inscrit dans un travail de fin d'études de Master en Didactiques des Sciences Physiques.

Comme motivation du choix du sujet, les données d'irradiations solaires acquises à l'aide de techniques de modélisation, constituent donc un complément indispensable à celles obtenues au moyen de techniques de mesure. Premièrement un intérêt certain de la disponibilité de telles données réside d'abord dans leur utilisation comme « entrées » de tout projet d'installation d'un système de conversion et/ou stockage de l'énergie du R.S.

Deuxièmement, elles sont également indispensables dans l'établissement des cartes (saisonnières ou annuelles) des irradiations solaires et de la durée d'ensoleillement d'un pays ou d'une région donné(e) ainsi que dans la définition d'une année météorologique typique pour chacune des grandeurs précédentes (entre autres).

Le but de ce travail de recherche est d'utiliser les variables durées relatives d'ensoleillement journalières s_r et s'_r pour estimer la grandeur indice de clarté journalier k_t (et donc irradiation solaire globale journalière G) en différents sites du Burundi.

CHAPITRE I. CADRE, DESCRIPTION DES CONCEPTS, PROBLEMATIQUE, FORMULATIONS CHOISIES ET OBJECTIFS

I.1. Cadre général du travail

Pour leurs besoins primaires en énergie, qui sont la cuisson et le chauffage des habitations, 98 % de la population au Burundi utilisent du bois et ses débris dans le milieu rural, puis du charbon de bois en milieu urbain [1]. Cette activité a un impact néfaste, non seulement sur la santé publique (via par exemple les maladies respiratoires qu'elle cause), mais aussi sur l'environnement (via la déforestation). Selon un rapport fourni en 2013 par le Ministère de l'Energie et Mines, dans la période 1993-2000, seulement 2,4 % de la population au Burundi avaient accès à l'utilisation de l'énergie électrique [2]. L'énergie électrique consommée annuellement dans ce pays (400 GWh selon les estimations de 2014) [3] provient à 98,2 % d'installations hydroélectriques et à 1,8 % de combustibles fossiles (produits pétroliers uniquement). La consommation électrique annuelle moyenne par habitant est de 30 kWh, ce qui est environ 5 fois plus faible que la même moyenne en Afrique [3]. La production électrique annuelle dans le pays (300 GWh selon les estimations de 2014) étant inférieure à la consommation, ce dernier doit importer la quantité manquante (100 GWh). Entretemps, des coupures de courant électrique sont fréquemment enregistrées dans plusieurs localités en raison de ce déficit de production. Pour résoudre ponctuellement les problèmes liés à ces coupures, divers établissements, institutions et services doivent utiliser des générateurs électriques (ex. : des moteurs Diesel) consommant des produits pétroliers importés. Cependant, les prix de ces produits ne cessent de croître au fil des ans. La pollution résultant des émissions des gaz (issus de la combustion de ces produits) impacte quant à elle négativement l'environnement.

Comme d'autres pays, le Burundi cherche à résoudre les problèmes précédents et à faire face à la forte demande énergétique liée à ses efforts de développement socio-économique. L'une des voies pour atteindre ces objectifs, et qui est clairement tracée dans divers plans gouvernementaux [1] [4] [5], est l'utilisation des énergies renouvelables (ERs) en complément (ou en substitution partielle) des sources traditionnelles d'énergie. En effet, le rapport cité plus haut [2] stipule qu'en dehors de l'énergie hydroélectrique qui représente 95 % des ERs, l'utilisation d'autres ERs reste très peu significative dans ce pays. Cette utilisation ne peut être efficace que quand elle résulte d'une recherche-développement (R-D) conséquente dans le domaine.

Le présent travail s'inscrit dans le cadre d'une R-D en énergie solaire (l'une des ERs) au Burundi, plus exactement dans la partie "gisement solaire" (disponibilité de la ressource).

I.2. Description des concepts

I.2.1. Durées d'ensoleillent

I.2.1.1. Durée réelle

En un site donné (de latitude Φ et de longitude L) et une date ou un jour numéro (n°) J de l'année (J varie de 1 à 365 ou 366), la durée réelle d'ensoleillement est aussi appelée l'insolation effective, notée S. C'est la période de temps, exprimée en heures, durant laquelle on peut voir le disque solaire et observer les ombres portées au sol, ou durant laquelle l'irradiance solaire globale (sur une surface horizontale) est supérieure à un certain seuil [6]. Cette grandeur S est mesurée à l'aide d'appareils appelés héliographes. Ceux-ci sont de divers types et diffèrent entre autres par leurs détecteurs. Parmi eux, l'héliographe de référence internationale est celui de Campbell-Stokes, dont le seuil de sensibilité est fixé à 120 Wm^{-2} (norme OMM, 1962) [7]. Plusieurs stations météorologiques dans divers pays sont dotées de ce type d'appareil et opèrent régulièrement des mesures de S.

I.2.1.2. Durée théorique

En un site donné et un jour considéré de l'année, p/r au plan horizontal (ou à l'horizon astronomique) de ce site, la durée théorique d'ensoleillement, également appelée longueur de la journée, est la période de temps, exprimée en heures, entre les instants du lever et du coucher solaires (en temps solaire vrai), notés t_{sr} et t_{ss} respectivement. Ainsi, cette durée notée S_O a pour expression :

$$S_O = t_{ss} - t_{sr} \qquad (1.1.1)$$

Dans l'équation (éq.) (1.1.1), les indices sr et ss se réfèrent respectivement aux termes anglais « sunrise » et « sunset ». En vue de déduire l'expression analytique de S_O, on part de la relation classique ci-après entre les coordonnées horizontales et équatoriales du soleil, ainsi que les coordonnées géographiques du site telles que représentées sur la figure (fig.) 1.1 [8] :

$$\sin h = \cos \theta = \sin \delta \sin \Phi + \cos \delta \cos \Phi \cos \omega_s \qquad (1.1.2)$$

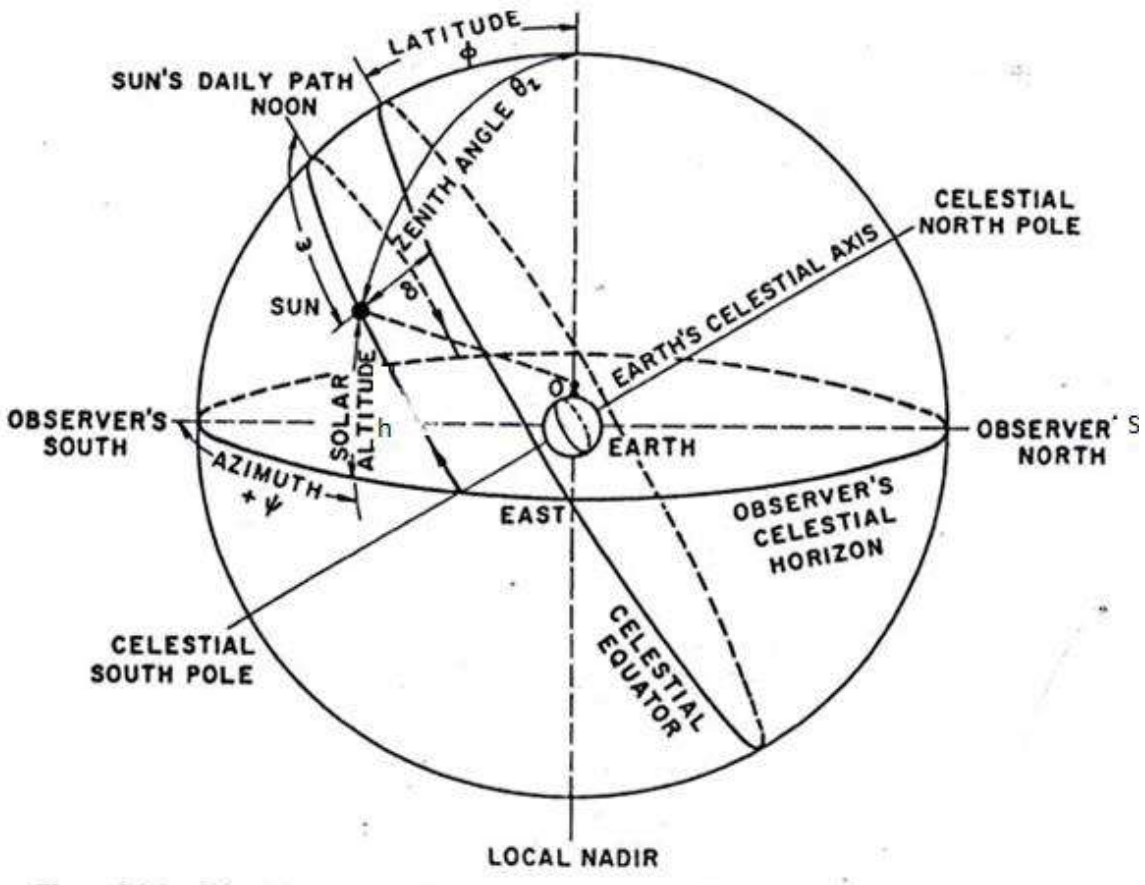


Figure 1 : Schéma représentant les coordonnées horizontales et équatoriales du soleil, plus la latitude du site

Dans l'éq.(1.1.2) et sur la fig. 1.1, h est la hauteur du soleil au-dessus de l'horizon (astronomique du site), θ est l'angle zénithal du soleil à l'instant considéré (c'est aussi le complément de h à 90°), δ est la déclinaison solaire (au jour considéré), ω_s est l'angle horaire du soleil (à l'instant précédent), puis Φ est la latitude du site. Au lever et au coucher solaires (p/r à l'horizon astronomique) ce jour considéré (pour le site), h vaut 0° et 180° respectivement, puis θ vaut quant à lui 90° et -90°. Ainsi, à partir de l'eq.(1.1.2), on obtient :

$$cos\omega_{sr/ss} = -(tg\ \delta\ )(tg\ \Phi) \qquad (1.1.\ 3)$$

L'une des deux conventions dont la littérature fait état pour le signe de l'angle horaire (ω_s) du soleil (et c'est celle-là qui est adoptée dans ce travail) est que ω_s est négatif avant midi (AM), nul à midi et positif après-midi (PM) [9]. Ainsi, on a :

$$\omega_{sr/ss} = \pm\, cos^{-1}(-tg\ \delta\ tg\ \Phi) \qquad (1.1.\ 4)\ ;\ \omega_{sr}=-\omega_{ss} \qquad (1.1.\ 4')$$

Selon cette même convention, ω_s (en degrés) est lié comme suit au temps solaire vrai (en heures) [10] :

$$\omega_s = 15(t - 12) \qquad (1.1.\ 5)$$

On obtient ainsi :

$$t_{ss} = 12 + \frac{\omega_{sr}}{15} \qquad (1.1.\ 6)\ ;\ t_{sr} = 12 + \frac{\omega_{sr}}{15} = 12 - \frac{\omega_{ss}}{15} \qquad (1.1.\ 7)$$

$$S_o = t_{ss} - t_{sr} = \frac{2\omega_{sr}}{15} = \frac{2}{15} cos^{-1}(-tg\ \delta\ tg\ \Phi) \qquad (1.1.\ 8)$$

Dans certaines références, on tient compte de l'horizon naturel d'un site (au lieu de son horizon astronomique) en considérant que les instants effectifs du lever et du coucher solaires en ce site correspondent à h=5° et 175° (et donc à θ=85° et -85°) respectivement. L'expression modifiée de la durée théorique d'ensoleillement est alors la suivante [11] [12] :

$$S'_0 = \frac{2}{15}\ cos^{-1}\left[\frac{\cos 85° - \sin\delta \sin\Phi}{\cos\delta \cos\Phi}\right] \qquad (1.1.\ 8')$$

Dans les éqs.(1.1. 2) à (1.1. 8'), δ dépend du numéro (J) du jour de l'année. Ses valeurs sont disponibles dans la littérature (ex .dans [8] [9]) sous forme de tables ou d'expressions analytiques.

I.2.1.3. Durée relative

En un site considéré et un jour donné de l'année, la durée relative d'ensoleillement est le rapport entre la durée réelle et la durée théorique d'ensoleillement. Selon qu'on considère S_0 ou S'_0 pour la longueur du jour, on a donc les expressions ci-après pour cette durée relative :

$$s_r = \frac{S}{S_O} \qquad (1.1.\ 9); \qquad s'_r = \frac{S}{S'_0} \qquad (1.1.\ 9')$$

Chacune de ces deux grandeurs est donc adimensionnelle et varie de 0 à 1. Ses valeurs caractérisent le type de conditions climatiques. Ainsi, un ciel est considéré comme serein lorsque s_r ou $s'_r \in$ [0,80 ; 1,0], moyen quand $s_r\ ou\ s'_r \in$[0,20 ; 0,80] et couvert lorsque $s_r\ ou\ s'_r$ est inférieure à 0,20 [13].

I.2.2. Irradiances et irradiations solaires, puis indices de clarté

I.2.2.1. Irradiances et irradiations solaires hors atmosphère

a) Origine du rayonnement solaire

Le rayonnement solaire (R.S) trouve son origine dans les transformations thermonucléaires des composants chimiques qui constituent le soleil. En effet, vu l'ordre de grandeur (o.d.g) des températures au sein de cette étoile (1,5 10^7 ℃ au centre et 6 10^3 ℃ à la surface), il se passe des réactions de fusion des éléments légers de cet astre suivant divers cycles (ou diverses chaînes), en particulier suivant le cycle de Bethe [7] [14]. Le défaut de masse dans l'éq. du bilan des réactions d'un de ces cycles, indique qu'une partie non négligeable (environ 0,67% de masse) de la matière concernée se transforme entre autres en électrons et radiations électromagnétiques (rad.é.m(s)). Ces dernières couvrent tous les domaines du spectre électromagnétique (é.m), c'est-à-dire (c-à-d) toutes les valeurs de l'une des variables ci-après: fréquence (υ), pulsation (ω), longueur d'onde (l.d.o) spatiale (λ), énergie (ε), nombre d'onde ($\bar{\nu}$) et module (k) du vecteur d'onde; qui sont liées par les relations suivantes (de l'hypothèse des quanta de Planck et celle des photons d'Einstein) :

$$\varepsilon = h\upsilon = \hbar\omega = \frac{hc}{\lambda} = hc\bar{\nu} = \hbar ck \qquad (1.2.1)$$

où h est la constante de Planck et ħ la constante réduite de Planck ($\hbar = \frac{h}{2\pi}$). Lors de sa propagation dans une direction quelconque de l'espace et au cours du temps (avec une vitesse de module « c »), chacune de ces rad.é.m(s) emporte une énergie ε et une impulsion $= \hbar k = \frac{\varepsilon}{c}$. La propagation de toutes ces rad.é.m(s) constitue donc le rayonnement électromagnétique (r.é.m) du soleil ou le R.S. Même si elle est située à une grande distance du soleil, la terre reçoit ce R.S sous forme de chaleur et de lumière, en marge d'autres phénomènes liés à l'activité solaire (ex. : les orages magnétiques, les aurores boréales, etc.)[14]

b) Spectre de l'irradiance solaire normale de référence hors atmosphère et sa valeur intégrée

Une irradiance solaire est la quantité d'énergie du R.S qui atteint une unité d'aire d'une surface (située à une distance « r » donnée du soleil), par unité de temps. Son unité du Système International (u.S.I) est le Wm^{-2}. De son côté, une irradiation solaire est la quantité d'énergie du R.S qui atteint une unité d'aire d'une surface (située à une distance « r » donnée

du soleil), pendant une durée supérieure à la seconde. Cette durée peut être par ex. : horaire, journalière, etc. Son u.S.I est le Jm^{-2}.

La distribution spectrale de la puissance (ou de l'énergie par unité de temps) du R.S reçue par une unité d'aire d'une surface perpendiculaire (ou normale) au faisceau du R.S direct et qui est placée à la distance moyenne terre-soleil (centre à centre), dans l'hypothèse d'absence de l'atmosphère terrestre, est appelée spectre de l'irradiance solaire normale hors atmosphère.

En termes de la variable l.d.o spatiale (λ), cette distribution spectrale est notée $\dot{I_{SC}}(\lambda)$, et s'exprime souvent en l'unité hors système ($u.h.s$) suivante: $Wm^{-2}\mu m^{-1}$. La distance moyenne précédente, qui est une unité astronomique (u. a) des distances, a pour valeur (en u.S.I.) :

$$r_o = 1u.a = 1{,}496\ 10^{11}\,\text{m} \qquad (1.2.2)$$

La distribution précédente a fait l'objet, pendant plusieurs dizaines d'années, d'un grand nombre d'études théoriques et expérimentales, avec des mesures (utilisant des instruments variés) effectuées au sol, en haute altitude, à partir d'engins aéroportés et même de satellites [6]. En termes de synthèse des études précédentes, la fig.1.2.a) montre ladite distribution telle que proposée par le « World Radiation Center »(WRC) de Davos (en Suisse) et le « National Aeronautics and Space Administration » (NASA, aux USA). Pour ce spectre, les données de ces deux institutions s'accordent bien et, comme le montre la fig.1.2.b), ces données coïncident bien avec celles du spectre du rayonnement d'un corps noir (C.N) en équilibre à une température de 5777 K [8].

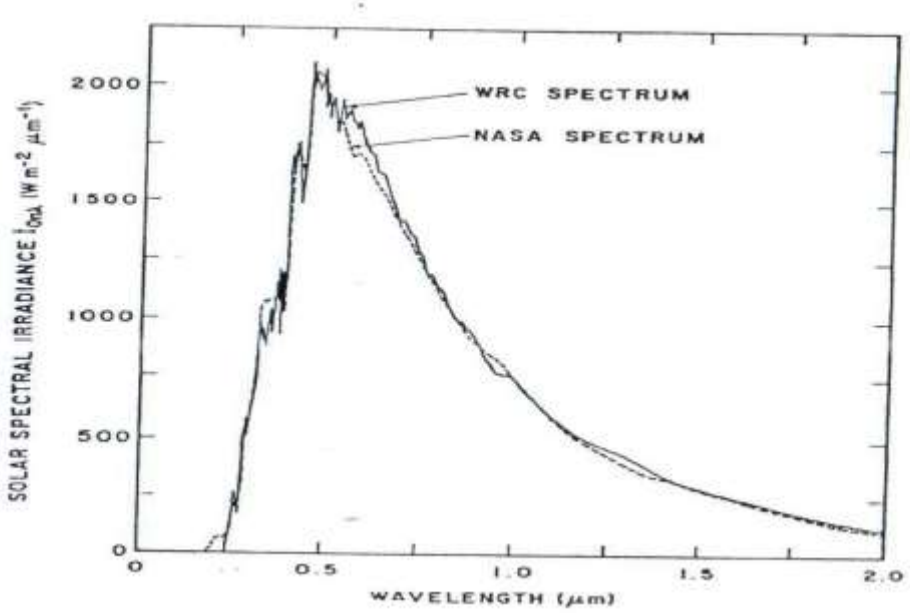


Figure 2 : Spectre de l'irradiance solaire normale de référence hors atmosphère : données du WRC et de la NASA.

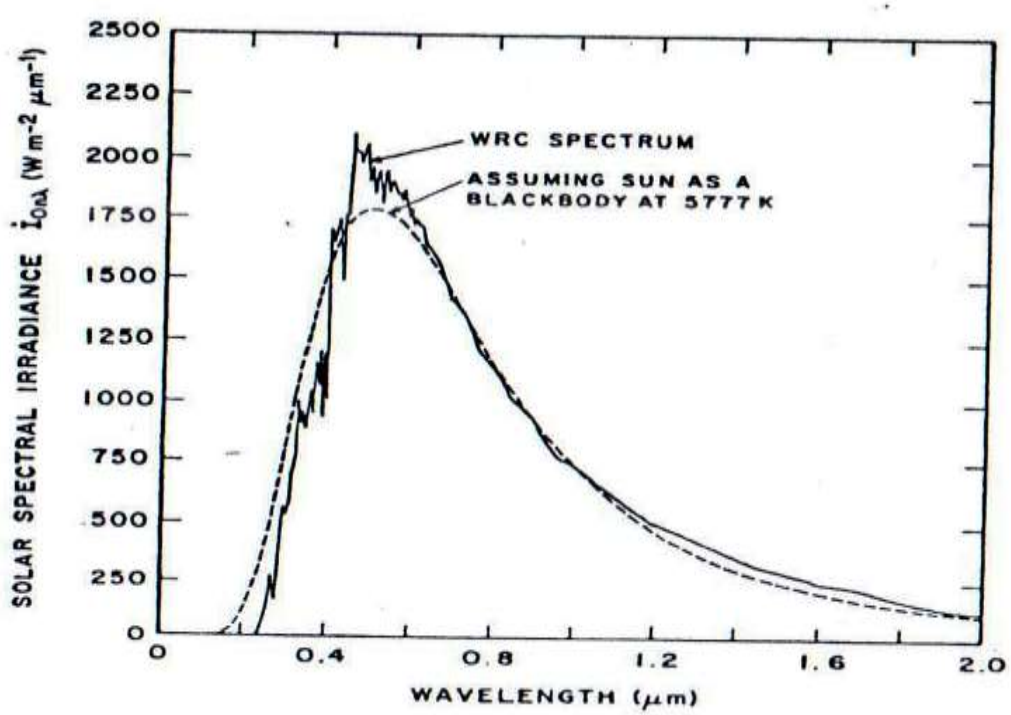


Figure 3 : Comparaison du spectre solaire normale de référence avec celui du rayonnement d'un corps noir à la température de 5777 K

L'intégrale sur tout le spectre du r.é.m, du spectre de l'irradiance solaire normale de référence hors atmosphère $\dot{I}_{sc}(\lambda)$, donne comme résultat la constante solaire, notée $\dot{I}_{sc}(\lambda)$, et dont l'u.S.I est le Wm^{-2}. Telle qu'elle a été adoptée en 1981 par la "Commission pour les Instruments et Méthodes d'Observation"(CIMO), sa valeur communément admise est la suivante [8] [9] :

$$\dot{I}_{sc} = 1367 \text{ wm}^{-2} \quad (1.2.3)$$

c) Irradiance solaire normale hors atmosphère à une distance quelconque

Comme la puissance (ou l'intensité) du r.é.m est inversement proportionnelle ($\sim$) au carré de la distance (r) de l'endroit où elle est mesurée par rapport (p /r) à la source (le soleil ici), l'irradiance solaire normale hors atmosphère en un site donné sur la terre et à un instant (t) donné d'un jour (ou d'une date) numéro J de l'année, est notée $\dot{I}_{on}$, puis donnée par la relation ci-après :

$$\dot{I}_{on} = \frac{{r_o}^2}{r^2}\dot{I}_{sc} = E_o\,\dot{I}_{sc} \quad (1.2.4)$$

Dans l'éq.(1.2.4), $E_o = \frac{{r_o}^2}{r^2}$ est le facteur de correction pour l'excentricité de l'orbite terrestre (autour du soleil). Il varie avec la date (J) de l'année (ou avec la déclinaison solaire δ). Ces variations (de E_O avec J ou δ) sont disponibles dans la littérature sous forme d'expressions analytiques ou de tables [8] [9].

d) Irradiance et irradiations solaires hors atmosphère sur une surface horizontale au sol à une distance quelconque

Selon les résultats des paragraphes précédents, pour une surface horizontale au sol (sur la terre) en un site donné et à une date (c-à-d un jour) considéré(e) de l'année, on obtient les relations ci-dessous [8] [9] [10] [15] :

1. Irradiance solaire hors atmosphère sur une surface horizontale au sol en un site donné, puis à un instant d'une date donnée :

$$\dot{I}_{oh} = (\sin h)\dot{I}_{on} = \cos\theta\,\dot{I}_{on} \quad (1.2.5)$$

où $\sin h$ (ou $\cos\theta$) $et\ \dot{I}_{on}$ sont donnés par les éqs.(1.1.2) et (1.2.4) respectivement ;

2. Irradiation solaire horaire hors atmosphère sur une surface horizontale au sol en un site et un jour donnés :

En appelant w_{si} l'angle horaire du soleil à l'instant correspondant au milieu de la durée horaire considérée, si on intègre l'éq.(1.2.5) sur cette durée horaire, les calculs donnent (en Jm^{-2}) :

$$G_{oh} = E_o \dot{I}_{sc}(36OO)\left[\sin\delta\sin\Phi + \frac{24}{\pi}\sin\frac{\pi}{24}\cos\delta\cos\Phi\, cos\omega_{si}\right] \quad (1.2.6)$$

Si par contre la durée horaire est considérée entre deux instants $t_1\ et\ t_2$ (en heures, temps solaires vrais), on obtient (en Jm^{-2}) :

$$G_{oh}|_{t_1}^{t_2} = E_o \dot{I}_{sc}(3600)\left\{(\sin\delta)(sin\,\Phi)(t_2 - t_1) + \frac{12}{\pi}(\cos\delta)(\cos\Phi)[\sin(15t_1) - sin(15t_2)]\right\} \quad (1.2.7)\ ;$$

3. Irradiation solaire journalière hors atmosphère sur une surface horizontale au sol en un site et un jour donnés :

En intégrant l'éq.(1.2.5) p/r au temps (t) entre les instants $t_{sr}\ et\ t_{ss}$ du lever et du coucher solaires, c- à-d p/r à l'angle horaire du soleil (w_s) entre les valeurs $w_{sr}\ et\ w_{ss}$, on obtient l'une ou l'autre des identités ci-après pour cette irradiation (en Jm^{-2}) :

$$G_o = \frac{24}{\pi}E_o \dot{I}_{sc}(3600)\left[\cos\delta\cos\Phi\sin\omega_{ss} + \frac{\pi}{180}\omega_{ss}\sin\delta\sin\Phi\right]$$

(1.2.8)

$$G_o = \frac{24}{\pi}E_o \dot{I}_{sc}(3600)\sin\delta\sin\Phi\left[\frac{\pi}{180}\omega_{ss} - tg\ \omega_{ss}\right]$$

(1.2.8')

$$G_o = \frac{24}{\pi}E_o \dot{I}_{sc}(36OO)\cos\delta\cos\Phi\left[\sin\omega_{ss} - \frac{\pi}{180}\omega_{ss}cos\ \omega_{ss}\right]$$

(1.2.8'')

I.2.2.2. Rayonnement solaire au sol après traversée d'une atmosphère sans nuages

a) Processus d'atténuation, puis irradiances solaires spectrales et totales au sol par ciel sans nuages

a.1) Aperçu sur ces processus et irradiance solaire directe spectrale au sol sur une surface normale au faisceau du R.S par ciel sans nuages

A la traversée d'une atmosphère sans nuages, le faisceau du R.S direct voit son intensité atténuée suite à divers phénomènes d'extinction (ou d'atténuation de cette intensité) par les constituants atmosphériques. Ces phénomènes constitués de diffusions et d'absorptions font que, sur une surface horizontale ou inclinée au sol, le R.S incident est global ; il comprend deux parties : le R.S direct et le R.S diffus. Dans le cas de l'irradiance solaire directe au sol sur une surface normale au faisceau (du R.S) par ciel sans nuages, la distribution spectrale est exprimée par la loi ci-après de Lambert-Bouguer-Beer [7] [16] :

$$\dot{I}_n(\lambda) = \dot{I}_{on}(\lambda) exp[-mk(\lambda)] \quad (1.3.1)$$

avec $\dot{I}_n(\lambda)\ et\ \dot{I}_{on}(\lambda)$ exprimées en $Wm^{-2}\ \mu m^{-1}$(comme u.h.s), puis :

$$mk(\lambda) = \sum_{i=1}^{j} m_i k_i(\lambda) \quad (1.3.2)$$

Dans les éqs.(1.3.1) et (1.3.2), $k_i(\lambda)$ est le coefficient d'extinction (ou d'atténuation, ou d'affaiblissement) monochromatique (de l'intensité du R.S direct normal) suite au i-ième processus d'extinction. Le produit $m_i k_i(\lambda)$ est l'épaisseur optique d'extinction monochromatique relative au i-ième processus ; il est adimensionnel. Si $k_i(\lambda)$ est adimensionnel, m_i l'est aussi et on l'appelle alors masse optique relative de l'espèce (ou du groupe d'espèces /constituants atmosphériques) responsable(s) du i-ième processus d'extinction. Si $k_i(\lambda)$ a la dimension de l'inverse d'une longueur (L^{-1}), alors m_i a celle d'une longueur (L) et on l'appelle masse optique ou longueur du trajet optique de l'espèce (ou du groupe d'espèces) précédente(s). Le facteur exponentiel dans l'éq.(1.3.1) est le coefficient de transmission monochromatique (du R.S direct) suite à tous les (j) processus. On le note :

$$\tau(\lambda) = \prod_{i=1}^{j} \tau_i(\lambda) \quad (1.3.3)$$

où le coefficient $\tau_i(\lambda)$, de transmission monochromatique (du R.S direct) suite i-ième processus, est donné par :

$$\tau_i(\lambda) = exp[-m_i k_i(\lambda)] \quad (1.3.4)$$

Parmi les nombreux processus d'extinction (de l'intensité du R.S direct) qui ont fait l'objet de diverses études théoriques et expérimentales dans la littérature [6] [7] [16], les principaux qu'on peut citer sont les suivants (avec les coefficients de transmission monochromatiques correspondants) :

1) la diffusion moléculaire (ou diffusion de Rayleigh) ; $\tau_r(\lambda)$;
2) la diffusion (plus un peu d'absorption) par les aérosols (ou diffusion de Mie) ; $\tau_a(\lambda)$;
3) l'absorption sélective par l'ozone : $\tau_o(\lambda)$;
4) l'absorption sélective par la vapeur d'eau : $\tau_{wa}(\lambda)$;
5) l'absorption sélective par tous les autres gaz atmosphériques supposés uniformément mélangés: $\tau_g(\lambda)$.

Avec ces données, l'éq.(1.3.1) se note aussi :

$$\dot{I}_n(\lambda) = \dot{I}_{on}(\lambda)\,\tau_r(\lambda)\,\tau_a(\lambda)\,\tau_o(\lambda)\,\tau_{\omega a}(\lambda)\,\tau_g(\lambda) \quad (1.3.5)$$

Dans les résultats des études (et littérature) citées plus haut sur la description des processus d'extinction et donc sur la connaissance des valeurs (ou expressions) des $m_i, k_i(\lambda) et\ \tau_i(\lambda)$, un grand nombre de paramètres interviennent. Ce sont notamment : la coordonnée h (ou son complément θ) du soleil; les profils (verticaux) de la température (T), la pression (p), l'indice de réfraction (n) et les concentrations relatives des molécules d'air dans une atmosphère standard ; l'épaisseur d'ozone (l_o) et la quantité d'eau condensable (ω) dans une colonne verticale d'atmosphère claire; la masse optique relative de l'air (m_r); celle des aérosols (m'_a); le coefficient de trouble atmosphérique d'Ångström (β) , l'exposant (α) de la l.d.o (dans la formule d'Ångström) ou tout autre indice du trouble atmosphérique ; l'albédo (ρ'_a) d'une atmosphère sans nuages ; l'albédo (ρ_g) du sol ; l'albédo (ω_o) de diffusion par les aérosols ; puis le facteur (F_C) de diffusion vers l'avant par les aérosols.

a.2) Autres irradiances solaires au sol par ciel sans nuages

Partant des données du paragraphe a.1) précédent, on définit comme suit les grandeurs ci-dessous [7] [8] [16] :

1) L'irradiance solaire directe totale au sol sur une surface normale (au faisceau du R .S) par atmosphère sans nuages

On l'exprime en Wm^{-2} (u.S.I) et on l'obtient en intégrant l'éq.(1.3.5) sur tout le spectre du r.é.m :

$$\dot{I}_n = \int_0^\infty \dot{I}_n(\lambda)\, d\lambda$$

(1.3.6)

2) L'irradiance solaire directe totale au sol sur une surface horizontale par ciel sans nuages

Elle s'exprime en Wm^{-2} (u.S.I) et est donnée par :

$$\dot{I}_b = \dot{I}_n \cos\theta = \dot{I}_n \sin h$$

(1.3.7)

où l'indice "b" se réfère au terme anglais "beam", puis $\cos\theta \; et \; \sin h$ sont donnés par l'éq. (1.1.2).

3) L'irradiance solaire diffuse spectrale au sol sur une surface horizontale par ciel sans nuages.

Elle résulte des contributions ci-après exprimées en en $Wm^{-2}\,\mu m^{-1}$ (u.h.s) chacune :

- celle due à la diffusion primaire (du R.S direct) par les molécules d'air (diffusion de Rayleigh) : $\dot{I}_{dr}(\lambda)$;

- celle due à la diffusion primaire (du R.S direct) par les aérosols (diffusion de Mie) : $\dot{I}_{da}(\lambda)$;

- celle due aux réflexions multiples (des R.S direct et diffus primaires) entre le sol et l'atmosphère : $\dot{I}_{dm}(\lambda)$;

d'où son expression :

$$\dot{I}_d(\lambda) = \dot{I}_{dr}(\lambda) + \dot{I}_{da}(\lambda) + \dot{I}_{dm}(\lambda)$$

(1.3.8)

4) L'irradiance solaire diffuse totale au sol sur une surface horizontale par le ciel sans nuages.

Elle s'exprime en Wm^{-2} (u.S.I) et est donnée par l'intégrale de l'éq.(1.3.8) sur tout le spectre du r.é.m :

$$\dot{I}_d = \int_0^{\infty} \dot{I}_d(\lambda) d\lambda$$

(1.3.9)

5) L'irradiance solaire globale spectrale au sol sur une surface horizontale par ciel sans nuages.

Elle s'exprime en $Wm^{-2}\,\mu m^{-1}$ (u.h.s) et est la résultante suivante :

$$\dot{I}(\lambda) = \dot{I}_b(\lambda) + \dot{I}_d(\lambda) = \dot{I}_n(\lambda)\cos\theta + \dot{I}_{dr}(\lambda) + \dot{I}_{da}(\lambda) + \dot{I}_{dm}(\lambda)$$

(1.3.10)

6) L'irradiance solaire globale totale au sol sur une surface horizontale par ciel sans nuages.

Elle s'exprime en Wm^{-2} (u.S.I) et s'obtient par intégration de l'éq.(1.3.10) sur tout le spectre du r.é.m :

$$\dot{I} = \int_o^{\infty} \dot{I}(\lambda) d\,\lambda \qquad (1.3.11)$$

En termes de remarques, on peut signaler que la littérature fait aussi état des expressions des irradiances solaires spectrales et totales, directes, diffuses et globales au sol mais sur des surfaces inclinées (p/r à l'horizontale), par atmosphère sans nuages. Dans la suite, seules des grandeurs énergétiques définies sur une surface horizontale seront considérées.

b) Irradiations solaires sur une surface horizontale au sol par ciel sans nuages.

On les obtient en intégrant les irradiances solaires totales correspondantes sur une durée de temps choisie supérieure à la seconde. Dans les deux cas particuliers où le choix de cette durée porte sur une heure, puis une journée, on définit les types suivants d'irradiations (en Jm^{-2} comme u.S.I) :

1) L'irradiation solaire directe horaire sur une surface horizontale au sol par ciel sans nuages :

$$I_{bh} = \int_{t_i}^{t_{i+1}} \dot{I}_b dt \qquad (1.4.1)$$

où t_i est un instant donné (exprimé en heures, temps solaire vrai) d'une journée au site considéré, puis t_{i+1} est l'instant d'une heure après t_i.

2) L'irradiation solaire directe journalière sur une surface horizontale au sol par ciel sans nuages :

$$I_b = \int_{t_{sr}}^{t_{ss}} \dot{I}_b dt = \int_{t_{sr}}^{t_{ss}} (\dot{I}_n \cos\theta)\, dt \qquad (1.4.2)$$

3) L'irradiation solaire diffuse horaire sur une surface horizontale au sol par ciel sans nuages :

$$D_h = \int_{t_i}^{t_{i+1}} \dot{I}_d\, dt \qquad (1.4.\ 3)$$

4) L'irradiation solaire diffuse journalière sur une surface horizontale au sol par ciel sans nuages :

$$D = \int_{t_{sr}}^{t_{ss}} \dot{I}_d\ dt \qquad (1.4.4)$$

5) L'irradiation solaire globale horaire sur une surface horizontale au sol par ciel sans nuages :

$$G_h = \int_{t_i}^{t_{i+1}} \dot{I}\, dt \qquad (1.4.5)$$

6) L'irradiation solaire globale journalière sur une surface horizontale au sol par ciel sans nuages :

$$G = \int_{t_{sr}}^{t_{ss}} \dot{I}\, dt \qquad (1.4.6)$$

I.2.2.3. Rayonnement solaire au sol après traversée d'une atmosphère réelle

a) Composantes de ce rayonnement

a.1) Description sommaire

En termes des composantes directes, diffuses et globales, diverses irradiances solaires (spectrales et totales) et irradiations solaires (horaires et journalières), sur une surface normale (au faisceau du R.S direct) ou horizontale au sol, ont été décrites par ciel sans nuages dans la sous-section 1.2.2.2. Par ciel (ou atmosphère) comportant des nuages, en tenant compte des effets de ces nuages en plus des processus d'extinction présentés dans ladite sous-section 1.2.2.2, on définit les mêmes types d'irradiances et irradiations que précédemment. On notera en particulier dans la suite par I_b, D et G respectivement, les irradiations solaires directe, diffuse et globale journalières sur une surface horizontale en un site donné, au sol, un jour (ou une date) J considéré(e).

a.2) Mesures

Un premier groupe de techniques d'acquisition des données sur les composantes du R.S est celui des mesures des grandeurs concernées. Pour les mesures directes au sol (ou in situ), on utilise les instruments ci-après (pour les trois types de composantes) dont le détail sur les modèles et leur principe de fonctionnement est disponible dans la littérature [6] [7] [8] :

1. **des pyrhéliomètres** à incidence normale (sigle en anglais: NIP) pour la mesure des irradiances solaires directes normales spectrales (si on associe des monochromateurs à ces appareils) et totales, ainsi que des irradiations solaires directes normales (si ces appareils sont munis d'intégrateurs) ;

2. **des pyranomètres** pour la mesure des irradiances solaires globales (sur une surface horizontale ou inclinée) spectrales (si on dispose des monochromateurs sur ces appareils) et totales, ainsi que des irradiations solaires globales (si ces pyranomètres sont munis d'intégrateurs);

3. **des pyranomètres** munis de dispositifs pare-soleil (en formes d'anneaux, bandes, rubans ou disques)pour la mesure des irradiances (spectrales et totales) et des irradiations solaires diffuses.

Les mesures des grandeurs précédentes peuvent également être menées à partir d'engins aéroportés (comme des ballons-sondes et des avions) et même à partir de satellites géostationnaires.

a.3) Indice de clarté journalier

En un site et un jour considérés, on définit l'indice de clarté journalier, noté k_t, par le rapport entre l'irradiation solaire globale journalière (G) sur une surface horizontale par ciel (ou atmosphère) réel(le) et l'irradiation solaire correspondante hors atmosphère (G_o) :

$$k_t = \frac{G}{G_O} \qquad (1.5.1)$$

b) Effet des nuages

L'extinction du R.S par les nuages est en général plus importante que celle effectuée par les autres constituants atmosphériques. Pour l'estimer, divers paramètres sont utilisés, entre autres : les dimensions de ces nuages (étendue et épaisseur), leur altitude, leur distribution spatiale, leur stationnarité ou intermittence, ainsi que leur nature (ou type) et donc leurs

propriétés optiques (ex : le coefficient de réflexion entre la base d'un nuage et le sol, la quantité d'eau liquide et de glace contenue dans un nuage, etc.). La connaissance de ces paramètres est obtenue à l'aide de méthodes complémentaires telles que l'observation à partir du sol, les sondages atmosphériques via des ballons et avions, la télédétection (usage de systèmes LIDAR par ex.), l'analyse d'images satellitaires et le traitement théorique [6].

La littérature fait état de divers modèles d'estimation de l'irradiation solaire globale (journalière en général: G) sur une surface horizontale en un site donné. Dans ces modèles, à l'image de ceux figurant dans [17], les parts du R.S direct (I_b) et du R.S diffus (D) sont généralement séparées. De plus, ces modèles varient selon le nombre et le type de paramètres (parmi ceux cités plus haut) utilisés. Dans la plupart de cas, les paramètres : nébulosité totale, notée Ne (ou exprimée en octas : $\frac{Ne}{8}$) ou alors C (fraction totale du ciel couvert par les nuages), puis visibilité horizontale, notée Vis (utilisée dans le cas de brumes, brouillards ou nuages très bas), sont souvent rencontrés [12] [17].

I.3. Problématique, cadre spécifique et motivation du sujet du travail

I.3.1. Problématique

Comme déjà montré à travers la section 1.2.2, en termes d'irradiances ou d'irradiations, l'énergie du R.S arrivant en un site donné au sol est entre autres définie p/r à la disposition de la surface réceptrice (de ce R.S), qui peut être horizontale ou inclinée (p/r à l'horizontale). De cette énergie, il est souvent nécessaire de séparer les parts des composantes directe et diffuse. C'est le cas notamment dans le calcul du R.S global sur une surface inclinée [8] [18], l'analyse du bilan radiatif dans une portion donnée d'atmosphère réelle (généralement nuageuse) [6], ainsi que la description des processus biologiques [17] [19].

Par ailleurs, des données fiables sur l'irradiation solaire globale sur une surface horizontale en un site donné (parfois avec ses composantes directe et diffuse), sont essentielles comme « entrées » (inputs en anglais) dans des projets (ex. : en termes de dimensionnement) d'installations de différents types de systèmes de conversion photothermique ou photovoltaïque, et /ou de stockage de l'énergie du R.S (en ce site). Ces projets peuvent concerner l'un ou l'autre des secteurs ci-après de la vie socio-économique : l'énergie, l'agriculture, l'habitat (architecture), le tourisme et les loisirs, les transports, les télécommunications, la santé, l'industrie et l'artisanat. Comme déjà signalé plus haut (sous-section I.2.2.3, paragraphe a.2) des mesures (directes au sol, à partir d'engins aéroportés ou à

partir de satellites) constituent un premier groupe de techniques d'acquisition des données sur l'irradiation solaire globale et (éventuellement) ses composantes directe et diffuse.

Cependant, en raison particulièrement des coûts très élevés des appareils correspondants de mesure (pyrhéliomètres et pyranomètres, avec leurs accessoires), des données sur les trois types précédents d'irradiations solaires sont souvent indisponibles aux sites où on souhaiterait installer les projets (de conversion et/ou de stockage) cités plus haut. Ou alors, des mesures sur ces irradiations (en des sites donnés) ont été opérées sur une certaine période, puis (pour une raison ou une autre) ont été interrompues. Dans les deux cas, l'installation de tels projets en ces sites devient problématique.

I.3.2. Alternative : cadre spécifique du présent travail

Dans une solution alternative à la problématique d'acquisition des données sur l'irradiation solaire globale (et ses composantes directe et diffuse), on recourt à un deuxième groupe de techniques, qui constituent la modélisation et dans laquelle s'inscrit spécifiquement le présent travail. Ces techniques consistent à estimer les irradiations précédentes, en des sites où les mesures (sur ces irradiations) n'ont pas été opérées ou ont été interrompues (dans ce temps), à l'aide de données (généralement disponibles) d'une ou plusieurs variable(s) climatologique (s).

Un très grand nombre de travaux de recherche sur ce groupe de techniques ont été (et sont encore actuellement) réalisés pour des sites à travers diverses régions du monde. Dans un sous-groupe de ces travaux, on estime, en termes de moyennes mensuelles ou annuelles sur une période de plusieurs années, l'irradiation solaire globale journalière (G) par atmosphère réelle (entre autres via la grandeur indice de clarté journalier k_t), à l'aide de diverses associations de variables climatologiques. La suite cite quelques exemples de réalisations dans ce sous-groupe.

I.3.2.1. Modèles d'estimation de G en fonction de S

Plusieurs types de modèles d'estimation de l'irradiation solaire globale journalière (G) par atmosphère réelle en fonction de la durée réelle d'ensoleillement (S) sont rencontrés dans la littérature. Ils se présentent sous l'une ou l'autre des formes ci-après :

1. Relation linéaire $\frac{G}{G_c}$ vs s_r , de type « modèle d'Ångström » [11] [12] [20], où $s_r = \frac{S}{S_O}$ est la durée relative d'ensoleillement et G_c l'irradiation solaire globale journalière par ciel très clair ;

2. Relation linéaire $k_t = \frac{G}{G_0}$ vs s_r, de type « modèle d'Ångström-Black » également appelé modèle d'Ångström-Prescott, dans laquelle la grandeur G_c (assez difficile à estimer) est remplacée par l'irradiation solaire globale journalière hors atmosphère (G_0) [11] [12] [21] [22] [23] [24] [25];

3. Relation entre $\frac{G}{G_c}$ d'une part, puis s_r ou $s'_r = \frac{S}{S'_O}$ d'autre part (où comme signalé à la fin de la sous –section 1.2.1.2, s'_O est l'expression de la durée théorique d'ensoleillement modifiée pour tenir compte de l'horizon naturel d'un site), en forme de : polynôme du second degré, polynôme du troisième degré, fonction puissance, fonction logarithmique et fonction exponentielle [26] ;

4. Relation entre k_t et s_r en termes des polynômes et fonctions précédents [27]-[32].

I.3.2.2. Modèles d'estimation de G en fonction d'autres variables ou groupes de variables climatologiques

On rencontre aussi des modèles où G est liée aux variables (ou associations de variables) climatologiques suivantes :

1. Fraction totale (notée C) du ciel couvert par les nuages, également appelée nébulosité (exprimée en octas et notée $\frac{N_e}{8}$) [8] [11] [12] ;

2. Quantité d'eau condensable dans l'atmosphère (W_{at}) et s_r [33] ;

3. Amplitude thermique (ΔT=$T_{max} - T_{min}$) [34]-[36] ;

4. Amplitude thermique (ΔT) et s_r [37]-[40];

5. Durée relative d'ensoleillement (s_r), température normalisée de l'air ($T_n = \frac{T}{T_{max}}$) et humidité relative de l'air (RH) [41] [42] ;

6. Durée relative d'ensoleillement (s_r), température de l'air, nombre de jours de pluies et précipitations [43] ;

7. Durée relative d'ensoleillement (s_r), température de l'air, précipitations et fraction du ciel couvert par les nuages [44].

I.3.3. Motivation de ce travail

Les données d'irradiations solaires (directe, diffuse et globale) acquises à l'aide de techniques de modélisation, constituent donc un complément indispensable à celles obtenues au moyen de techniques de mesure. Un intérêt certain de la disponibilité de telles données réside d'abord dans leur utilisation (comme déjà mentionné dans la section 1.3.1) comme « entrées » de tout projet d'installation d'un système de conversion et/ou stockage de l'énergie du R.S. Elles sont également indispensables dans l'établissement des cartes (saisonnières ou annuelles) des irradiations solaires (ex: $\mathrm{I}_b, D\ et\ G$) et de la durée d'ensoleillement (S) (d'un pays ou d'une région donné(e)) [45] [46], ainsi que dans la définition d'une année météorologique typique (sigle en anglais: TMY) [47] pour chacune des grandeurs précédentes (entre autres).

Or, un coup d'œil sur l'utilisation de l'énergie du R.S au Burundi montre les faits suivants entre autres. A cause des coupures fréquentes du courant du secteur dans différentes villes et agglomérations, il y a augmentation (au fil du temps) du nombre d'installations de systèmes de conversion photovoltaïque et photothermique (de l'énergie du R.S) à des fins de consommation (de l'énergie et de l'eau chaude) au niveau tant communautaire que des ménages. Cependant, un tel engouement est contrarié par une quasi-inexistence d'études sérieuses sur la faisabilité des installations, qui inclut entre autres : la disponibilité de la ressource (énergie du R.S) aux sites ad hoc ; les caractéristiques des dispositifs installés (notamment en termes de rapports coût /rendement/longévité) qui sont par ex. des champs de panneaux de cellules solaires photovoltaïques, ceux de panneaux de capteurs solaires photothermiques, des chauffe-eau solaires, des batteries de stockage (d'électricité) ; l'efficacité de la gestion (ex. : suivi et maintenance) des installations. En rapport avec la disponibilité de la ressource (solaire) par exemple, on note une rareté de publications scientifiques sur la modélisation des irradiations solaires en des sites du Burundi. Pourtant, une consultation des archives dans les services de l'Institut Géographique du Burundi (IGEBU en sigle) sis à Gitega, indique la disponibilité de données de mesure effectuées sans (ou parfois avec) interruption sur plusieurs dizaines d'années, pour diverses variables météorologiques. Celles-ci sont notamment la durée réelle d'ensoleillement journalier (S), l'irradiation solaire globale journalière (G), la pression atmosphérique (p), la vitesse du vent

en module (v) et orientation (θ), des caractéristiques de l'air (température moyenne (T), température maximum (T_{max}) et minimum (T_{min}), température de rosée (T_d), pression de vapeur (p_v), humidité relative (RH)) et la précipitation (en mm). La même consultation (via plusieurs séjours à l'IGEBU) montre que les données sur S sont plus disponibles (c-à-d plus nombreuses et complètes) que celles de G en divers sites.

C'est donc toutes ces raisons qui ont motivé le choix du présent travail de recherche, à savoir : l'estimation, en divers sites du Burundi, de l'irradiation solaire globale journalière (G) (via l'indice de clarté k_t), en fonction de la durée réelle d'ensoleillement (S) (via les deux types de durée relative d'ensoleillement (s_r) et (s'_r), chacun à son tour).

I.4. Description du sujet de recherche choisi et des objectifs

I.4.1. Formulations choisies

Parmi les modèles d'estimation de G en fonction de S décrits dans la sous-section 1.3.2.1, le choix du présent travail porte sur cinq les couples ci-après de relations $k_t \; vs \; s_r \; et \; k_t \; vs \; s'_r$: (i) relations linéaires, (ii) fonctions logarithmiques, (iii) fonctions exponentielles,

(iv) fonctions puissances et (v) polynômes du second degré. En termes des moyennes mensuelles des valeurs journalières (moyennes opérées sur une assez longue période) des grandeurs en jeu ici, ces formulations se notent comme suit :

1. Relations linéaires simples :

$y = a_1 x + b_1$ (1.6.1); $y = a'_1 x' + b'_1$ (1.6.1')

2. Fonctions logarithmiques :

$y = a_2 \ln x + b_2$ (1.6.2); $y = a'_2 \ln x' + b'_2$ (1.6.2′)

3. Fonctions exponentielles :

$y = b_3 e^{a_3 x}$ (1.6.3); $y = b'_3 e^{a'_3 x'}$ (1.6.3′)

4. Relations en forme de puissance :

$y = b_4 (x)^{a_4}$ (1.6.4); $y = b'_4 x'^{a'_4}$ (1.6.4′)

5. Polynômes du second degré :

$y = a_5 + b_5 x_1 + c_5 x_2$ (1.6.5); $y = a'_5 + b'_5 x'_1 + c'_5 x'_2$ (1.6.5')

Dans ces formulations, $y = \bar{k}_t$, $x = x_1 = \bar{s}_r$, $x' = x'_1 = \bar{s'}_r$, $x_2 = x^2 = (\bar{s}_r)^2$,

$x'_2 = x'^2 = (\overline{s'}_r)^2$.

I.4.2. Objectifs

I.4.2.1. Objectif principal

En termes de moyennes mensuelles (sur une assez longue période de temps) des valeurs journalières des grandeurs en jeu, le présent travail vise l'utilisation des variables durées relatives d'ensoleillement journalières s_r et s'_r (et donc durée réelle d'ensoleillement S) pour estimer la grandeur indice de clarté journalier k_t (et donc irradiation solaire globale journalière G) en différents sites du Burundi.

I.4.2.2. Objectifs spécifiques

Plus spécifiquement, sur base des données des variables $x = \bar{s}_r,\ x' = \overline{s'}_r\ et\ y = \bar{k}_t$ qu'on aura obtenues pour chacun des sites identifiés, il s'agira de :

(i) établir tous les cinq couples de formulations de la section 1.4.1. (éqs.(1.6.1)) à (1.6.5')) via la détermination des coefficients $(a_i, b_i, c_i)\ et\ (a'_i, b'_i, c'_i)$ (i variant de 1 à 5) qui y figurent ;

(ii) évaluer l'efficacité de chacune de ces formulations à l'aide du calcul de divers tests statistiques complémentaires ;

(iii) interpréter les résultats obtenus, puis formuler des conclusions et des perspectives.

CHAPITRE II : DEMARCHE METHODOLOGIQUE

II.1. Données de base

II.1.1. Moyennes mensuelles de S et G

II.1.1.1. Sites et périodes d'études choisis

Les données sur les moyennes mensuelles $\bar{S}$ et $\overline{G}$, respectives de la durée réelle d'ensoleillement journalier et de l'irradiation solaire journalière utilisées dans ce travail, nous ont été aimablement prodiguées par les services de l'IGEBU. L'objectif à ce niveau-ci était de sélectionner des sites du Burundi pour lesquels des données de $\overline{S}$ et $\bar{G}$ étaient simultanément disponibles et sur une plus longue période récente possible (de plusieurs années), avec le moins de discontinuités (des mesures) possibles. De plus, ces sites devaient couvrir le plus de régions éco-climatiques distinctes possibles du pays.

Les premières données offertes par l'IGEBU (lors d'un premier séjour au sein de ses services) ayant (après analyse) été trouvées invraisemblables (et donc inutilisables), il a fallu un deuxième séjour pour obtenir des données plus fiables (mais moins récentes). Le résultat de la sélection (et donc de la collecte des données en questions ici) se rapporte en définitive aux six sites montrés sur la carte de la figure 2.1 (fig. 2.1) et pour lesquels les périodes des moyennes $\bar{S}$ et $\bar{G}$ disponibles sont assez longues (d'une vingtaine d'années environ). Ces périodes, ainsi que les coordonnées des sites considérés, sont représentées dans le tableau 2.1, où l'altitude est comptée au-dessus du niveau de la mer (n.d.l.m).

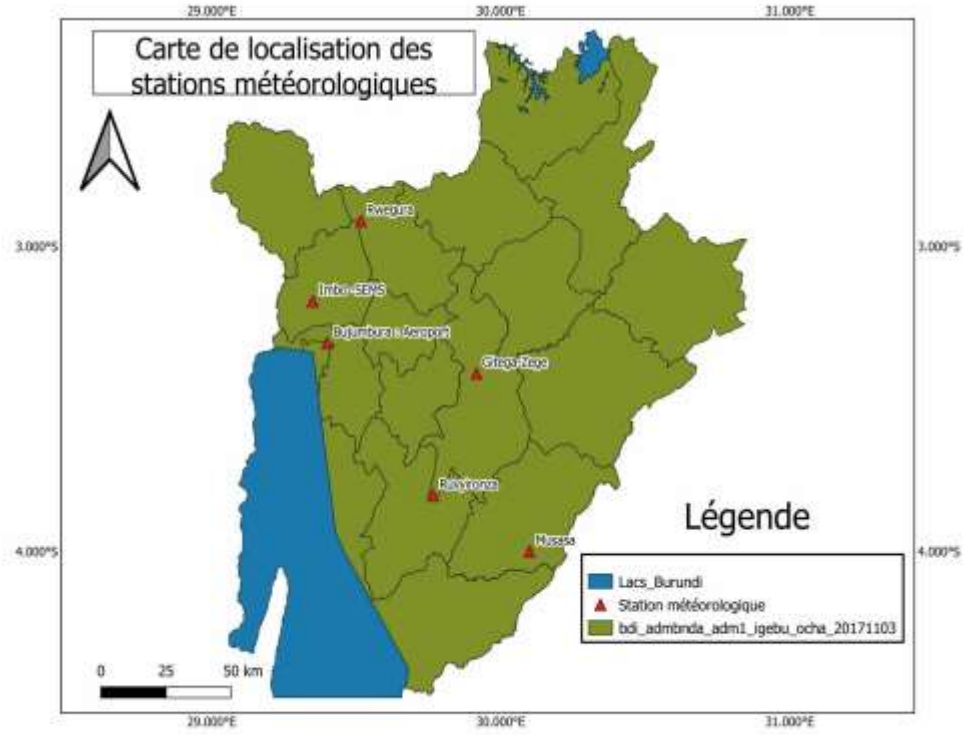


Figure 4 : Représentation des six sites de cette étude sur une carte du Burundi

Tableau 1 : Périodes des données de $\overline{S}$ et $\overline{G}$, puis coordonnées géographiques des sites de cette étude

Site	Période des données	Durée (ans)	Altitude (m)	Latitude Φ (Sud)	Longitude L (Est)
Bujumbura-Aéroport	1968-1988	21	783	$3°19' S$	$29°19'$
Gitega-Zege	1968-1989	22	1663	$3°24' S$	$29°55'$
Imbo-SEMS	1968- 1987	20	820	$3°11' S$	$29°21'$
Musasa	1968-1987	20	1260	$4°00' S$	$30°06'$
Ruvyironza	1968-1988	21	1822	$3°49' S$	$29°46'$
Rwegura	1968-1990	23	2120	$2°55' S$	$29°31'$

II.1.1.2. Valeurs de $\overline{S}$ et $\overline{G}$

Telles que recueillies auprès de l'IGEBU, pour les sites et périodes de cette étude, les moyennes mensuelles $\bar{S}$ (exprimées en heures) et $\bar{G}$ (initialement en cal/cm^2/jour, puis converties en J/cm^2/jour) sont présentées dans les tableaux 2 et 3 respectivement.

Tableau 2 : Valeurs de $\bar{S}$ (en heures)

Sites / Mois	Bujumbura-Aéroport	Gitega-Zege	Imbo-SEMS	Musasa	Ruvyironza	Rwegura
Janvier	5,323	5,010	4,806	5,310	4,458	4,429
Février	6,150	5,164	5,354	6,257	4,700	5,514
Mars	5,848	5,103	5,471	5,942	4,923	4,671
Avril	6,193	4,983	5,580	5,953	4,840	3,740
Mai	7,355	6,277	6,474	6,932	6,061	4,300
Juin	8,960	8,053	7,817	9,000	7,717	6,733
Juillet	8,897	9,242	7,877	9,465	8,016	7,816
Août	8,523	8,306	7,684	9,174	7,294	7,065
Septembre	7,207	6,233	6,633	7,940	6,240	6,043
Octobre	6,603	6,016	6,139	6,274	5,677	5,590
Novembre	5,533	4,817	5,010	5,950	4,283	4,220
Décembre	5,910	5,235	4,977	5,887	4,216	4,977
Moyenne Annuelle	**6,881**	**6,214**	**6,154**	**7,009**	**5,711**	**5,425**

Tableau 3 : Valeurs de $\bar{G}$ (J/cm^2/jour)

Sites / Mois	Bujumbura-Aéroport	Gitega-Zege	Imbo-SEMS	Musasa	Ruvyironza	Rwegura
Janvier	1709,62	1504,8	1560,93	1738,88	1345,96	1467,18
Février	1759,78	1592,58	1588,40	1751,42	1463,00	1442,10
Mars	1876,64	1651,10	17231,12	1918,62	1592,58	1421,20
Avril	1830,84	1550,78	1677,97	1826,66	1500,62	1262,36
Mai	1835,02	1554,96	1726,94	1885,18	1580,04	1270,72
Juin	1818,30	1663,64	1707,83	1943,70	1642,74	1471,36
Juillet	1818,30	1717,98	1695,69	1897,72	1630,20	1559,14
Août	1860,10	1617,66	1700,07	1943,70	1634,38	1609,30
Septembre	1872,64	1634,38	1672,00	1931,16	1613,48	1546,60
Octobre	1851,74	1538,24	1681,55	1830,84	1513,16	1492,26
Novembre	1784,86	1500, 62	1585,41	1726,34	1383,58	1358,50
Décembre	1793,22	1588,40	1634,38	1776,50	1429,56	1429,56
Moyenne Annuelle	**1793,43**	**1549,39**	**1663,53**	**1847,56**	**1512,81**	**1444,19**

Selon les données des deux tableaux précédents, les six sites se classent comme suit par ordre décroissant des moyennes annuelles de $\bar{S}$: Musasa, Bujumbura-Aéroport, Gitega- Zege, Imbo-SEMS, Ruvyironza et Rwegura. Par ordre décroissant des moyennes annuelles de $\bar{G}$, le classement des six sites reste pratiquement le même, à part que Gitega-Zege et Imbo-SEMS

inter-changent leurs places. Par ailleurs, pour les six sites, les mois les plus ensoleillés se classent grosso modo comme suit (par ordre décroissant des valeurs de $\bar{S}$) : juillet, août, juin, septembre, mai et octobre. Cela correspond aux périodes les plus sèches (et donc les moins pluvieuses ou humides) de l'année au Burundi. Pour les six sites également, les valeurs les plus élevées de $\bar{G}$ apparaissent dans le mois d'août. Cependant, dépendamment du site, l'un de ces six mois est remplacé : soit par mars (le plus souvent), soit alors (le moins souvent) par janvier, février ou avril. Ce dernier fait semble normal vu que ces mois correspondent grosso modo à la petite saison sèche en divers endroits du Burundi et se situent en plus au voisinage de la date d'un des deux maxima de la grandeur G_o pour les latitudes des sites d'étude ici (voir dans la section II.1.2).

II.1.2. Moyennes mensuelles de S_o , S'_o et G_o , puis valeurs de x, x′et y

II.1.2.1. Démarche pour leur détermination

Pour chacun des six sites, les valeurs de S_o et S'_o des 365 jours d'une année normale ont été calculées au moyen des éqs. (1.1.8) et (1.1.8′) respectivement. Dans ces calculs, il a été tenu compte des valeurs des latitudes (Φ) de ces sites (voir tableau 2.1 où Φ est négative pour tout site de l'hémisphère Sud), puis des valeurs de la déclinaison solaire (δ) (données par une table [8] en fonction du jour ou de la date (J) de l'année). Des valeurs journalières ainsi obtenues, on a déduit pour chaque site 12 moyennes mensuelles: $\overline{S}_o$ (de S_o), puis $\overline{S'}_o$ (de S'_o). Ensuite, vu les données du tableau 2.2, on a déduit pour chaque site 12 moyennes mensuelles: $\overline{s_r}$ (de s_r), puis $\overline{s'_r}$ (de s'_r) selon les relations ci-après analogues aux éqs.(1.1.9) et (1.1.9') respectivement:

$$x=\bar{s}_r = \frac{\bar{S}}{\bar{S}_O} \qquad (2.1.1)\ ; \qquad x' = \overline{s'_r} = \frac{\bar{S}}{\bar{S}'_0} \qquad (2.1.1')$$

De leur côté, les valeurs de G_o des 365 jours d'une année normale ont été calculées, pour chacun des six sites, à l'aide de l'éq. (1.2.8), dans laquelle ω_{ss} est donnée par l'éq. (1.1.4), la constante solaire $\dot{I}_{sc}$ par l'éq. (1.2.3), puis les valeurs du facteur E_0 de correction pour l'excentricité de l'orbite données en fonction de J par une table [8].

Pour chaque site, des valeurs journalières ainsi obtenues, on a déduit 12 moyennes mensuelles $\bar{G}_O$ (de G_O) . Puis, compte tenu des données du tableau 3, pour chaque site on a déterminé 12 moyennes mensuelles $\bar{k}_t$ (de k_t), selon la relation ci-après (analogue à l'éq.(1.5.1)) :

$$y = \bar{k_t} = \frac{\bar{G}}{\bar{G_0}} \qquad (2.1.2)$$

II.1.2.2. Résultats et commentaires

a) Sur S_0, $\overline{S_0}$ et $\bar{s_r}$

a.1) Sur S_0

Du fait de leur nombre élevé ((365)(6)=2190), les valeurs de S_0 obtenues dans cette étude ne sont pas reproduites ici. Cependant, juste en guise d'illustration, les variations de S_0 avec le n° (J) du jour (et donc de la date) de l'année, sont proposées sur la figure 5 pour l'un des six sites, ex. : Musasa.0

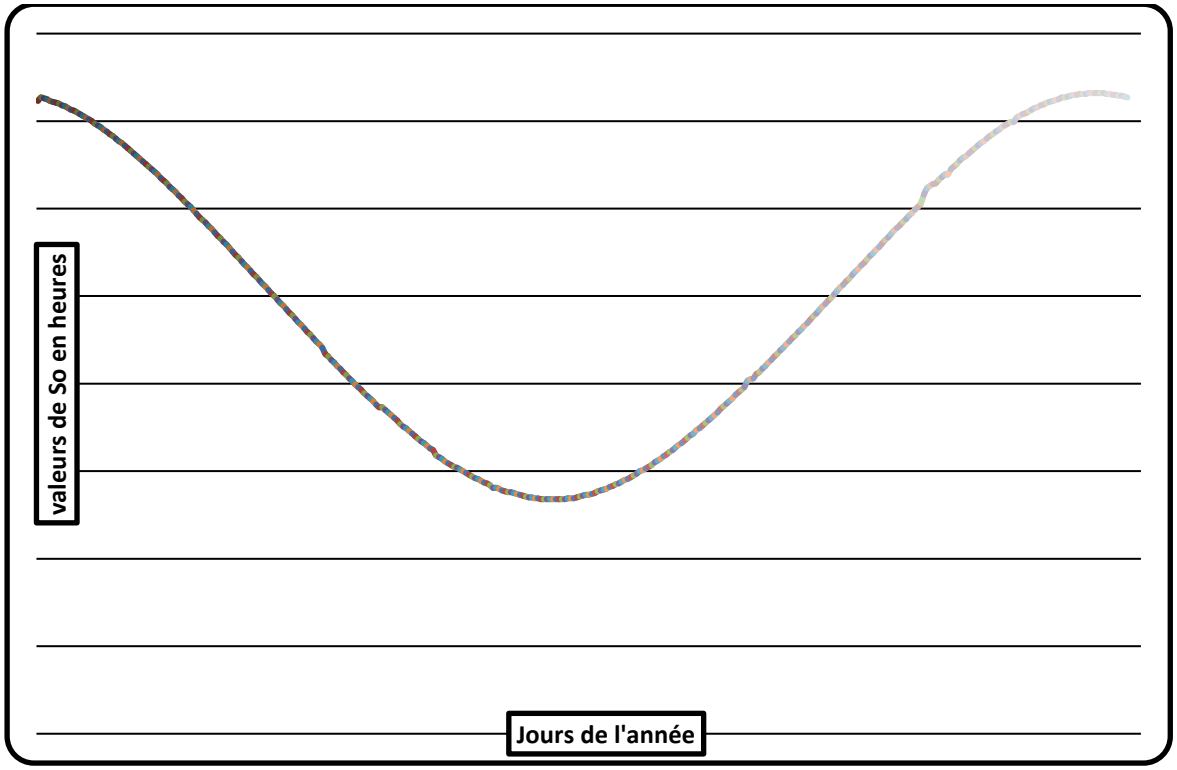


Figure 5 : Variations de S_0 avec J, exemple du site de Musasa.

L'allure de cette courbe S_0 vs J est tout à fait la même pour chacun des autres cinq sites. Elle présente un minimum absolu entre le 19 et le 26 juin (c-à-d au voisinage du 21 juin ou du solstice d'hiver en hémisphère Sud) ; un maximum absolu entre le 19 et le 25 décembre (c-à-d au voisinage du 21 décembre ou du solstice d'été en hémisphère Sud) ; puis un maximum secondaire au début de cette courbe (le 1ier janvier). Ces faits s'expliquent théoriquement comme suit : selon les éqs. (1.1.3) et (1.1.4) et la convention sur le signe de ω_s adoptée dans ce travail, avec Φ qui est négative pour les sites de cette étude, le maximum de δ (=+23,49°

selon les tables [8]) correspond, pour Φ fixée, au maximum de $\cos\omega_{sr}$ ou $\cos\omega_{ss}$, c-à-d au minimum de ω_{ss} et donc de S_0 (vu l'éq.(1.1.8)). Cela a justement lieu entre le 21 et le 23 juin (sur les tables de δ). A l'opposé, selon les éqs. précédentes, le minimum de δ (=-23,49°) correspond, pour Φ fixée, au minimum de $\cos\omega_{sr}$ ou $\cos\omega_{ss}$, c-à-d au maximum de ω_{ss} et donc de S_0. Cela a lieu entre le 21 et le 23 décembre (sur les tables de δ).

a.2) Sur $\overline{S_0}$

Les moyennes mensuelles (et annuelles) $\overline{S_0}$ des données précédentes de S_0 pour les six sites sont proposées dans le tableau 4.

Tableau 4 : Valeurs de $\overline{S}_o$

Sites / Mois	Bujumbura-Aéroport	Gitega-Zege	Imbo – SEMS	Musasa	Ruvyironza	Rwegura
Janvier	12,168	12,173	12,162	12,203	12,194	12,148
Février	12,101	12,104	12,101	12,127	12,117	12,089
Mars	12,015	12,015	12,014	12,016	12,017	12,015
Avril	11,924	11,922	11,927	11,912	11,913	11,933
Mai	11,848	11,844	11,855	11,822	11,826	11,868
Juin	11,811	11,806	11,821	11,772	11,786	11,834
Juillet	11,828	11,823	11,834	11,792	11,801	11,842
Août	11,991	11,882	11,895	11,868	11,874	11,904
Septembre	11,977	11,976	11,977	11,961	11,972	11,979
Octobre	12,066	12,069	12,064	12,081	12,077	12,059
Novembre	12,147	12,147	12,546	12,177	12,169	12,129
Décembre	12,188	12,193	12,181	12,227	12,217	12,165
Moyenne annuelle	**12,005**	**11,996**	**12,031**	**11,997**	**11,997**	**11,997**

Les données de ce tableau 4 montrent que la courbe de variation de $\bar{S}_o$ avec le mois de l'année, suit une allure semblable à celle de la courbe de S_o en fonction de J. Elle présente un maximum secondaire à son début (mois de janvier), un minimum absolu en juin (mois du solstice d'hiver en hémisphère Sud) et un maximum absolu en décembre (mois du solstice d'été en hémisphère Sud). De plus, on note que les moyennes annuelles de $\bar{S}_o$ ont pratiquement la même valeur pour les six sites, à savoir 12,0 heures.

a.3) Sur $\bar{s}_r$

Les valeurs de $x = \bar{s}_r$ pour les six sites sont présentées dans le tableau 5.

Tableau 5 : Valeurs de $x = \bar{s}_r$

Sites / Mois	Bujumbura-Aéroport	Gitega-Zege	Imbo- SEMS	Musasa	Ruvyironza	Rwegura
Janvier	0,437	0,412	0,395	0,435	0,366	0,365
Février	0,508	0,427	0,442	0,516	0,388	0,456
Mars	0.487	0,425	0,455	0,495	0,410	0,389
avril	0,519	0,418	0,468	0,500	0,406	0,313
Mai	0,621	0,530	0,546	0,586	0,513	0,362
Juin	0,759	0,682	0,661	0,765	0,655	0,569
Juillet	0,752	0,782	0,666	0,803	0,679	0,660
Août	0,711	0,699	0,646	0,773	0,614	0,593
Septembre	0,602	0,520	0,554	0,664	0,521	0,504
Octobre	0,547	0,498	0,509	0,519	0,470	0,464
Novembre	0,456	0,387	0,399	0,489	0,352	0,348
Décembre	0,485	0,429	0,409	0,481	0,345	0,409
Moyenne annuelle	**0,573**	**0,518**	**0,513**	**0,586**	**0,476**	**0,453**

La courbe de variation de de $x = \bar{s}_r$ avec le mois de l'année présente une allure similaire à celle de $\bar{S}$, c-à-d une augmentation depuis janvier jusqu'à un maximum en juin (pour Bujumbura-Aéroport) ou juillet (pour les cinq autres sites), puis une diminution jusqu'en décembre. Une telle allure est à l'opposé de celles des variations saisonnières de $\bar{S}_o$. En outre, dans l'ordre décroissant des moyennes annuelles de x $= \bar{s}_r$, les six sites se classent exactement de la même façon qu'ils se classent vis-à-vis des moyennes annuelles de $\bar{S}$ (voir sous-section II.1.1.2). Enfin, ces moyennes annuelles de $x = \bar{s}_r$ varient entre 0,453 et 0,586, ce qui indique que le ciel est moyennement ensoleillé aux sites de cette étude.

b) Sur S'_0, $\overline{S'}_O$ et $\overline{s}'_r$

b.1) Sur S'_0

En raison de leur nombre élevé (2190), les valeurs calculées de S'_0 ne sont pas toutes présentées ici. Par contre, un exemple de courbe de S'_0 avec le numéro (J) du jour de l'année, est proposée sur la figure 2.3 pour l'un des six sites : le même que pour S_0, c-à-d Musasa.

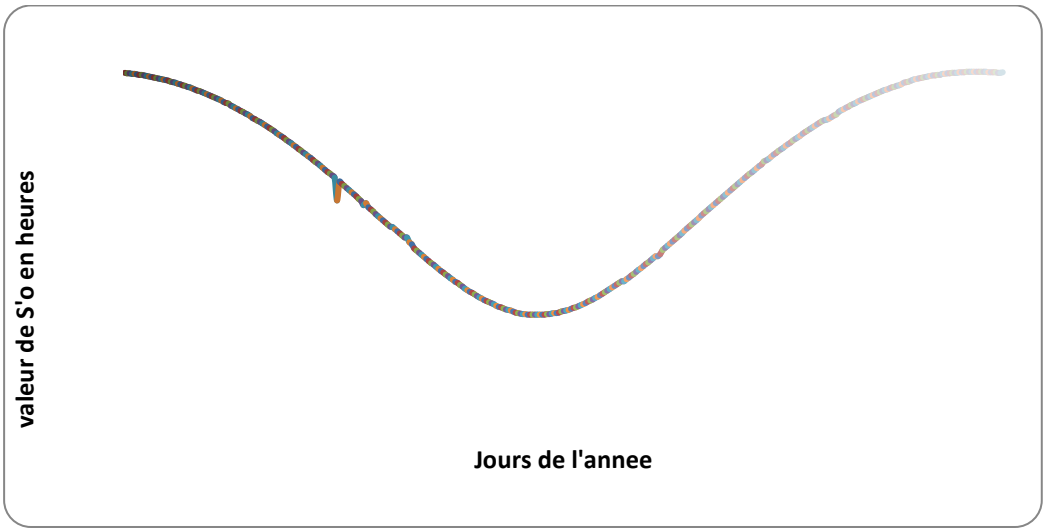


Figure 6 : Variations de S'_0 avec J, exemple du site de Musasa

On observe la même allure des courbes S'_0 vs J pour les cinq autres sites. De plus, cette allure est tout à fait semblable à celle des variations de S_0 avec J, la seule différence résidant dans le fait que, comme l'indique l'éq. (1.1.8') comparée à l'éq.(1.1.8), les valeurs de S'_0 sont systématiquement légèrement inférieures à celles de S_0. Ladite allure présente un minimum absolu au voisinage du 21 juin (solstice d'hiver aux sites de cette étude), un maximum absolu au voisinage du 21 décembre (solstice d'été aux mêmes sites), puis un maximum secondaire au début (le 1er janvier). Ces faits s'expliquent théoriquement de la même façon que pour S_0 (voir paragraphe 2.1.2.2.a.1) à partir des éqs.(1.1.3) et (1.1.8').

b.2) Sur $\overline{S}'_O$

Des données précédentes de S'_0 , les moyennes mensuelles $\overline{S}'_O$ sont présentées dans le tableau 6.

Tableau 6 : Valeurs de $\overline{S'}_O$ (en heures)

Sites / Mois	Bujumbura Aéroport	Gitega-Zege	Imbo – SEMS	Musasa	Ruvyironza	Rwegura
Janvier	11,454	11,458	11,447	11,488	11,480	11,434
Février	11,416	11,419	11,412	11,437	11,431	11,403
Mars	11,346	11,346	11,345	11,348	11,347	11,344
Avril	11,246	11,244	11,249	11,231	11,234	11,254
Mai	11,142	11,136	11,147	11,110	11,119	11,161
Juin	11,082	11,077	11,090	11,044	11,050	11,107
Juillet	11,109	11,105	11,116	11,072	11,082	11,131
Août	11,204	11,199	11,207	11,178	11,185	11,215
Septembre	11,306	11,305	11,329	11,296	11,303	11,310
Octobre	11,391	11,393	11,390	11,404	11,402	11,383
Novembre	11,443	11,447	11,437	11,473	11,465	11,425
Décembre	11,463	11,468	11,455	11,502	11,492	11,441
Moyenne annuelle	**11,300**	**11,300**	**11,302**	**11,299**	**11,298**	**11,301**

Les données de ce tableau montrent que, pour chacun des six sites, la courbe de variation de $\overline{S'}_O$ avec le mois de l'année a une allure semblable à celle de S_0 en fonction de J. Elle présente un minimum absolu en juin (mois du solstice d'hiver pour ces sites), un maximum absolu en décembre (mois du solstice d'été pour ces sites), puis un maximum secondaire au début (en janvier). On remarque en plus que les moyennes annuelles de $\overline{s'}_O$ ont pratiquement la même valeur en ces six sites, qui est de 11,30 heures.

b.3) Sur $x' = \overline{s'}_r$

Les valeurs obtenues de $x' = \overline{s'}_r$ pour les six sites sont proposées dans le tableau 7.

Tableau 7 : Valeurs de $x' = \overline{s'}_r$

Sites / Mois	Bujumbura-Aéroport	Gitega-Zege	Imbo-SEMS	Musasa	Ruvyironza	Rwegura
Janvier	0,465	0,437	0,420	0,462	0,388	0,387
Février	0,539	0,452	0,469	0,547	0,411	0,484
Mars	0.515	0,450	0,482	0,524	0,434	0,412
Avril	0,551	0,443	0,496	0,530	0,431	0,332
Mai	0,660	0,564	0,581	0,624	0,545	0,385
Juin	0,809	0,727	0,705	0,815	0,698	0,606
Juillet	0,801	0,832	0,709	0,855	0,723	0,702
Août	0,761	0,742	0,686	0,821	0,652	0,630
Septembre	0,637	0,551	0,585	0,703	0,552	0,534
Octobre	0,580	0,528	0,539	0,550	0,498	0,491
Novembre	0,484	0,421	0,438	0,519	0,374	0,369
Décembre	0,516	0,456	0,434	0,512	0,367	0,435
Moyenne annuelle	**0,610**	**0,550**	**0,545**	**0,622**	**0,506**	**0,481**

Les données de ce tableau montrent que (comme pour celles de $x = \bar{s}_r$) la courbe de variation de $x' = \overline{s'}_r$ avec le mois de l'année a une allure semblable à celle de $\bar{S}$: elle croît depuis janvier jusqu'à un maximum absolu en juin (pour Bujumbura-Aéroport) ou en juillet (pour les cinq sites restants), puis elle décroît jusqu'en décembre. Cette allure est à l'opposé de celle des variations saisonnières de $\overline{S'}_o$. De plus, dans l'ordre décroissant des moyennes annuelles de $x' = \overline{s'}_r$ (comme pour celles de $x = \bar{s}_r$), les six sites se classent exactement de la même façon qu'ils se classent vis-à-vis des moyennes annuelles de $\overline{S}$ (voir sous-section II.1.1.2). Enfin, le fait que ces moyennes annuelles de $x' = \overline{s'}_r$ varient entre 0,481 et 0,622 est un autre indicateur d'un ciel moyennement ensoleillé dans chacun des six sites considérés.

c) Sur G_O, $\overline{G}_0$ et $\overline{k}_t$

c.1) Sur G_O

Vu également leur nombre élevé (2190), les valeurs de G_O (déterminées comme indiqué dans la sous-section II.1.2.1) ne sont pas non plus reproduites ici. Par contre, en termes d'illustration, la figure 2.4 propose la courbe de variation de G_O avec le numéro (J) du jour de l'année pour l'un des six sites, le même que dans les deux figures précédentes: Musasa.

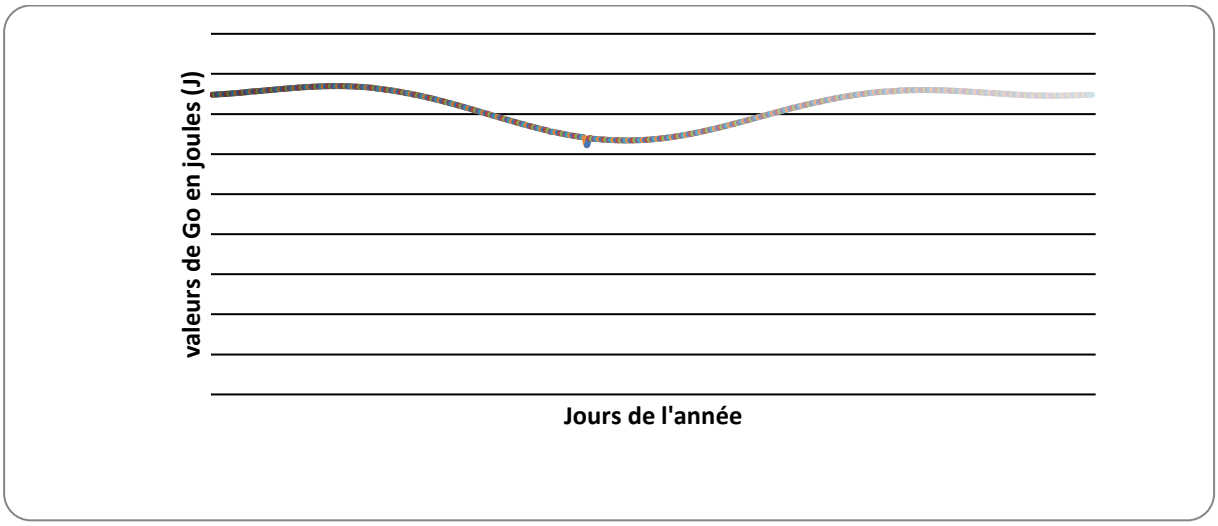


Figure 7 : Variations de G_O avec J, exemple du site de Musasa

On observe la même allure des courbes de variation de G_O avec J pour les cinq autres sites. Sur chacune de ces courbes, G_O croît depuis le 1er jour jusqu'à un maximum absolu entre le 21 et le 24 février, décroît ensuite jusqu'à un minimum absolu entre le 21 et le 24 juin, remonte après jusqu'à un maximum secondaire entre le 21 et le 24 octobre, puis redescend jusqu'à un minimum secondaire entre le 15 et le 17 décembre, avant de remonter enfin (en passant par le 1er janvier) jusqu'au maximum absolu cité plus ci-haut (les 21-24 février). On peut avancer les arguments ci-après comme contribution à l'explication théorique des faits précédents. Le minimum absolu de G_O a lieu au voisinage du 21 juin, c-à-d du solstice d'hiver pour les sites considérés. Là, le soleil décrit son cercle horaire en sa position (sur la sphère céleste) la plus éloignée côté Nord (des sites précédents): au-dessus du tropique du cancer (δ_{max}=+23,49^0). Le maximum secondaire de G_O a lieu au voisinage du 21 décembre, c-à-d du solstice d'été pour les mêmes sites. Là, le soleil décrit son cercle horaire en sa position la plus éloignée côté Sud (des sites en question) : au-dessus du tropique du capricorne (δ_{min}=-23,46^0).

Enfin, les deux maxima (l'un absolu observé vers les 21-24 février, l'autre secondaire observé vers les 21-24 octobre) doivent avoir lieu, selon l'éq.(1.2.8), aux jours J tels que: (i) le soleil décrive son cercle horaire en une position (sur la sphère céleste) la plus proche possible du dessus du site (c-à-d telle que δ soit très voisine de Φ) ; (ii) le facteur E_o soit le plus proche possible de sa valeur maximale ($E_{O\,max} = 1{,}0351$).

c.2) Sur $\bar{G}_0$

Les valeurs obtenues de $\bar{G}_0$ pour les six sites sont présentées dans le tableau 2.8.

Tableau 8 : Valeurs de $\bar{G}_0$ (en J/cm²/jour)

Sites / Mois	Bujumbura-Aéroport	Gitega- Zege	Imbo-SEMS	Musasa	Ruvyironza	Rwegura
Janvier	3748,34	3754,57	3747,06	3775,14	3768,90	3750,01
Février	3825,70	3827,37	3823,08	3844,00	3835,58	3821,32
Mars	3803,33	3795,75	3796,69	3795,30	3798,71	3796,16
Avril	3612,58	3612,40	3613,77	3597,95	3595,00	3622,30
Mai	3365,37	3362,39	3372,89	3340,94	3347,50	3378,48
Juin	3206,22	3211,57	3212,83	3186,10	3195,14	3231,64
Juillet	3346,34	3261,37	3257,60	3235,05	3246,92	3379,91
Août	3376,77	3480,63	3467,78	3439,11	3456,87	3483,40
Septembre	3360,58	3681,84	3683,78	3679,62	3679,17	3685,92
Octobre	3785,88	3781,01	3784,44	3787,82	3785,64	3785,46
Novembre	3751,60	3753,71	3747,53	3776,77	3768,89	3740,85
Décembre	3706,98	3796,14	3696,55	3730,29	3726,33	3685,99
Moyenne	**3574,14**	**3609,90**	**3600,33**	**3599,01**	**3600,39**	**3613,45**

La courbe de variation de $\bar{G}_0$ avec le mois de l'année présente, pour chacun des six sites, une allure semblable à celle de variation de G_0 avec J. Elle possède un maximum absolu en février, un minimum absolu en juin, puis un maximum secondaire en octobre et un minimum secondaire en décembre. De plus, par ordre décroissant des moyennes annuelles de $\bar{G}_0$, les six sites se classent comme suit: Rwegura, Gitega-Zege, Ruvyironza, Imbo-SEMS, Musasa et Bujumbura-Aéroport.

c.3) Sur y =$\bar{k}_t$

Les valeurs obtenues de y =$\bar{k}_t$ sont proposées dans le tableau 2.9.

Tableau 9 : Valeurs de y =$\bar{k}_t$

Sites / Mois	Bujumbura-Aéroport	Gitega-Zege	Imbo-SEMS	Musasa	Ruvyironza	RWEGURA
Janvier	0,4561	0,4008	0,4166	0,4606	0,3571	0,3912
Février	0,4600	0,4161	0,4155	0,4556	0,3814	0,3774
Mars	0,4924	0,4350	0,4560	0,5055	0,4192	0,3744
Avril	0,5068	0,4293	0,4643	0,5077	0,4174	0,3485
Mai	0,5453	0,4625	0,5120	0,5643	0,4720	0,3760
Juin	0,5671	0,5180	0,5316	0,6101	0,5141	0,4553
Juillet	0,5434	0,5268	0,5205	0,5866	0,5021	0,4613
Août	0,5509	0,4648	0,4902	0,5652	0,4728	0,4620
Septembre	0,5572	0,4439	0,4539	0,5248	0,4385	0,4196
Octobre	0,4891	0,4068	0,4443	0,4833	0,3997	0,3942
Novembre	0,4758	0,3998	0,4231	0,4571	0,3671	0,3632
Décembre	0,4837	0,4281	0,4421	0,4762	0,3836	0,3878
Moyenne	**0.5107**	**0,4443**	**0,4642**	**0,5164**	**0,4271**	**0,4009**

Les données de ce tableau montrent que, pour chacun de ces sites, la courbe de variation de y =$\bar{k}_t$ avec le mois de l'année présente l'allure ci-après:

(i) un maximum absolu en juin pour Bujumbura, Imbo, Musasa et Ruvyironza ; en août pour Gitega et en septembre pour Rwegura ;

(ii) un minimum secondaire en novembre pour tous ces sites ;

(iii) un maximum secondaire en décembre pour tous les six sites (à part pour Rwegura où ce maximum a lieu en janvier) ;

(iv) un minimum absolu en janvier pour Bujumbura, Gitega et Ruvyironza ; en février pour Imbo et Musasa; puis en avril pour Rwegura.

Par ailleurs, dans l'ordre décroissant des moyennes annuelles de y =$\bar{k}_t$, les six sites se classent de la même façon que vis-à-vis des moyennes annuelles de $\bar{G}$ (voir dans la section 2.1.1.2), à savoir comme suit: Musasa, Bujumbura-Aéroport, Imbo-SEMS, Gitega-Zege, Ruvyironza et Rwegura. Enfin, on note que les moyennes annuelles de y =$\bar{k}_t$ sont comprises entre 0,40 et 0,52 pour ces sites. D'après certains classements (dans la littérature) des conditions atmosphériques selon l'indice de clarté journalier [24] [46] [48], les valeurs précédentes

caractérisent un ciel (ou une atmosphère) moyennement clair(e) ou partiellement nuageux (nuageuse) aux sites considérés.

II.2. Détermination des coefficients des régressions choisies et des variables statistiques de test des performances de ces régressions

II.2.1. Objectifs visés

Comme signalé dans la sous- section 1.4.2.2, pour chacune des dix (et donc des cinq couples de) formulations choisies (éqs. (1.6.1) à (1.6.5') dans la section 1.4.1), on cherche d'abord à déterminer les coefficients (de régression) (a_i, b_i, c_i) ou (a'_i, b'_i, c'_i) , i=1 à 5, le coefficient de corrélation et celui de détermination. Ensuite, à partir des coefficients de régression (et donc de la formulation) obtenus (obtenue) pour chaque site, on calcule différentes variables statistiques complémentaires en vue d'évaluer les performances de cette formulation.

Dans la suite, la détermination des variables et coefficients précédents sera conduite en considérant deux catégories de formulations : les formulations réductibles à des régressions linéaires simples, puis celles réductibles à des régressions linéaires doubles.

II.2.2. Formulations réductibles à des régressions linéaires simples

II.2.2.1. Formulations concernées et leur réduction

Dans cette catégorie, on reprend ci-dessous les groupes de formulations concernées.

1. Relations linéaires simples

Ce sont les éqs. (1.6.1) et (1.6.2) renumérotées comme suit :

$$y = a_1 x + b_1 \qquad (2.2.1); \qquad y = a'_1 x' + b'_1 \qquad (2.2.1')$$

2. Fonctions logarithmiques

Ce sont les éqs. :

$$y = a_2 \ln x + b_2 \qquad (1.6.2); \qquad y = a'_2 \ln x' + b'_2 \qquad (1.6.2')$$

qui se réduisent aux régressions linéaires simples ci- après :

$$y = a_2 X + b_2 \qquad (2.2.2) \quad et \quad y = a'_2 X' + b'_2 \qquad (2.2.2')$$

$$avec: \quad X = \ln x, X' = \ln x', \; x = \bar{s_r}, \; x' = \overline{s'_r}, \quad y = \bar{k}_t.$$

3. Fonctions exponentielles

Ce sont les éqs. :

$y = b_3 e^{a_3 x}$ (1.6.3) et $y = b'_3 e^{a'_3 x'}$ (1.6.3′)

dont les formes de régressions linéaires simples sont les suivantes:

$y = ax + b$ (2.2.3) et $y = a'x' + b'$ (2.2.3')

avec : a = a_3 , b = $\ln b_3$, a′ = a'_3, $b' = \ln b'_3$, $Y = \ln y$, puis x, x' et y qui continuent à avoir les mêmes significations que précédemment.

4. Relations en forme de puissance

Les éqs. :

$y = b_4 (x)^{a_4}$ (1.6.4) et $y = b'_4 x'^{a'_4}$ (1.6.4′)

prennent quant à elles les formes ci-après de régressions linéaires simples :

$Y = aX + b$ (2.2.4) et $Y = a'X' + b'$ (2.2.4')

avec : a = a_4, $b = \ln b_4$, a′ = a'_4, $b' = \ln b'_4$, $X = \ln x$ et $Y = \ln y$, puis x, x' et y qui gardent les mêmes significations que précédemment.

II.2.2.2. Expressions des coefficients à déterminer

On note comme suit une relation de cause à effet où une variable dépendante « y » (l'effet) est supposée être linéairement corrélée à une variable indépendante « x » (la cause) :

$y = ax + b$ (2. 3.1)

La droite de régression linéaire qui s'ajuste le mieux au nuage des points de coordonnées (x_i , y_i) d'un échantillon aléatoire de « n » observations de ces deux variables (« i » varie de 1 à n), est déterminée au moyen de la méthode des moindres carrés, qui aboutit aux expressions classiques ci-après des estimations des coefficients « a » et « b » de cette régression [49] [50]:

$a = \frac{cov(x,y)}{var(x)} = \frac{m_{xy}}{m_{xx}}$ (2.3.2) ; $cov(x,y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = m_{xy}$ (2.3.3)

$var(x) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = m_{xx}$ (2.3.4)

$b = \bar{y} - a\bar{x}$ (2.3.5) ; $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ (2.3.6) ; $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ (2.3.7)

Le coefficient de corrélation « R » entre x et y est quant à lui donné par :

$$R = \frac{cov\,(x,y)}{([var\,(x)][var\,(y)])^{1/2}} = \frac{m_{xy}}{(m_{xx}m_{yy})^{1/2}} \qquad (2.3.8)$$

Il montre (ou mesure) le degré de corrélation entre x et y. Il varie sur l'intervalle [0 ; 1] lorsque les deux variables croissent dans le même sens, puis sur l'intervalle [-1 ; 0] lorsque les valeurs de l'une des deux variables tendent à croître alors que celles de l'autre décroissent. Une corrélation positive est d'autant meilleure (ou tend plus vers la perfection) que R tend vers 1. Une corrélation négative est d'autant meilleure (ou tend plus vers la perfection) que R tend vers -1 [51].

Enfin, le coefficient de détermination R^2 (donné par le carré des expressions dans l'éq. (2.3.8)), qui varie sur l'intervalle [0 ; 1], mesure la fraction (ou le pourcentage) de la variance (ou de la variabilité) de la variable dépendante « y » qui peut être prédite (ou expliquée)par la variable indépendante « x ».

II.2.2.3. Autres variables statistiques utilisées pour le test des performances des formulations recherchées

En complément des coefficients de R et R^2 précédents, les variables statistiques décrites ci-dessous et dont l'utilité a été démontrée à travers différents travaux [8] [10] [27] [31] [33] [39] [48] [52]-[62], ont été déterminées pour tester la performance de chacune des formulations recherchées.

a) Ecart-moyen

On calcule la valeur de « l'écart-moyen » ci-après (dont le sigle en anglais est MBE, qui signifie « mean bias error »), entre les valeurs calculées $y_{i,calc}$ (selon la formulation obtenue) et les valeurs observées ou mesurées $y_{i,mes}$ (d'un échantillon considéré des valeurs de la variable y) :

$$\text{MBE}= \frac{1}{n}\sum_{i=1}^{n} di \qquad (2.4.1)\ ; \qquad d_i = \ y_{i,calc} - y_{i,mes} \qquad (2.4.1')$$

Le test de l'écart-moyen donne une information sur la performance à long terme de la formulation théorique (y_{calc}) obtenue. L'ajustement (de cette formulation au nuage des points expérimentaux ($x_{i,}y_{i,}$)) est d'autant meilleur que MBE est égale à (ou très proche de) zéro. Une valeur positive de MBE indique qu'en moyenne les $y_{i,calc}$ surestiment les $y_{i,mes}$, tandis

qu'une valeur négative de MBE indique une sous-estimation (moyenne) des $y_{i,mes}$ par les $y_{i,calc}$.

b) Ecart quadratique moyen

On détermine également « l'écart quadratique moyen » ci –après (dont le sigle en anglais est RMSE, qui signifie « root mean square error ») entre les $y_{i,calc}$ et $les\ y_{i,mes}$ (du même échantillon que précédemment des valeurs de la variable y) :

$$RMSE = \left[\frac{1}{n}\sum_{i=1}^{n} {d_i}^2\right]^{\frac{1}{2}} \qquad (2.4.2)$$

Ce test informe cette fois sur la performance à court terme de la formulation obtenue, car il compare point par point les écarts entre les $y_{i,calc}$ et les $y_{i,mes}$. L'ajustement est d'autant meilleur que la valeur de RMSE est faible. Cependant, si dans le tableau des valeurs de x_i, $y_{i,calc}$ et $y_{i,mes}$ il existe même un petit nombre de couples ($y_{i,calc}$; $y_{i,mes}$) qui présentent des écarts importants, alors la valeur de RMSE est élevée. De plus, en considérant simultanément les variables MBE et RMSE, on peut être en présence de l'un ou l'autre des quatre cas suivants :

1. des valeurs de MBE et RMSE simultanément très proches de zéro indiquent que la courbe des $y_{i,calc}$ ajuste quasi-parfaitement celle des $y_{i,mes}$;

2. le cas d'une très faible valeur de MBE et d'une valeur élevée de RMSE correspond à une forte dispersion des points (x_i ; $y_{i,mes}$) de part et d'autre de la courbe des $y_{i,calc}$;

3. le cas d'une valeur élevée (en valeur absolue) de MBE et d'une faible valeur de RMSE indique une faible surestimation (si MBE >0) ou une faible sous-estimation (si MBE < 0) $des\ y_{i,mes}\ par\ les\ y_{i,calc}$;

4. le cas où MBE et RMSE ont simultanément des valeurs élevées est par contre un indicateur d'un mauvais ajustement des $y_{i,mes}$ par la courbe des $y_{i,calc}$.

c) Variable « t » de Student

Même si les tests du MBE et du RMSE se complètent bien pour comparer raisonnablement une formulation théorique donnée avec les données expérimentales, ils ne montrent pas objectivement si oui ou non, du point de vue statistique, les résultats théoriques ($y_{i,calc}$) sont significativement différents de leurs contreparties expérimentales ($y_{i,mes}$).

D'où la nécessité d'un troisième test, qui est ici celui de la variable « t » de Student, dont l'une des relations de définition est la suivante :

$$t= \left[\frac{(n-1)(MBE)^2}{(RMSE)2-(MBE)^2}\right]^{½} \quad (2.4.3)$$

Pour une formulation donnée (relative au même échantillon que précédemment des valeurs de la variable y), on calcule « t » selon cette éq.(2.4.3). On lit ensuite dans des tables appropriées (ex. dans [63] [64]) la valeur critique $t_c\ (de\ t)$ qui correspond à un seuil de confiance α qu'on se fixe (ex .ici : $\alpha = 0{,}005$) et donc à un niveau de confiance $\gamma = 1-\alpha$ (ici : $\gamma = 0{,}995 = 99{,}5\ \%$) , ainsi qu'à un nombre de degrés de liberté (d.d.l) égal à n-1. Pour qu'on puisse affirmer que, à des erreurs statistiques près, la performance théorique obtenue conduit à des valeurs $y_{i,calc}$ qui sont peu différentes des données expérimentales $y_{i,mes}$ (de la variable « y » de l'échantillon considéré), il faut que l'inégalité ci-après soit satisfaite (ou vérifiée) :

$$t<t_c = t\ (n-1,\alpha) \quad (2.4.4)$$

II.2.3. Formulations réductibles à des régressions linéaires doubles

II.2.3.1. Formulations en question et leur réduction

Il s'agit ici des deux groupes de relations $\bar{k}_t\ vs\ \bar{s}_r\ et\ \bar{k}_t\ vs\ \overline{s'}_r$ en termes de polynômes du second degré (de la variable $\bar{s}_r\ ou\ \overline{s'}_r$). Dans la section 1.4.1, ces deux groupes de relations sont déjà réduites à des régressions linéaires doubles via les éqs. (1.6.5) et (1.6.5'). Pour la suite de la démarche, ces éqs. sont notées sous les formes ci-après :

$$y = a + b_1x_1 + b_2x_2 \quad (2.5.1)\ ; \quad y = a' + b'_1x'_1 + b'_2x'_2 \quad (2.5.1')$$

II.2.3.2. Expressions des coefficients à déterminer

On peut noter comme suit une relation de cause à effet où une variable dépendante « y » (l'effet) est supposée être linéairement corrélée à plusieurs variables indépendantes (les causes) x_i (« i » variant de 1 à k ; $k \geq 2$) :

$$y = a + b_1x_1 + b_2x_2 + \cdots + b_kx_k \quad (2.5.2)$$

Pour déterminer la régression linéaire multiple de ce type (éq. (2.5.2)) qui s'ajuste le mieux aux données expérimentales ($x_{ij}\ ; y_j$) d'un échantillon de « N » observations (« j » varie de 1 à N) de ces k+1 variables ($x_i\ et\ y$), on utilise la méthode des moindres carrés.

Celle-ci aboutit au système ci-après de k éqs. algébriques linéaires régulières (c-à-d indépendantes, compatibles et non homogènes) aux k inconnues ($b_1, b_2, \dots, b_k$)[15] [65] [66] :

$$m_{1y} = b_1 m_{11} + b_2 m_{21} + \cdots + b_k m_{k1} \quad (2.5.3.1)$$

$$m_{2y} = b_1 m_{12} + b_2 m_{22} + \cdots + b_k m_{k2} \quad (2.5.3.2)$$

.

.

.

$$m_{iy} = b_1 m_{1i} + b_2 m_{2i} + \cdots + b_k m_{ki} \quad (2.5.3.i)$$

.

.

.

$$m_{ky} = b_1 m_{1k} + b_2 m_{2k} + \cdots + b_k m_{kk} \quad (2.5.3.k)$$

Ces éqs. contiennent des moments simples ou liés d'ordre 2 pour les (k+1) variables ($x_i \; et \; y$) définis comme suit :

$$m_{iy} = m_{yi} = \frac{1}{N}\sum_{j=1}^{N}\left(x_{ij} - \bar{x}_i\right)\left(y_j - \bar{y}\right) \quad (2.5.4.1)$$

$$m_{ii} = \frac{1}{N}\sum_{j=1}^{N}\left(x_{ij} - \bar{x}_i\right)^2 \quad (2.5.4.2)$$

$$m_{ik} = m_{ki} = \frac{1}{N}\sum_{j=1}^{N}\left(x_{ij} - \bar{x}_i\right)\left(x_{kj} - \bar{x}_k\right) \quad (2.5.4.3)$$

$$m_{yy} = \frac{1}{N}\sum_{j=1}^{N}\left(y_{ij} - \bar{y}\right)^2 \quad (2.5.4.4)$$

Dans l'hypothèse où N≥k+1, il suffit de résoudre le système d'éqs. (2.5.3.1)-(2.5.3.k) pour obtenir des estimations des coefficients ($b_1, b_2, \dots, b_k$) de la régression linéaire multiple proposée. On en déduit ensuite l'expression suivante du coefficient « a » (en termes d'estimation) :

$$a = \bar{y} - (b_1\bar{x}_1 + b_2\bar{x}_2 + \cdots + b_k\bar{x}_k) \quad (2.5.5.1)$$

où

$\bar{x_i} = \frac{1}{N}\sum_{j=1}^{N} x_{ij}$ (2.5.5.2) ; $\bar{y} = \frac{1}{N}\sum_{j=1}^{N} y_j$ (2.5.5. 3)

Le coefficient de corrélation multiple (R) est donné par :

$$R_{y.12\ldots k} = \left[\frac{\sum_{i=1}^{k} b_i m_{iy}}{m_{yy}}\right]^{\frac{1}{2}} \quad (2.5.6)$$

Il mesure le degré d'ajustement (d'un échantillon de N observations des (k+1) variables x_i et y) par la relation linéaire multiple considérée. Il mesure aussi, sans se prononcer sur l'ordre de causalité, le degré de relation linéaire entre la variable dépendante y et l'ensemble des k variables x_i prises globalement. Quant à lui, le coefficient de détermination multiple (R^2) (donné par les membres de l'éq.(2.5.6) élevés au carré) est une mesure du pourcentage (ou de la proportion) de la variance (ou de la variabilité) de y, qui est expliqué (e) ou prédit (e) par les k variables x_i prises globalement. Plus R^2 tend vers 1, plus ce pourcentage est élevé et meilleur est l'ajustement.

Dans le cadre du présent travail, on considère k=2 dans les éqs. (2.5.2) à (2.5.6). De plus, on considère (et on utilise) des relations tout à fait similaires aux précédentes (éqs. précédentes) à partir de la corrélation linéaire multiple ci-après (à la place de celle de l'éq.(2.5.2)) :

$$y = a' + b'_1 x'_1 + b'_2 x'_2 + \cdots + b'_k x'_k \quad (2.5.7)$$

II.2.3.3. Autres variables statistiques à déterminer

Comme pour les groupes de formulations réductibles à des régressions linéaires simples (section II.2.2), ici aussi, en complément aux coefficients précédents de corrélation multiple et détermination multiple, on détermine également les mêmes autres variables statistiques (MBE, RMSE et t) que dans la sous- section II.2.2.3.

CHAPITRE III. RESULTATS OBTENUS ET LEUR INTERPRETATION

III.1. Exemples des calculs et résultats sur un site pour chacun des couples des formulations recherchées

Dans chacune des sections du présent sous-chapitre, le site de Musasa sera pris comme exemple (des calculs et résultats) et seule la première formulation (du couple de formulations) sera considérée, étant donné que le traitement de la deuxième formulation (du couple) est tout à fait semblable à celui de la première.

III.1.1. Relations linéaires simples $\bar{k}_t$ *vs* $\bar{s}_r$ et $\bar{k}_t$ *vs* $\bar{s'}_r$

Il s'agit du couple suivant d'éqs. de la sous-section 2.2.2.1 :

$$y = a_1 x + b_1 \qquad (2.2.1); \quad y = a'_1 x' + b'_1 \qquad (2.2.1')$$

Le traitement de l'éq. (2.2.1), basé sur la démarche décrite aux sous-sections II.2.2.2 et II.2.2.3, a été mené selon les calculs indiqués dans les tableaux 10 et 11 ci-dessous. Les résultats de ces calculs et leur interprétation directe sont proposés sous ces tableaux.

Tableau 10 : Données nécessaires au calcul des coefficients de la régression linéaire (2.2.1) et des coefficients R et R^2 pour le site de Musasa

Grandeur / Mois	x_i	y_i	$y_i - \bar{y}$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
Janvier	0,435	0,4606	-0,0558	-0,151	0,022801	0,00311364	0,0084258
Février	0,516	0,4556	-0,0608	-0,070	0,004900	0,00369664	0,0042560
Mars	0,495	0,5055	-0,0109	-0,091	0,008281	0,00011881	0,0009919
Avril	0,500	0,5077	-0,0087	-0,086	0,007396	0,00007569	0,0007482
Mai	0,586	0,5643	0,0479	-0,000	0,00000	0,00229441	0,000000
Juin	0,765	0,6101	0,0937	0,179	0,032041	0,00877969	0,0167723
Juillet	0,803	0,5866	0,0702	0,217	0,047089	0,00492804	0,0152334
Août	0,773	0,5652	0,0488	0,187	0,034969	0,00238144	0,0091256
Septembre	0,664	0,5248	0,0084	0,078	0,006084	0,00007056	0,0006552
Octobre	0,519	0,4833	-0,0331	-0,067	0,004489	0,00109561	0,0022177
Novembre	0,489	0,4571	-0,0593	-0,097	0,009409	0,00351649	0,0057521
Décembre	0,481	0,4762	-0,0402	-0,105	0,011025	0,00161604	0,0042210
Somme	**7,026**	**6,1970**			**0,1884840**	**0,03168706**	**0,0683992**
Moyenne	**0,586**	**0,5164**					

$\bar{x} =$ 0,586 ; $\overline{y} =$ 0,5164 ; $a = \frac{\sum_{i=1}^{12}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{12}(x_i-\bar{x})^2}$ = 0,362891279 ;

a≈ 0,3629 (3.1.1)

$b = \bar{y} - a\bar{x}$ = 0,3037406

b≈ 0,3037 (3.1.2)

La régression linéaire simple est donc ici : $y = 0{,}3629x + 0{,}3037$ (3.1.3)

Le coefficient de corrélation est :

$$R = \frac{\sum_{i=1}^{12}(x_i - \bar{x})\,(y_i - \bar{y})}{([\sum_{i=1}^{12}(x_i - \bar{x})^2]\ [\sum_{i=1}^{12}(y_i - \bar{y})^2])^{\frac{1}{2}}}$$

R=0,885060185 (3.1.4)

Le coefficient de détermination est: R^2=0,783331531 (3.1.5)

Ceci indique que 78,3% de la variance de $y = \bar{k}_t$ sont expliqués (ou prédits) par la variable indépendante $x = \bar{s}_r$ ici (c-à-d dans l'échantillon considéré des données expérimentales de $x = \bar{s}_r$ et $y = \bar{k}_t$ pour le site de Musasa et la période d'étude correspondante).

Tableau 11 : Données nécessaires au calcul des autres variables statistiques : relation (2.2.1), site de Musasa

Grandeur / Mois	x_i	$y_{i\,mes}$	$y_{i\,calc}$	d_i	$(d_i)^2$
Janvier	0,435	0,4606	0,4616	0,0010	0,00000100
Février	0,516	0,4556	0,4910	0,0354	0,00125316
Mars	0,495	0,5055	0,4833	-0,0222	0,00049284
Avril	0,500	0,5077	0,4852	-0,0225	0,00050625
Mai	0,586	0,5643	0,5164	-0,0479	0,00229441
Juin	0,765	0,6101	0,5813	-0,0288	0,00082944
Juillet	0,803	0,5866	0,5951	0,0085	0,00007225
Août	0,773	0,5652	0,5842	0,0190	0,00036100
Septembre	0,664	0,5248	0,5447	0,0199	0,00039601
Octobre	0,519	0,4833	0,4920	0,0087	0,00007569
Novembre	0,489	0,4571	0,4812	0,0241	0,00058081
Décembre	0,481	0,4762	0,4783	0,0021	0,00000441
Somme	**7,032**	**6,1970**	**6,1943**	**-0,0027**	**0.00686727**
Moyenne	**0,586**	**0,5164**	**0,5162**	**-0,000225**	**0,000572272**

MBE=-0,000225 (3.1.6) ; RMSE=0,023922217 (3.1.7)

t= $[\frac{(n-1)(MBE)^2}{(RMSE)^2-(MBE)^2}]^{½}$ où n=12;

t=0,031195837 (3.1.8)

Or, la lecture des tables sur les valeurs critiques de la variable « t » de Student (ex. dans [63] [64]), c-à-d $t_c = t$(n-1; α) (avec ici et pour tous les échantillons des données de x et y ou x' et y de tout le présent travail : $\alpha = 0{,}005$ et donc $\gamma = 1 - \alpha = 0{,}995 = 99{,}5\,\%$; n=12 ; n-1=11), donne :

$t_c = 3{,}106$ (3.1.9)

On conclut donc qu'ici, l'inégalité $t < t_c$ (2.4.4) est bien satisfaite.

En conclusion d'étape, la formulation (régression linéaire simple) obtenue (éq.(3.1.3)) ajuste très bien les données expérimentales de $x = \bar{s}_r$ et $y = \bar{k}_t$ du site de Musasa et de la période considérée (voir dans le tableau 2.1). En effet, les valeurs de R et R^2 sont assez élevées (p /r à leur maximum qui est 1), celles de MBE et RMSE sont très faibles (proches de zéro), puis t est très faible p/r à sa valeur critique t_c=3,106.

III.1.2. Fonctions logarithmiques $\bar{k}_t$ vs $\bar{s}_r$ *et* $\bar{k}_t$ vs $\bar{s'}_r$

Il s'agit du couple des éqs. suivantes de la section 1.4.1 :

$y = a_2 \ln x + b_2$ (1.6.2); $y = a'_2 \ln x' + b'_2$ (1.6.2′)

qui se réduisent aux régressions linéaires simples ci-après de la sous-section 2.2.2.1 :

$y = a_2 X + b_2$ (2.2.2); $y = a'_2 X' + b'_2$ (2.2.2')

Le traitement de l'éq. (2.2.2) est aussi basé sur la démarche décrite dans les sous-sections (II.2.2.2) et (II.2.2.3). Les calculs y relatifs sont consignés dans les tableaux 12 et 13 ci-dessous. Les résultats de ces calculs et des commentaires directs sur eux sont donnés sous ces tableaux.

Tableau 12 : Données nécessaires au calcul des coefficients de la régression linéaire (2.2.2) et des coefficients R et R^2 pour le site de Musasa

Grandeur / Mois	x_i	y_i	$X_i = \ln x_i$	$X_i - X$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
Janvier	0,435	0,4606	-0,8324	-0,2755	0,07590025	0,00311364	0,01537290
Février	0,516	0,4556	-0,5516	-0,1047	0,01096209	0,00369664	0,00636576
Mars	0,495	0,5055	-0,7032	-0,1463	0,02140369	0,00011881	0,00159467
Avril	0,500	0,5077	-0,6931	-0,1362	0,01855044	0,00007569	0,00118494
Mai	0,586	0,5643	-0,5344	0,0225	0,00050625	0,00229441	0,00107775
Juin	0,765	0,6101	-0,2679	0,2890	0,08352100	0,00877969	0,02707930
Juillet	0,803	0,5866	-0,2194	0,3375	0,11390625	0,00492804	0,023692550
Août	0,773	0,5652	-0,2575	0,2994	0,08964036	0,00238144	0,01461072
Septembre	0,664	0,5248	-0,4095	0,1474	0,02172676	0,00007056	0,00123816
Octobre	0,519	0,4833	-0,6559	-0,9900	0,00980100	0,00109561	0,00327690
Novembre	0,489	0,4571	-0,7154	-0,1585	0,02512225	0,00351649	0,00939905
Décembre	0,481	0,4762	-0,7319	-0,1750	0,03062500	0,00161604	0,00703500
Somme	**7,026**	**6,1970**	**-6,6822**		**0,050166534**	**0,03168706**	**0,11192765**
Moyenne	**0,586**	**0,5164**	**-0,5569**				

$\bar{x} = -0{,}5569$; $\bar{y} = 0{,}5164$; $a = \frac{\sum_{i=1}^{12}(X_i-\bar{X})(y_i-\bar{y})}{\sum_{i=1}^{12}(X_i-\bar{X})^2} = 0{,}223112184$

a≈ 0,2231 (3.2.1)

$b = \bar{y} - a\bar{X}$ = 0,64064439

b≈ 0,6406 (3.2.2)

La régression linéaire simple (2.2.2) est donc ici : $y = 0{,}2231\ X + 0{,}6406$ (3.2.3)

Le coefficient de corrélation est :

$$R = \frac{\sum_{i=1}^{12}(X_i-\bar{X})(y_i-\bar{y})}{([\sum_{i=1}^{12}(X_i-\bar{X})^2][\sum_{i=1}^{12}(y_i-\bar{y})^2])^{\frac{1}{2}}}$$

R=0,887747308 (3.2.4)

Le coefficient de détermination est : R^2=0,788095283 (3.2.5)

Cela indique que 78,8% de la variance de $y = \bar{k}_t$ sont expliqués (ou prédits) par la variable $X = \ln x = \ln \bar{s}_r$ ici.

Tableau 13 : Données nécessaires au calcul des autres variables statistiques ; relation (2.2.2), site de Musasa

Grandeur / Mois	X_i	$y_{i\,mes}$	$y_{i\,calc}$	d_i	$(d_i)^2$
Janvier	-0,8324	0.4606	0,4549	-0,0057	0,00003249
Février	-0,6616	0,4556	0,4930	0,0374	0,00139876
Mars	-0,7032	0,5055	0,4837	-0,0228	0,00051984
Avril	-0,6931	0,5077	0,4860	-0,0217	0,00047089
Mai	-0,5344	0,5643	0,5214	-0,0429	0,00184041
Juin	-0,2679	0,6101	0,5808	-0,0293	0,00085849
Juillet	-0,2194	0,5866	0,5917	0,0051	0,00002601
Août	-0,2575	0,5652	0,5832	0,0180	0,00032400
Septembre	-0,4095	0,5248	0,5492	0,0244	0,00059536
Octobre	-0,6559	0,4833	0,4943	0,0110	0,00012100
Novembre	-0,7154	0,4571	0,4810	0,0239	0,00057121
Décembre	-0,7319	0,4762	0,4773	0,0011	0,00000121
Somme				**-0,0015**	**0,00675967**
Moyenne				**-0,000125**	**0,000563305**

MBE=-0,000125 (3.2.6) ; RMSE=0,023734064 (3.2.7)

$t= \left[\frac{(n-1)(MBE)^2}{(RMSE)^2-(MBE)^2}\right]^{½}$;

t=0,017467882 (3.2.8)

La lecture des tables citées plus haut (réfs.[63] [64]) sur les valeurs critiques t_c de la variable « t » de Student donne (pour le même nombre de d.d.l (n-1=11) et le même niveau de confiance $\gamma = 99{,}5\%$) : $t_c = 3{,}106$. Par conséquent, l'inégalité $t < t_c$ (2.4.4) est aussi satisfaite ici.

En conclusion d'étape, la régression linéaire simple (éq(3.2.3)) obtenue ici, c-à-d la fonction logarithmique correspondante (si on remplace X par $\ln x$), ajuste très bien les données expérimentales de $x = \bar{s}_r$ et $y = \bar{k}_t$ du site de Musasa et de la période d'étude.

En effet, les valeurs de R et R^2 sont assez élevées (p/r à leur maximum qui est « 1 »), celles de MBE et RMSE sont très faibles (proches de zéro), puis t est très inférieure à sa valeur critique t_c=3,106.

III.1.3. Fonctions exponentielles $\bar{k}_t\ vs\ \bar{s}_r$ et $\bar{k}_t\ vs\ \bar{s}'_r$

Sont concernées ici le couple d'éqs. ci-après de la section 1.4.1 :

$y = b_3 e^{a_3 x}$ (1.6.3) ; $y = b'_3 e^{a'_3 x'}$ (1.6.3′)

dont la réduction à des régressions linéaires simples donne le couple d'éqs. suivantes de la sous- section II.2.2.1 :

$Y = ax + b$ (2.2.3) ; $Y = a'x' + b'$ (2.2.3')

avec : a=a_3 , $b = \ln b_3$, a′=a'_3, $b' = \ln b'_3$, $Y = \ln y$, $x = \bar{s}_r$, $x' = \overline{s'}_r$ et $y = \bar{k}_t$.

Sur base de la démarche décrite dans les sous-sections (II.2.2.2) et (II.2.2.3), l'éq. (2.2.3) a été traitée et les calculs y relatifs sont présentés dans les tableaux 14 et 15 ci-dessous. Les résultats de ces calculs et leur interprétation sommaire sont donnés en dessous de ces tableaux.

Tableau 14 : Données nécessaires au calcul des coefficients de la régression linéaire (2.2.3) et des coefficients R et R^2 pour le site de Musasa

Grandeur / Mois	Y_i	$Y_i - \bar{Y}$	$(Y_i - \bar{Y})^2$	x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})\,(Y_i - \bar{Y})$
Janvier	-0,7752	-0,1092	0,01199025	0,435	-0,151	0,022801	0,0165345
Février	-0,7861	-0,1204	0,01449616	0,516	-0,070	0,004900	0,0084280
Mars	-0,6822	-0,0165	0,00027225	0,495	-0,091	0,008281	0,0015015
Avril	-0,6779	-0,0122	0,00014884	0,500	-0,096	0,007396	0,0010492
Mai	-0,5722	0,0935	0,00874225	0,586	-0,000	0,00000	0,000000
Juin	-0,4941	0.1716	0,02944656	0,765	0,179	0,032041	0,0307164
Juillet	-0,5334	0,1323	0,01750329	0,803	0,217	0,047089	0,0287091
Août	-0,5706	0,0951	0,00904401	0,773	0,187	0,034969	0,0177837
Septembre	-0,6447	0,0210	0,00044100	0,664	0,078	0,006084	0,0016380
Octobre	-0,7271	-0,0614	0,00376996	0,519	-0,067	0,004489	0,0041138
Novembre	-0,7829	-0,1172	0,01373584	0,489	-0,097	0,009409	0,0113684
Décembre	-0,7419	-0,0762	0,00580644	0,481	-0,105	0,011025	0,0080010
Somme	**-7,2883**	-	**0,11539685**	**7,026**	-	**0,1884840**	**0,1298436**
Moyenne	**-0,6657**	-	-	**0,586**	-	-	-

$\bar{x} = 0{,}586$; $Y = -0{,}6657$; $a = \frac{\sum_{i=1}^{12}(x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^{12}(x_i - \bar{x})^2}$ = 0,688883937;

a≈ 0,6889 (3.3.1)

$b = \bar{Y} - a\bar{x}$ = -1,0693954 ;

b≈ $-1{,}0694$ (3.3.2)

La régression linéaire simple (2.2.3) se note donc ici :

$Y = 0{,}3629x - 1{,}0694$ (3.3.3)

Dans l'éq (1.6.3), $a_3 = a$ et $b_3 = e^b \approx 0{,}3608$ (3.3.4)

Le coefficient de corrélation entre x et Y est ici:

$$R = \frac{\sum_{i=1}^{12}(x_i - \bar{x})(Y_i - \bar{Y})}{([\sum_{i=1}^{12}(x_i - \bar{x})^2][\sum_{i=1}^{12}(Y_i - \bar{Y})^2])^{\frac{1}{2}}} \quad ; \qquad R=0{,}880412755 \qquad (3.3.5)$$

Le coefficient de détermination a quant à lui comme valeur :

R^2=0,77512662 (3.3.6)

Cette valeur indique que 77,5% de la variance de la variable Y sont expliqués (ou prédits) par la variable x (pour les données de l'échantillon des variables $x = \bar{s}_r$ et $y = \bar{k}_t$, avec $Y = \ln y$, qui se rapportent au site de Musasa et à la période d'étude considérée dans le tableau 2.1).

Tableau 15 : Données nécessaires au calcul d'autres variables statistiques ; relation (2.2. 3), site de Musasa

Grandeur / Mois	x_i	$Y_{i\,mes}$	$Y_{i\,calc}$	d_i	$(d_i)^2$
Janvier	0,435	-0,7752	-0,7697	0,0058	0,00003025
Février	0,516	-0,7861	-0,7139	0,0722	0,00521284
Mars	0,495	-0,6822	-0,7284	-0,0462	0,00213444
Avril	0,500	-0,6779	-0,7250	-0,0471	0,00221841
Mai	0,586	-0,5722	-0,6657	-0,0935	0,00874225
Juin	0,765	-0,4941	-0,5424	-0,04935	0,00233289
Juillet	0,803	-0,5334	-0,5162	0,0172	0,00029584
Août	0,773	-0,5706	-0,5361	0,0345	0,00119025
Septembre	0,664	-0,6447	-0,6120	0,0327	0,00106929
Octobre	0,519	-0,7271	-0,7119	0,0152	0,00023104
Novembre	0,489	-0,7829	-0,7325	0,0504	0,00254016
Décembre	0,481	-0,7419	-0,7380	0,0039	0,00001521
Somme	**7,026**			**-0,0035**	**0,02601287**
Moyenne	**0,586**			**-0,000291666**	**0,002167739**

MBE=-0,000291666 (3.3.7) ; RMSE=0,046558985 (3.3.8)

$t = [\frac{(n-1)(MBE)^2}{(RMSE)^2-(MBE)^2}]^{½}$;

t=0,020777206 (3.3.9)

Or, selon les tables sur les valeurs critiques « t_c» de la variable « t » de Student, pour un nombre de ddl n-1=11 et un niveau de confiance $\gamma = 1 - \alpha = 99{,}5\%$, on lit t_c=3,106. Donc ici, l'inégalité t<t_c (2.4.4) est très bien satisfaite. Par ailleurs, les valeurs obtenues de R et R^2 sont ici assez élevées (p/r à leur maximum qui vaut « 1»), celles de MBE et RMSE sont très proches de zéro, puis t est très faible devant t_c. On conclut par conséquent que la régression linéaire (3.3.3) obtenue ici, c-à-d aussi la fonction exponentielle (1.6.3) (dans laquelle a=$a_3 \approx 0{,}6889$, $b_3 = e^b \approx 0{,}3608$, $x = \bar{s}_r \; et \; y = \bar{k}_t$) ajuste très bien les données expérimentales de x et y du site de Musasa et de la période correspondante d'étude (voir tableau 1).

III.1.4. Fonctions puissances $\bar{k}_t$ vs $\bar{s}_r$ et $\bar{k}_t$ vs $\bar{s'}_r$

Il s'agit ici du couple d'éqs ci-après de la section I.4.1 :

$y = b_4 (x)^{a_4}$ (1.6.4) ; $y = b'_4 \, x'^{a'_4}$ (1.6.4′)

qui se réduisent aux régressions linéaires simples suivantes de la sous-section II.2.2.1 :

$Y = aX + b$ (2.2.4) ; $Y = a'X' + b'$ (2.2.4')

où

a=a_4, $b = \ln b_4$, a′=a'_4, $b' = \ln b'_4$, $X = \ln x$, $Y = \ln y$, $x = \bar{s}_r, x' = \bar{s'}_r \; et \; y = \bar{k}_t$.

La relation (2.2.4) a été traitée selon la démarche décrite dans les sous-sections II.2.2.2 et II.2.2.3 ; les calculs de ce traitement sont consignés dans les tableaux 16 et 17 ci-dessous. Les résultats de ces calculs et leur interprétation immédiate sont proposés en dessous de ces tableaux.

Tableau 16 : Données nécessaires au calcul des coefficients de la régression linéaire (2.2.4) et des coefficients R et R^2 pour le site de Musasa

Grandeur / Mois	Y_i	$Y_i - \bar{Y}$	$(Y_i - \bar{Y})^2$	X_i	$X - \bar{X}$	$(X_i - \bar{X})^2$	$(x_i - \bar{x})(Y_i - \bar{Y})$
Janvier	-0.7752	-0,1092	0,01199025	-0,8324	-0,2755	0.07590025	0,03016725
Février	-0,7861	-0,1204	0,01449616	-0,5516	-0,1047	0,01096209	0,01260588
Mars	-0,6822	-0,0165	0,00027225	-0,7032	-0,1463	0,02140369	0,00241395
Avril	-0,6779	-0,0122	0,00014884	-0,6931	-0,1362	0,01855044	0,00166164
Mai	-0,5722	0,0935	0,00874225	-0,5344	0,0225	0,02140369	0,00210375
Juin	-0,4941	0.1716	0,02944656	-0,2679	0,2890	0,08352100	0,04959240
Juillet	-0,5334	0,1323	0,01750329	-0,2194	0,3375	0,11390625	0,04465125
Août	-0,5706	0,0951	0,00904401	-0,2575	0,2994	0,08964036	0,02847294
Septembre	-0,6447	0,0210	0,00044100	-0,4095	0,1474	0,02172676	0,00309540
Octobre	-0,7271	-0,0614	0,00376996	-0,6559	-0,9900	0,00980100	0,00607860
Novembre	-0,7829	-0,1172	0,01373584	-0,7154	-0,1585	0,02512225	0,01857620
Décembre	-0,7419	-0,0762	0,00580644	-0,7319	-0,1750	0,03062500	0,01333500
Somme	**-7,9883**		**0,11539685**	**-6,6822**		**0,50166534**	**0,21275426**
Moyenne	**-0,6657**			**-0,5569**		**0,041805445**	

$\bar{X} = -0{,}5569$; $\bar{Y} = -0{,}6657$; $a = \frac{\sum_{i=1}^{12}(x_i-\bar{x})(Y_i-\bar{Y})}{\sum_{i=1}^{12}(x_i-\bar{x})^2}$ = 0,424095992 ;

$a \approx 0{,}4241$ (3.4.1)

$b = \bar{Y} - a\bar{X}$ = -0,42952094

b≈ $-0{,}4295$ (3.4.2)

La régression linéaire simple (2.2.4) se note donc ici :

$Y = 0{,}4241\, X - 0{,}4295$ (3.4.3)

Dans l'éq. (1.6.4), $a_4 = a$ et $b_4 = e^b \approx 0{,}6508$ (3.4.4)

Le coefficient de corrélation entre X et Y est ici:

$R = \frac{\sum_{i=1}^{12}(X_i-\bar{X})(Y_i-\bar{Y})}{([\sum_{i=1}^{12}(X_i-\bar{X})^2][\sum_{i=1}^{12}(Y_i-\bar{Y})^2])^{\frac{1}{2}}}$; R=0,884248312 (3.4.5)

Le coefficient de détermination est donc ici:

R^2=0,781895077 (3.4.6)

Cela indique que 78,2 % de la variance de la variable Y sont prédits (ou expliqués) par la variable x (pour les données de l'échantillon des variables $x = \bar{s}_r$ et $y = \overline{k_t}$, avec $X = \ln x$ et $Y = \ln y$, qui se rapportent au site de Musasa et à la période d'étude correspondante montrée dans le tableau 1).

Tableau 17 : Données nécessaires au calcul d'autres variables statistiques ; relation (2.2.4), site de Musasa

Grandeur / Mois	X_i	$Y_{i\,mes}$	$Y_{i\,calc}$	d_i	$(d_i)^2$
Janvier	-0,8324	-0,7725	-0,7825	-0,0073	0,00005329
Février	-0,6616	-0,7861	-0,7101	0,0760	0,00577600
Mars	-0,7032	-0,6822	-0,7277	-0,0455	0,00207025
Avril	-0,6931	-0,6779	-0,7234	-0,0455	0,00207025
Mai	-0,5344	-0,5722	-0,6561	-0,0839	0,00703921
Juin	-0,2679	-0,4941	-0,5431	-0,0490	0,00240100
Juillet	-0,2194	-0,5334	-0,5225	0,0109	0,00011881
Août	-0,2575	-0,5706	-0,5387	0,0319	0,00101761
Septembre	-0,4095	-0,6447	-0,6032	0,0415	0,00172225
Octobre	-0,6559	-0,7271	-0,7077	0,0194	0,00037636
Novembre	-0,7154	-0,7829	-0,7329	0,0500	0,00250000
Décembre	-0,7319	-0,7419	-0,7399	0,0020	0,00000400
Somme				**0,0005**	**0,02514903**
Moyenne				**0,000041666**	**0,002095752**

MBE = 0,000041666 (3.4.7)

RMSE = 0,045779389 (3.4.8)

$t = \left[\frac{(n-1)(MBE)^2}{(RMSE)^2-(MBE)^2}\right]^{½}$ où n=12 ;

t=0,003018619 (3.4.9)

Or, la lecture des tables des valeurs critiques « t_c » de la variable « t » de Student, pour (ici) un nombre de ddl n-1=11 et un niveau de confiance $\gamma = 1 - \alpha = 99{,}5$ %, donne comme résultat t_c=3,106. Donc, l'inégalité t<t_c (2.4.4) est ici très bien satisfaite.

En conclusion d'étape, vu que les valeurs obtenues ici de R et R^2 ne sont pas très éloignées de leur maximum « 1 » ; que celles de MBE et RMSE sont très faibles (proches de zéro) et que t est très inférieure à t_c, on peut affirmer que la régression linéaire simple (3.4.3) (et par conséquent la fonction puissance (1.6.4) dans laquelle a= a_4, $b_4 = e^b$ des $éqs.$ (3.4.1) et (3.4.4), $x = \bar{s}_r$ et $y = \bar{k}_t$) ajuste très bien les données expérimentales de x et y du site et de la période considérés ici.

III.1.5. Polynômes du second degré $\bar{k}_t$ vs $\bar{s}_r$ et $\bar{k}_t$ vs $\bar{s'}_r$

Les deux types de polynômes concernés ont leurs formes réduites à des régressions linéaires doubles données par les éqs. (1.6.5) et (1.6.5') dans la section (I.4.1), ou alors par le couple d'éqs. suivantes de la sous-section (II.2.3.1) :

$$y = a + b_1 x_1 + b_2 x_2 \qquad (2.5.1)\,; \quad y = a' + b'_1 x'_1 + b'_2 x'_2 \qquad (2.5.1')$$

où: $x_1 = x = \bar{s}_r$, $x_2 = x^2 = (\bar{s}_r)^2$, $y = \bar{k}_t$, $x'_1 = x' = \bar{s'}_r$ et $x'_2 = x'^2 = (s'_r)^2$.

La relation (2.5.1) a été traitée selon la démarche décrite dans les sous-sections II.2.3.2 et II.2.3.3. Les calculs y relatifs sont donnés dans les tableaux 18 et 19 ci-dessous. Les résultats de ces calculs et leur interprétation immédiate sont proposés en dessous de ces tableaux.

Tableau 18 : Données nécessaires au calcul des coefficients de la régression linéaire double (2.5.1) et des coefficients $\mathbf{R = R_{y.12}}$ et $\mathbf{R^2}$

Grandeur / Mois	x_i	y_i	$y_i - \bar{y}$	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
Janvier	0,435	0,4606	-0,0558	-0,151	0,022801	0,00311364	0,0084258
Février	0,516	0,4556	-0,0608	-0,070	0,004900	0,00369664	0,0042560
Mars	0,495	0,5055	-0,0109	-0,091	0,008281	0,00011881	0,0009919
Avril	0,500	0,5077	-0,0087	-0,096	0,007396	0,00007569	0,0007482
Mai	0,586	0,5643	0,0479	-0,000	0,00000	0,00229441	0,000000
Juin	0,765	0,6101	0,0937	0,179	0,032041	0,00877969	0,0167723
Juillet	0,803	0,5866	0,0702	0,217	0,047089	0,00492804	0,0152334
Août	0,773	0,5652	0,0488	0,187	0,034969	0,00238144	0,0091256
Septembre	0,664	0,5248	0,0084	0,078	0,006084	0,00007056	0,0006552
Octobre	0,519	0,4833	-0,0331	-0,067	0,004489	0,00109561	0,0022177
Novembre	0,489	0,4571	-0,0593	-0,097	0,009409	0,00351649	0,0057521
Décembre	0,481	0,4762	-0,0402	-0,105	0,011025	0,00161604	0,0042210
Somme	**7,026**	**6,1970**			**0,1884840**	**0,03168706**	**0,0683992**
Moyenne	**0,586**	**0,5164**			**0,015707**	**0,0026140588**	**0,0056999**

Suite du tableau 18

grandeur / Mois	$x_i{}^2$	$x_i{}^2 - \overline{x^2}$	$(x_i{}^2 - \overline{x^2})^2$	$(x_i - \bar{x})(x_i{}^2 - \overline{x^2})$	$(x_i{}^2 - \overline{x^2})\ (y_i - \bar{y})$
Janvier	0,189225	-0,169292	0,028659781	0,025563092	0,009446493
Février	0,266256	-0,092261	0,008512092	0,00645827	0,005609468
Mars	0,245025	-0,113492	0,012880434	0,010327772	0,001237062
Avril	0,250000	-0,108517	0,011775939	0,009332462	0,000944097
Mai	0,343396	-0,015121	0,000228644	-0,0000000	-0,000724295
Juin	0,585225	0,226708	0,051396517	0,040580732	0,021242539
Juillet	0.644809	0,286292	0,081963109	0,062125364	0,020097698
Août	0,597529	0,239012	0,057126736	0,044695244	0,011663785
Septembre	0,440896	0,082379	0,006786299	0,006425562	0,0000691983
Octobre	0,269361	-0,089156	0,007948792	0,005973452	0,002951063
Novembre	0,239121	-0,119396	0,014255404	0,011581412	0,007080182
Décembre	0,231361	-0,127156	0,016168648	0,01335138	0,005111671
Somme	**4,302204**		**0,297702399**	**0,236414742**	**0,085351752**
Moyenne	**0,358517**		**0,024808533**	**0,019701228**	**0,007112646**

$\bar{y} = 0{,}5164$ (3.5.1) ; $\bar{x}_1 = 0{,}586$ (3.5.2) ; $\bar{x}_2 = 0{,}358517$ (3.5.3)

$m_{yy} = 0{,}002640588$ (3.5.4) ; $m_{11} = 0{,}0157070$ (3.5.5)

$m_{1y} = 0{,}005699933$ (3.5.6) ; $m_{12} = m_{21} = 0{,}019701228$ (3.5.7)

$m_{22} = 0{,}024808533$ (3.5.8) ; $m_{2y} = 0{,}007112646$ (3.5.9)

Les coefficients $b_1\ et\ b_2$ (de la régression (2.5.1)) sont déterminés via la résolution (par la méthode de Cramer par ex.) du système ci-après d'éqs. algébriques régulières :

$$\begin{cases} m_{11}b_1 + m_{21}b_2 = m_{1y} \\ m_{12}b_1 + m_{22}b_2 = m_{2y} \end{cases}$$

$\Delta = m_{11}m_{22} - (m_{12})^2$; $\Delta = 0{,}000001529$ (3.5.10.1)

$\Delta_1 = m_{1y}m_{22} - m_{2y}m_{12}$; $\Delta_1 = 0{,}000001279$ (3.5.10.2)

$\Delta_2 = m_{11}m_{2y} - m_{21}m_{1y}$; $\Delta_2 = -0{,}000000577$ (3.5.10.3)

$b_1 = \frac{\Delta_1}{\Delta} = 0{,}836569914$; $b_1 \approx 0{,}8366$ (3.5.11)

$b_2 = \frac{\Delta_2}{\Delta} = -0{,}377599015$; $b_2 \approx -0{,}3776$ (3.5.12)

Le coefficient « a » (de la régression (2.5.1)) vaut donc ici :

$a = \bar{y} - (b_1\bar{x}_1 + b_2\bar{x}_2) = 0{,}161528419$; $a \approx 0{,}1615$ (3.5.13)

La régression linéaire double recherchée est donc ici :

$y = 0{,}1615 + 0{,}8366\,x_1 - 0{,}3776\,x_2$ (3.5.14)

où x_1, x_2 et y ont les significations écrites sous l'éq. (2.5.1').

Le coefficient de corrélation double a pour valeur :

$R_{y.12} = \left(\frac{b_1m_{1y} + b_2m_{2y}}{m_{yy}}\right)^{1/2}$; $R_{y.12} = 0{,}88809478$ (3.5.15)

Quant à lui, le coefficient de détermination double est :

$R^2{}_{y.12} = 0{,}788712338$ (3.5.16)

Cette valeur indique que 78,9 % de la variance de y sont prédits (ou expliqués) par les deux variables $x_1\ et\ x_2$ prises globalement (pour l'échantillon des variables $x = x_1 = \bar{s}_r$; $x_2 = x^2 = (\bar{s}_r)^2$ et $y = \bar{k}_t$ relatives au site de Musasa et à la période d'observation stipulée dans le tableau 1).

Tableau 19 : Données nécessaires au calcul d'autres variables statistiques ; relation 2.2.4, site de Musasa

Grandeur / Mois	x_i	$y_{i\,mes}$	$y_{i\,calc}$	d_i	$(d_i)^2$
Janvier	0,435	0,4606	0,4540	-0,0066	0,00004356
Février	0,516	0,4556	0,4931	0,0375	0,00140625
Mars	0,495	0,5055	0,4831	-0,0224	0,00050176
Avril	0,500	0,5077	0,4854	-0,0223	0,00049720
Mai	0,586	0,5643	0,5221	-0,0422	0,00178084
Juin	0,765	0,6101	0,5805	-0,0296	0,00087616
Juillet	0,803	0,5866	0,5898	0,0032	0,00001024
Août	0,773	0,5652	0,5826	0,0174	0,00030276
Septembre	0,664	0,5248	0,5505	0,0257	0,00066049
Octobre	0,519	0,4833	0,4940	0,0107	0,00011449
Novembre	0,489	0,4571	0,4803	0,0232	0,00053824
Décembre	0,481	0,4762	0,4765	0,0003	0,00000009
Somme	**7,026**	**6,1970**		**-0,0051**	**0,00673217**
Moyenne	**0,586**	**0,5164**		**-0,000425**	**0,000561014**

MBE = -0,000425 (3.5.17) ; RMSE=0,023685737 (3.5.18)

$t= [\frac{(n-1)(MBE)^2}{(RMSE)^2-(MBE)^2}]^{½}$; t=0,059520734 (3.5.19)

Etant donné que, dans les tables sur les valeurs critiques « t_c » de la variable « t » de Student, on lit t_c=3,106 pour un nombre de ddl n-1=11 et un niveau de confiance $\gamma = 1 - \alpha =$ 99,5 %, l'inégalité t<t_c (2.4.4) est satisfaite ici. Or en plus, les valeurs obtenues de MBE et RSME sont très faibles (proches de zéro), celles de R=$R_{y.12}$ et R^2 sont assez élevées (p/r à leur maximum qui est « 1 »), puis t est très inférieure à t_c. On conclut donc que la régression linéaire double (3.5.14) obtenue (et donc le polynôme du second degré correspondant :

$y = a + b_1 x + b_2 x^2, avec\ x = \bar{s}_r,\ x^2 = (\bar{s}_r)^2\ et\ y = \bar{k}_t$, $puis\ a = 0{,}1615,\ b_1 = 0{,}8366\ et\ b_2 = -0{,}3776$) ajuste très bien les données expérimentales de x et y du site et de la période considérés ici.

III.2. Résultats des calculs sur tous les sites

Des calculs similaires à ceux présentés dans les sous-sections III.1.1 à III.1.5, ont été effectués sur chacun des cinq couples (2.2.1)-(2.2.1'), (2.2.2)-(2.2.2'), (2.2.3)-(2.2.3'), (2.2.4)-(2.2.4') et (2.5.1)-(2.5.1') de formulations, non seulement pour Musasa, mais aussi pour les sites restants (du tableau 1). Le tableau 20 ci-dessous présente, pour chacun de ces six sites, une synthèse des résultats de la détermination de tous les cinq couples de (c-à-d toutes les dix) formulations, ainsi que des cinq variables statistiques complémentaires utilisées pour tester la performance de chacune de ces formulations.

Tableau 20 : Synthèse, pour les six sites, des formulations obtenues et des valeurs correspondantes des variables statistiques du test des performances de ces formulations

Formulations et variables statistiques	Bujumbra-Aéroport	Gitega-Zege	Imbo-SEMS
1.1) $\bar{k}_t = a_1\bar{s}_r + b_1$.	$\bar{k}_t = 0{,}3086\,\bar{s}_r + 0{,}3336$	$\bar{k}_t = 0{,}2881\,\bar{s}_r + 0{,}2951$	$\bar{k}_t = 0{,}3494\,\bar{s}_r + 0{,}2850$
R	0,8841	0,8975	0,8776
R^2.	0,7816	0,8055	0,7702
MBE	-0,000008	0,000083	-0,000108
RMSE	0,017873	0,017850	0,018582
t	0,001546	0,015484	0,019337
1.2) $\bar{k}_t = a'_1\bar{s}'_r + b'_1$.	$\bar{k}_t = 0{,}2862\,\bar{s}'_r + 0{,}3361$	$\bar{k}_t = 0{,}2701\,\bar{s}'_r + 0.2957$	$\bar{k}_t = 0{,}3321\,\bar{s}'_r + 0{,}2832$
R	0,8812	0,8993	0,8802
R^2.	0,7765	0,8088	0,7747
MBE	-0,000017	-0,000008	0,000083
RMSE	0,018076	0,017679	0,018399
t	0,003058	0,001563	0,015022

2.1) $\bar{k}_t = a_2 \ln \bar{s_r} + b_2$.	$\bar{k}_t = 0{,}1855 \ln \bar{s_r} + 0{,}6170$	$\bar{k}_t = 0{,}1600 \ln \bar{s_r} + 0{,}5537$	$\bar{k}_t = 0{,}1867\ ln\ \bar{s}_r + 0{,}5889$
R	0,9008	0,8921	0,8751
R^2.	0,8114	0,7959	0,7657
MBE	0,000008	-0,000025	0,000083
RMSE	0,016884	0,018385	0,018774
t	0,001637	0,004510	0,014722
2.2) $\bar{k}_t = a'_2 \ln \overline{s'_r} + b'_2$.	$\bar{k}_t = 0{,}1831 \ln \overline{s'_r} + 0{,}6045$	$\bar{k}_t = 0{,}1563\ \ln \overline{s'_r} + 0{,}5419$	$\bar{k}_t = 0{,}1847\ \ln \overline{s'_r} + 0{,}5794$
R	0,8948	0,8766	0,8782
R^2.	0,8007	0,7684	0,7713
MBE	0,000033	-0,000017	0,000042
RMSE	0,017069	0,018166	0,018559
t	0,006477	0,003043	0,007446
3.1) $\bar{k}_t = b_3 \exp(a_3 \bar{s}_r)$.	$\bar{k}_t = 0{,}3604 \exp(0{,}6020\ \bar{s}_r)$	$\bar{k}_t = 0{,}3203 \exp(0{,}6239\ \bar{s}_r)$	$\bar{k}_t = 0{,}3156 \exp(0{,}7452\ \bar{s}_r)$
R	0,8805	0,8939	0,8806
R^2.	0,7752	0,7990	0,7754
MBE	-0,000192	0,000092	-0,000350
RMSE	0,035512	0,039428	0,039593

t	0,017901	0,007711	0,029320
3.2) $\bar{k}_t = b'_3 \exp(a'_3 \bar{s'}_r)$.	$\bar{k}_t = 0{,}3623 \exp(0{,}5582\ \bar{s'}_r)$	$\bar{k}_t = 0{,}3208 \exp(0{,}5850\ \bar{s'}_r)$	$\bar{k}_t = 0{,}3150 \exp(0{,}7051\ \bar{s'}_r)$
R	0,8775	0,8956	0,8790
R^2.	0,7700	0,8021	0,7727
MBE	-0,000108	-0,000250	0,000233
RMSE	0,035915	0,038616	0,039305
t	0,010004	0,021472	0,019690
4.1) $\bar{k}_t = b_4 (\bar{s}_r)^{a_4}$	$\bar{k}_t = 0{,}6267\ (\bar{s}_r)^{0{,}3623}$	$\bar{k}_t = 0{,}5113\ (\bar{s'}_r)^{0{,}3474}$	$\bar{k}_t = 0{,}6036\ (\bar{s'}_r)^{0{,}3879}$
R	0.8981	0.8904	0.8789
R^2.	0,8066	0,7928	0,7725
MBE	0,000083	-0,000075	-0,000017
RMSE	0,033112	0,040043	0,039778
t	0,008347	0,006212	0,001390

4.2) $\overline{k}_t = b'_3(\overline{s'}_r)^{a'_4}$	$\overline{k}_t = 0{,}6116\,(\overline{s'}_r)^{0{,}3575}$	$\overline{k}_t = 0{,}5492\,(\overline{s'}_r)^{0{,}3458}$	$\overline{k}_t = 0{,}5910\,(\overline{s'}_r)^{0{,}3927}$
R	0.8919	0.8919	0.8785
R^2.	0,7956	0,7955	0,7718
MBE	0,000008	-0,000025	0,000025
RMSE	0,033864	0,039764	0,039393
t	0,001216	0,002085	0,002105
5.1) $\overline{k}_t = a_5 + b_5(\bar{s}_r) + c_5\,(s_r)^2$	$\overline{k}_t = -0{,}0061 + 1{,}4729(\bar{s}_r) - 0{,}9634(s_r)^2$	$\overline{k}_t = 0{,}3306 + 0{,}1574(\bar{s}_r) + 0{,}1131(s_r)^2$	$\overline{k}_t = 0{,}2987 + 0{,}2959(\bar{s}_r) + 0{,}0504(s_r)^2$
$R_{y,12}$	0.9133	0.8978	0.8779
${R_{y,12}}^2$	0,8342	0,8061	0,7707
MBE	-0,000425	-0,000017	-0,000367
RMSE	0,015620	0,017778	0,018346
t	0,090334	0,003109	0,066301
5.2) $\overline{k}_t = a'_5 + b'_5(\overline{s'}_r) + c'_5\,(s'_r)^2$	$\overline{k}_t = -0{,}0135 + 1{,}4130\,(\overline{s'}_r) - 0{,}8755\,(s'_r)^2$	$\overline{k}_t = 0{,}3342 + 0{,}1367(\overline{s'}_r) - 0{,}1089(s'_r)^2$	$\overline{k}_t = 0{,}2848 + 0{,}3263\,(\overline{s'}_r) - 0{,}0052\,(s'_r)^2$
$R_{y,12}$	0,9122	0.9002	0.8803
${R_{y,12}}^2$	0,8321	0,8103	0,7750

MBE RMSE	0,000183 0,015685	0,000033 0,017631	0,000175 0,018430
t Formulations et variables statistiques	Musasa 0,038769	Ruvyironza 0,006270	Rwegura 0,031494
1.1) $\bar{k}_t = a_1 \bar{s}_r + b_1$.	$\bar{k}_t = 0{,}3629\ \bar{s}_r + 0{,}3037$	$\bar{k}_t = 0{,}4163\ \bar{s}_r + 0{,}2285$	$\bar{k}_t = 0{,}3416\ \bar{s}_r + 0{,}2462$
R R^2.	0,8851 0,7833	0,9407 0,8848	0,9483 0,8993

MBE	-0,000225	-0,000175	-0,000108
RMSE	0,023922	0,017199	0,012113
t	0,031196	0,033748	0,029664
1.2) $\bar{k}_t = a'_1 \bar{s'}_r + b'_1$	$\bar{k}_t = 0{,}3406\ \bar{s'}_r + 0{,}3045$	$\bar{k}_t = 0{,}3901\ \bar{s'}_r + 0{,}2297$	$\bar{k}_t = 0{,}3207\ \bar{s'}_r + 0{,}2466$
R	0,8871	0,9414	0,9491
R^2.	0,7869	0,8862	0,9008
MBE	-0,000125	0,000033	-0,000192
RMSE	0,023723	0,017097	0,011874
t	0,017476	0,006466	0,053544
2.1) $\bar{k}_t = a_2\ ln\bar{s}_r + b_2$.	$\bar{k}_t = 0{,}2231\ ln\bar{s}_r + 0{,}6406$	$\bar{k}_t = 0{,}2047\ ln\bar{s}_r + 0{,}5845$	$\bar{k}_t = 0{,}1577\ ln\bar{s}_r + 0{,}5299$
R	0,8877	0,9411	0,9486
R^2	0,7881	0,8857	0,8999
MBE	-0,000125	0,000058	-0,000050
RMSE	0,023734	0,017133	0,012786
t	0,017468	0,011292	0,012970

2.2) $\bar{k}_t = a'_2\, ln\bar{s'}_r + b'_2$.	$\bar{k}_t = 0{,}2226\, ln\bar{s'}_r + 0{,}6270$	$\bar{k}_t = 0{,}2041\, ln\bar{s'}_r + 0{,}5718$	$\bar{k}_t = 0{,}1573\, ln\bar{s'}_r + 0{,}5202$
R	0,8897	0,9424	0,9417
R^2.	0,7916	0,8880	0,8868
MBE	0,000017	-0,00030	-0,000042
RMSE	0,023458	0,017286	0,012668
t	0,002356	0,057570	0,010908
3.1) $\bar{k}_t = b_3 \exp(a_3 \bar{s}_r)$.	$\bar{k}_t = 0{,}3608 \exp(0{,}6889\, \bar{s}_r)$	$\bar{k}_t = 0{,}2686 \exp(0{,}9578\, \bar{s}_r)$	$\bar{k}_t = 0{,}2815 \exp(0{,}7708\, \bar{s}_r)$
R	0,8804	0,9338	0,8761
R^2.	0,7751	0,8720	0,7675
MBE	-0,000292	-0,000408	-0,000267
RMSE	0,046559	0,042013	0,030279
t	0,020777	0,032236	0,029211
3.2) $\bar{k}_t = b'_3 \exp(a'_3 \bar{s'}_r)$	$\bar{k}_t = 0{,}3438 \exp(0{,}6465\, \bar{s'}_r)$	$\bar{k}_t = 0{,}2694 \exp(0{,}8973\, \bar{s'}_r)$	$\bar{k}_t = 0{,}2740 \exp(0{,}7821\, \bar{s'}_r)$
R	0.8822	0,9343	0,9477
R^2.	0,7783	0,8729	0,8981
MBE	-0,000092	0,001700	-0,000325
RMSE	0,113109	0,043209	0,029365

t	0,002688	0,130589	0,036710
4.1) $\overline{k}_t = b_4(\bar{s}_r)^{a_4}$	$\overline{k}_t = 0{,}6508(\bar{s}_r)^{0{,}4241}$	$\overline{k}_t = 0{,}6100(\bar{s}_r)^{0{,}4726}$	$\overline{k}_t = 0{,}5474(\bar{s}_r)^{0{,}3858}$
R	0.8842	0.9375	0.9425
R^2.	0,7819	0,8789	0,8883
MBE	0,000042	0,001575	0,000000
RMSE	0,045779	0,041801	0,030746
t	0,003019	0,125056	0,000000
4.2) $\overline{k}_t = b'_4(\overline{s'}_r)^{a'_4}$	$\overline{k}_t = 0{,}6341(\overline{s'}_r)^{0{,}4230}$	$\overline{k}_t = 0{,}5924(\overline{s'}_r)^{0{,}4712}$	$\overline{k}_t = 0{,}5346(\overline{s'}_r)^{0{,}3849}$
R	0.8860	0.9385	0.9434
R^2.	0,7851	0,8808	0,8901
MBE	0,000017	0,001625	0,000017
RMSE	0,045466	0,041493	0,030521
t	0,001216	0,129988	0,001811

5.1) $\overline{k}_t = a_5 + b_5(\overline{s}_r) + c_5\,(s_r)^2$	$\overline{k}_t = 0{,}1615 + 0{,}8366(\overline{s}_r) - 0{,}3776(s_r)^2$	$\overline{k}_t = 0{,}1742 + 0{,}6501(\overline{s}_r) - 0{,}2381(s_r)^2$	$\overline{k}_t = 0{,}2624 + 0{,}2707(\overline{s}_r) + 0{,}0735(s_r)^2$
$R_{y,12}$	0.8881	0.9328	0.9485
${R_{y.12}}^2$	0,7887	0,8700	0,8996
MBE	-0,000425	-0,000558	0,000383
RMSE	0,023686	0,017045	0,012293
t	0,059521	0,108699	0,103472
5.2) $\overline{k}_t = a'_5 + b'_5(\overline{s'}_r) + c'_5\,(s'_r)^2$	$\overline{k}_t = 0{,}1610 + 0{,}7890(\overline{s'}_r) - 0{,}3362(s'_r)^2$	$\overline{k}_t = 0{,}1669 + 0{,}6364(\overline{s'}_r) - 0{,}2279(s'_r)^2$	$\overline{k}_t = 0{,}2602 + 0{,}2647(\overline{s'}_r) + 0{,}0548(s'_r)^2$
$R_{y,12}$	0.8901	0.9428	0.9494
${R_{y,12}}^2$	0,7923	0,8889	0,9013
MBE	-0,000817	0,000125	-0,000167
RMSE	0,023429	0,016898	0,011872
t	0,115676	0,024535	0,046564

III.3. Interprétation des résultats

Dans cette partie, chacun des dix types de formulations obtenues sera cité au moyen de son numéro d'ordre (tel qu'il apparaît dans le tableau 4.1) c-à-d : (1.1) relation linéaire simple $\bar{k}_t$ vs $\bar{s}_r$; (1.2) relation linéaire simple $\bar{k}_t$ vs $\overline{s'}_r$; (2.1) fonction logarithmique $\bar{k}_t$ vs $\bar{s}_r$; (2.2) fonction logarithmique $\bar{k}_t$ vs $\overline{s'}_r$; (3.1) fonction exponentielle $\bar{k}_t$ vs $\bar{s}_r$; (3.2) fonction exponentielle $\bar{k}_t$ vs $\overline{s'}_r$; (4.1) fonction puissance $\bar{k}_t$ vs $\bar{s}_r$; (4.2) fonction puissance $\bar{k}_t$ vs $\overline{s'}_r$; (5.1) polynôme du second degré $\bar{k}_t$ vs $\bar{s}_r$; (5.2) polynôme du second degré $\bar{k}_t$ vs $\overline{s'}_r$.

III.3.1. Appréciation des formulations obtenues vis-à-vis des valeurs de R et R2, puis selon les sites

III.3.1.1. Classement de ces formulations, site par site, par ordre décroissant des valeurs

L'analyse des données du tableau 20 conduit au classement des dix formulations dans l'ordre ci-après (décroisant des valeurs de R ou de R^2), site par site (tableaux 21).

Tableaux 21. Classement des 10 types de formulations dans l'ordre décroissant des valeurs de R ou R^2

Tableau 21.1 : Bujumbura-Aéroport

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(5.1)	(5.2)	(2.1)	(4.1)	(2.2)	(4.2)	(1.1)	(1.2)	(3.1)	(3.2)

Bilan pour R et R^2 : R∈ [0,8775 ; 0,9133] ; R^2∈ [0,7700 ; 0,8342] ; $R_{moy} = 0{,}8934$

Tableau 21.2 : Gitega-Zege

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(5.2)	(1.2)	(5.1)	(1.1)	(3.2)	(3.1)	(2.1)	(4.2)	(4.1)	(2.2)

Bilan pour R et R^2 : R∈ [0,8766; 0,9002]; R^2∈ [0,7684; 0,8103] ; $R_{moy} = 0{,}8935$

Tableau 21.3 : Imbo-SEMS

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(3.1)	(5.2)	(1.2)	(3.2)	(4.1)	(4.2)	(2.2)	(5.1)	(1.1)	(2.1)

Bilan pour R et R^2 : R∈ $[0{,}8751; 0{,}8806]$; R^2∈ $[0{,}7657; 0{,}7705]$; $R_{moy} = 0{,}8786$

Tableau 21.4. : Musasa

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(5.2)	(2.2)	(5.1)	(2.1)	(1.2)	(4.2)	(1.1)	(4.1)	(3.2)	(3.1)

Bilan pour R et R^2 : R∈ $[0{,}8804 ; 0{,}8901]$; R^2∈ $[0{,}7751 ; 0{,}7923]$; $R_{moy} = 0{,}8862$

Tableau 21.5 : Ruvyironza

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(5.2)	(2.2)	(1.2)	(2.1)	(1.1)	(4.2)	(4.1)	(3.2)	(3.1)	(5.1)

Bilan pour R et R^2 : R∈ $[0{,}9328; 0{,}9428]$; R^2∈ $[0{,}8700; 0{,}9013]$; $R_{moy} = 0{,}9385$

Tableau 21.6 : Rwegura

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(5.2)	(1.2)	(2.1)	(5.1)	(1.1)	(3.2)	(4.2)	(4.1)	(2.2)	(3.1)

Bilan pour R et R^2 : R∈ $[0{,}8761; 0{,}9494]$; R^2∈ $[0{,}7675; 0{,}9013]$; $R_{moy} = 0{,}9395$

Ces résultats montrent que, sur les dix types de formulations obtenues, les six sites se classent comme suit (tableau 21.7) dans l'ordre décroissant des valeurs de la moyenne de R (et donc de R^2) :

Tableau 21.7 : Classement des six sites par ordre décroissant de R_{moy} sur les dix formulations

Sites	Rwegura	Ruvyironza	Gitega-Zege	Bujumbura-Aéroport	Musasa	Imbo-SEMS
R_{moy}	0,9395	0,9385	0,8935	0,8934	0,8862	0,8786

Ces mêmes résultats montrent également que, sur toutes les 60 (c-à-d 10 x 6) formulations obtenues (pour les 6 sites), R et R^2 varient entre les valeurs limites suivantes : $R_{min} = 0{,}8751$; ${R^2}_{min} = 0{,}7705$ (site d'Imbo-SEMS) et $R_{max} = 0{,}9494$; ${R^2}_{max} = 0{,}9013$ (site de Rwegura). Ces faits indiquent que, dans sa forme réduite à une régression linéaire (simple ou double selon le cas), chacune de ces 60 formulations présente un degré très élevé (car variant entre 87,51 % et 94,94 %) de corrélation positive entre la variable indépendante $x = \bar{s}_r$ ou $x' = \overline{s'}_r$ et la variable dépendante $y = \bar{k}_t$ (N.B : la corrélation est positive car les 2 variables x et y ou x' et y croissent dans le même sens). En d'autres termes, dans chacune de ces 60 formulations, un grand pourcentage (entre 77,05% et 90,13%) de la variance de la variable dépendante $y = \bar{k}_t$ est expliqué (ou prédit) par la variable indépendante $x = \bar{s}_r \ ou \ x' = \overline{s'}_r$.

III.3.1.2. Classement de ces formulations sur l'ensemble des six sites.

Les résultats des tableaux 21.1 à 21.6 de la sous-section III.3.1.1 précédente permettent de déduire, pour l'ensemble des six sites, le tableau ci-après (tableau 21.8) des fréquences d'apparition en différentes positions (du classement selon l'ordre décroissant des valeurs de R ou R^2) de chacun des dix (c-à-d des cinq couples de) types de formulations obtenues.

Tableau 21.8 : Fréquences d'apparition des 10 formulations dans diverses positions de leur classement, sur l'ensemble des six sites, par ordre décroissant des valeurs de R ou (de R^2)

N^{os} des positions	1	2	3	4	5	6	7	8	9	10
N^{os} des formula- tions et leurs fréquences d'apparition en ces positions	(5.2) 4 (5.1) 1 (3.1) 1 - -	(5.2) 2 (2.2) 2 (1.2) 2 - -	(5.1) 2 (2.1) 2 (1.2) 2 - -	(2.1) 2 (5.1) 1 (3.2) 1 (1.1) 1 (4.1) 1	(1.1) 2 (1.2) 1 (4.1) 1 (3.2) 1 (2.2) 1	(4.2) 4 (3.1) 1 (3.2) 1 - -	(1.1) 2 (2.1) 1 (2.2) 1 (4.1) 1 (4.2) 1	(4.1) 2 (3.2) 1 (5.1) 1 (4.2) 1 (1.2) 1	(3.1) 2 (2.2) 1 (3.2) 1 (1.1) 1 (4.1) 1	(3.1) 2 (5.1) 1 (2.1) 1 (2.2) 1 (3.2) 1

Un examen minutieux de ce tableau permet de dégager, sur l'ensemble des six sites, le classement très probable ci-après (classement sûr au moins jusqu'au 3-ième numéro d'ordre) des dix types de formulations obtenues, dans l'ordre décroissant des performances vis-à-vis de la variable statistique R ou R^2 : 1) polynôme du second degré $\bar{k}_t \; vs \; \overline{s'}_r$ (5.2) ; 2) relation linéaire simple $\bar{k}_t \; vs \; \overline{s'}_r$ (1.2) ; 3) fonction logarithmique $\bar{k}_t \; vs \; \bar{s}_r$ (2.1) ; 4) fonction logarithmique $\bar{k}_t \; vs \; \overline{s'}_r$ (2.2) ; 5) polynôme du second degré $\bar{k}_t \; vs \; \bar{s}_r$ (5.1) ; 6) relation linéaire simple $\bar{k}_t \; vs \; \bar{s}_r$ (1.1) ; 7) fonction puissance $\overline{k_t}$ vs $\overline{s'_r}$ (4.2) ; 8) fonction puissance $\overline{k_t}$ vs $\overline{s_r}$ (4.1) ; 9) fonction exponentielle $\bar{k}_t \; vs \; \bar{s}_r$ (3.1) ; 10) fonction exponentielle $\bar{k}_t \; vs \; \overline{s'}_r$ (3.2) .

De ce classement sur base de la variable statistique R et R^2, de l'ensemble des 60 formulations obtenues pour les sites et périodes de cette étude, il se dégage les deux faits suivants :

1) une prédominance des modèles $\bar{k}_t \; vs \; \bar{s}_r$ (ou $\overline{k}_t \; vs \; \overline{s'}_r$) des types polynôme du second degré, relation linéaire simple et fonction logarithmique (dans cet ordre) sur ceux des types fonction puissance et fonction exponentielle ;

2) un léger avantage de l'utilisation de la variable indépendante $\overline{s'}_r = \frac{\bar{s}}{\bar{s}_o}$ à la place de $\bar{s}_r = \frac{\bar{s}}{\bar{s}_o}$ dans la modélisation de $\bar{k}_t = \frac{\bar{G}}{\bar{G}_o}$.

III.3.2. Appréciation des formulations obtenues vis -à-vis des valeurs des autres variables statistiques

III.3.2.1. Vis-à-vis des valeurs de MBE

a) Classement de ces formulations, site par site, par ordre croissant des valeurs absolues de MBE

A partir des données du tableau 20, les dix types de formulations se classent comme suit par ordre croissant des valeurs absolues de MBE, site par site.

Tableaux 22 : Classement des 10 types de formulations par ordre croissant des valeurs absolues de MBE

Tableau 22.1 : Bujumbura-Aéroport

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(1.1)	(2.1)	(4.2)	(1.2)	(2.2)	(4.1)	(3.2)	(5.2)	(3.1)	(5.1)

Bilan pour MBE : MBE$\in [-0{,}000425; +0{,}000183]$; $\overline{MBE}$=-0,0000435

Tableau 22.2 : Gitega-Zege

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(1.2)	(2.2)	(5.1)	(2.1)	(4.2)	(5.2)	(4.1)	(1.1)	(3.1)	(3.2)

Bilan pour : MBE$\in [-0{,}000250; +0{,}000092]$; $\overline{MBE}$=-0,0000215

Tableau 22.3 : Imbo-SEMS

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(4.1)	(4.2)	(2.2)	(2.1)	(1.2)	(1.1)	(5.2)	(3.2)	(3.1)	(5.1)

Bilan pour MBE :$MBE \in [-0{,}000367; +0{,}000233]$; $\overline{MBE}$=-0,0000201

Tableau 22.4 : Musasa

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(2.2)	(4.2)	(4.1)	(3.2)	(1.2)	(2.1)	(1.1)	(3.1)	(5.1)	(5.2)

Bilan pour MBE: MBE$\in [-0{,}000817; +0.000042]$; $\overline{MBE}$=-0,0002025

Tableau 22.5 : Ruvyironza

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(1.2)	(2.1)	(5.2)	(1.1)	(2.2)	(3.1)	(5.1)	(4.1)	(4.2)	(3.2)

Bilan pour R et R^2 : R$\in [-0{,}000558; +0{,}001700]$; $\overline{MBE} = +0.0003675$

Tableau 22.6 : Rwegura

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(4.1)	(4.2)	(2.2)	(2.1)	(1.1)	(5.2)	(1.2)	(3.1)	(3.2)	(5.1)

Bilan pour MBE : MBE $\in [-0{,}000325; +0{,}000383]$; $\overline{MBE} = -0{,}0000535$.

Les 60 données de MBE montrent que, d'une formulation à une autre (parmi les 60 obtenues), l'écart moyen (MBE) entre les valeurs calculées ($y_{i;calc}$) et celles mesurées ($y_{i;mes}$) de la variable dépendante y (de la formulation réduite à sa forme linéaire) est : (i) tantôt positif, ce qui signifie une surestimation moyenne des $y_{i;mes}$ *par les* $y_{i;calc}$; *ii*) *tantôt négatif, ce qui signifie une* sous-estimation des $y_{i;cal}$ par les $y_{i,mes}$. Par ailleurs, pour toutes les 60 formulations obtenues, les valeurs de MBE sont très proches de (ou égales à) zéro. Donc, chacune de ces formulations ajuste très bien le nuage des points expérimentaux ($x_i; y_i$). Enfin, en considérant l'ordre croissant des moyennes $\overline{MBE}$ (des valeurs absolues de MBE) sur l'ensemble des 10 formulations par site, on dégage le classement ci-après (tableau 22.7) des six sites par ordre décroissant des performances de l'ensemble de ces 10 formulations.

Tableau 22.7 : Classement des six sites par ordre décroissant des performances, sur l'ensemble des 10 formulations, selon les valeurs de $\overline{MBE}$

N°d'ordre	1	2	3	4	5	6
Site	Gitega-Zege	Imbo-SEMS	Bujumbura-Aéroport	Rwegura	Musasa	Ruvyironza
$\overline{MBE}$	-0,0000175	-0,0000201	-0,0000435	-0,0000535	-0,0002025	+0,0003675

b) Ordre de performance de ces formulations sur l'ensemble des six sites

Des données des tableaux 22.1 à 21.6 du paragraphe précédent, on déduit pour l'ensemble des six sites le tableau ci-dessous (tableau 22.8) des fréquences d'apparition en différentes positions (du classement par ordre croissant des valeurs absolues de MBE) de chacun des dix types de formulations obtenues.

Tableau 22.8 : Fréquences d'apparition des 10 formulations dans diverses positions de leur classement, sur l'ensemble des six sites, par ordre croissant des valeurs absolues de MBE

N^{os} des positions	1	2	3	4	5	6	7	8	9	10
N^{os} des formulations et leurs fréquences d'apparition en ces positions	(4.1) 2 (1.2) 2 (1.1) 1 (2.2) 1 - -	(4.2) 3 (2.1) 2 (2.2) 1 - - -	(2.2) 2 (5.2) 1 (4.2) 1 (5.1) 1 (4.1) 1 -	(2.1) 3 (1.2) 1 (1.1) 1 (3.2) 1 - -	(2.2) 2 (1.2) 2 (4.2) 1 (1.1) 1 - -	(5.2) 2 (4.1) 1 (1.1) 1 (2.1) 1 (3.1) 1 -	(3.1) 1 (4.1) 1 (5.2) 1 (1.1) 1 (5.1) 1 (1.2) 1	(3.1) 2 (5.2) 1 (1.1) 1 (3.2) 1 (4.1) 1 -	(3.1) 3 (5.1) 1 (4.2) 1 (3.2) 1 - -	(5.1) 3 (3.2) 2 (5.2) 1 - - -

Une analyse des données de ce tableau conduit, sur l'ensemble des six sites, à l'ordre décroissant suivant des performances (vis-à-vis de la variable MBE) des dix types de formulations obtenues (réduites à leurs formes linéaires) : 1) fonction puissance $\bar{k}_t \; vs \; \bar{s}_r$ (4.1) ; 2) fonction puissance $\bar{k}_t \; vs \; \bar{s'}_r$ (4.2) ; 3) fonction logarithmique $\bar{k}_t \; vs \; \bar{s'}_r$ (2.2) ; ; 4)

relation linéaire simple $\bar{k}_t \; vs \; \overline{s'}_r$ (1.2) ; 5) fonction logarithmique $\bar{k}_t \; vs \; \bar{s}_r$ (2.1) ; 6) relation linéaire simple $\bar{k}_t \; vs \; \bar{s}_r$ (1.1) ; 7) polynôme du second degré $\bar{k}_t \; vs \; \overline{s'}_r$ (5.2) ; 8) polynôme du second degré $\bar{k}_t \; vs \; \bar{s}_r$ (5.1) ; 9) fonction exponentielle $\bar{k}_t \; vs \; \overline{s'}_r$ (3.2) ; 10) fonction exponentielle $\bar{k}_t \; vs \; \bar{s}_r$ (3.1). Ce classement ne change cependant pas le fait (déjà signalé) que, vis-à-vis de leurs valeurs de MBE, toutes les 60 formulations obtenues sont très performantes.

III.3.2.2. Vis-à-vis des valeurs de RMSE

a) Classement de ces formulations, site par site, par ordre croissant des valeurs de RMSE

Site par site, les données du tableau 4.1 permettent le classement ci-dessous (tableaux 4.4) des 10 types de formulations obtenues par ordre croissant des valeurs de RMSE.

Tableaux 23 : Classement des 10 types de formulations par ordre croissant des valeurs de RMSE

Tableau 23.1 : Bujumbura-Aéroport

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(5.1)	(5.2)	(2.1)	(2.2)	(1.1)	(1.2)	(4.1)	(4.2)	(3.1)	(3.2)

Bilan pour RMSE: RMSE $\in [0{,}015610; +0{,}035915]$; $\overline{RMSE}$ = 0,024

Tableau 23.2 : Gitega-Zege

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(5.2)	(1.2)	(5.1)	(1.1)	(2.2)	(2.1)	(3.2)	(3.1)	(4.2)	(4.1)

Bilan pour RMSE : RMSE$\in [0{,}017631; +0{,}040043]$; $\overline{RMSE}$ = 0,025

Tableau 23.3 : Imbo-SEMS

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(5.1)	(1.2)	(5.2)	(2.2)	(1.1)	(2.1)	(3.2)	(4.2)	(3.1)	(4.1)

Bilan pour RMSE :$RMSE \in [0{,}018346; +0{,}039778]$; $\overline{RMSE}$=0,027

Tableau 23.4 : Musasa

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(5.2)	(2.2)	(5.1)	(1.2)	(2.1)	(1.1)	(4.2)	(4.1)	(3.1)	(3.2)

Bilan pour RMSE: RMSE$\in [0{,}023429; +0.113109]$; $\overline{RMSE}$ = 0,039

Tableau 23.5 : Ruvyironza

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(5.2)	(5.1)	(1.2)	(2.1)	(1.1)	(2.2)	(4.2)	(4.1)	(3.1)	(3.2)

Bilan pour RMSE: RMSE $\in [0{,}016898; +0{,}043209]$; $\overline{RMSE} = +0{,}027$

Tableau 23.6 : Rwegura

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(5.2)	(1.2)	(1.1)	(5.1)	(2.2)	(2.1)	(3.2)	(3.1)	(4.2)	(4.1)

Bilan pour RMSE: RMSE $\in [0{,}011872; +0{,}030746]$; $\overline{RMSE} = +0{,}020$.

On remarque que pour chacune des 60 formulations obtenues, la valeur de l'écart quadratique moyen (RMSE) entre les $y_{i,cal}$ et $y_{i,mes}$ (dans la formulation réduite à sa forme linéaire) est très faible. Comme en plus l'écart moyen (MBE) est également très faible, on conclut que les valeurs calculées ($y_{i,cal}$) via chacune de ces formulations, ajustent très bien les valeurs expérimentales correspondantes ($y_{i,mes}$). De plus, si on considère l'ordre croissant des valeurs moyennes $\overline{RMSE}$ (de la variable RMSE) sur l'ensemble des 10 formulations par site, on est conduit au classement suivant (tableau 23.7) des six sites (par ordre décroissant des performances de l'ensemble de ces dix formulations).

Tableau 23.7 : Classement des six sites par ordre décroissant des performances de l'ensemble des 10 types de formulations, selon les valeurs croissantes de $\overline{RMSE}$

N°d'ordre	1	2	3	4	5	6
Site	Rwegura	Bujumbura-Aéroport	Gitega-Zege	Imbo-SEMS	Ruvyironza	Musasa
$\overline{RMSE}$	0,020	0,024	0,025	0,027	0,027	0,039

b) Classement de ces formulations sur l'ensemble des six sites

A partir des données des tableaux 23.1 à 23.6 du paragraphe précédent, on dégage pour l'ensemble des six sites, le tableau ci-après (tableau 23.7) des fréquences d'apparition en différentes positions (du classement par ordre croissant des valeurs de RMSE) de chacun des dix types de formulations obtenues.

Tableau 23. 8 : Fréquences d'apparition des 10 formulations dans diverses positions de leur classement, sur l'ensemble des six sites, par ordre croissant des valeurs de RMSE

N^{os} des positions	1	2	3	4	5	6	7	8	9	10
N^{os} des formulations et leurs fréquences d'apparition en ces positions	(5.2) 4 (5.1) 2 - - -	(1.2) 3 (5.2) 2 (5.1) 1 (2.2) 1 - -	(5.2) 2 (5.1) 2 (1.2) 1 (1.1) 1 - -	(2.2) 2 (1.2) 1 (1.1) 1 (2.1) 1 (5.1) 1 -	(1.1) 3 (2.2) 2 (2.1) 1 - - -	(2.1) 3 (1.2) 1 (1.1) 1 (2.2) 1 - -	(3.2) 3 (4.2) 2 (4.1) 1 - - -	(4.2) 2 (4.1) 2 (3.1) 2 - - -	(3.1) 4 (4.2) 2 - - - -	(5.1) 3 (3.2) 3 - - - -

Un examen fouillé de ces données conduit, sur l'ensemble des six sites, à l'ordre décroissant ci-après des performances (vis-à-vis de la variable RMSE) des 10 types de formulations obtenues (réduites à leurs formes linéaires) : 1) polynôme du second degré $\bar{k}_t \; vs \; \bar{s'}_r$ (5.2) ; 2) polynôme du second degré $\bar{k}_t \; vs \; \bar{s}_r$ (5.1) ; 3) relation linéaire simple $\bar{k}_t \; vs \; \bar{s'}_r$ (1.2) ; 4)

fonction logarithmique $\bar{k}_t \; vs \; \overline{s'}_r$ (2.2) ; 5) relation linéaire simple $\bar{k}_t \; vs \; \bar{s}_r$ (1.1) ; 6) fonction logarithmique $\bar{k}_t \; vs \; \bar{s}_r$ (2.1) ; 7) fonction exponentielle $\bar{k}_t \; vs \; \overline{s'}_r$ (3.2) ; 8) fonction puissance $\bar{k}_t \; vs \; \overline{s'}_r$ (4.2) ; 9) fonction puissance $\bar{k}_t \; vs \; \bar{s}_r$ (4.1) ; 10) fonction exponentielle $\bar{k}_t \; vs \; \bar{s}_r$ (3.1).

III.3.2.3. Vis-à-vis des valeurs de t

a) Classement de ces formulations, site par site, par ordre croissant des valeurs de t

A partir des données du tableau 20, on obtient site par site le classement ci-dessous (tableaux 24) des dix types de formulations obtenues par ordre croissant des valeurs de la variable t de Student.

Tableaux 24 : Classement des 10 types de formulations par ordre croissant des valeurs de t pour les six sites

Tableau 24.1 : Bujumbura-Aéroport

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(4.2)	(1.1)	(2.1)	(1.2)	(2.2)	(4.1)	(3.2)	(3.1)	(5.2)	(5.1)

Bilan pour t : t ∈ [0,000816; +0,090334]; $\bar{t}$ = 0,0179

Tableau 24.2 : Gitega-Zege

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(1.2)	(4.2)	(2.2)	(5.1)	(2.1)	(4.1)	(5.2)	(3.1)	(1.1)	(3.2)

Bilan pour t : t ∈ [0,001563; +0,021472]; $\bar{t}$ = 0,0072

Tableau 24.3 : Imbo-SEMS

N° **d'ordre**	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(4.1)	(4.2)	(2.2)	(2.1)	(1.2)	(1.1)	(3.2)	(3.1)	(5.2)	(5.1)

Bilan pour t: $t \in [0,001390; +0,066301]$; $\bar{t}$ = 0,0207

Tableau 24.4 : Musasa

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(4.2)	(2.2)	(3.2)	(4.1)	(2.1)	(1.2)	(3.1)	(1.1)	(5.1)	(5.2)

Bilan pour t: t ∈ [0,001216; +0,115676]; $\bar{t}$ = 0,0272

Tableau 24.5 : Ruvyironza

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(1.2)	(2.1)	(5.2)	(3.1)	(1.1)	(2.2)	(5.1)	(4.1)	(4.2)	(3.2)

Bilan pour t: t ∈ [0,006466; 0,130589]; $\bar{t} = +0{,}0660$

Tableau 24.6 : Rwegura

N° d'ordre	1	2	3	4	5	6	7	8	9	10
N° de la formulation	(4.1)	(4.2)	(2.2)	(2.1)	(3.1)	(1.1)	(3.2)	(5.2)	(1.2)	(5.1)

Bilan pour t: t ∈ [0,000000; 0,103472]; $\bar{t} = 0{,}0320$.

On note que pour chacune des 60 formulations obtenues, la valeur de la variable t de Student est très inférieure à sa valeur critique (pour n-1 = 11 ddl et un niveau de confiance $\gamma = 1 - \alpha = 99{,}5\ \%$) $t_c = 3{,}106$. Ce fait indique que, à des erreurs statistiques près, les valeurs $y_{i,calc}$ déterminées à l'aide de chacune de ces 60 formulations, sont très peu différentes des valeurs expérimentales correspondantes $y_{i,mes}$ (dans la formulation réduite à sa forme linéaire). En plus, en considérant l'ordre croissant des valeurs moyennes $\bar{t}$ (de la variable t) sur l'ensemble des dix formulations par site, on aboutit au classement ci-après (tableau 24.7) des six sites (par ordre décroissant des performances de cet ensemble des dix formulations).

Tableau 24.7 : Classement des six sites par ordre décroissant des performances, sur l'ensemble des dix formulations, selon les valeurs moyennes $\bar{t}$ de la variable t

N° d'ordre	1	2	3	4	5	6
Site	Gitega-Zege	Bujumbura-Aéroport	Imbo-SEMS	Musasa	Rwegura	Ruvyironza
$\bar{t}$	0,0072	0,0179	0,0207	0,0272	0,0320	0,0660

Dans ce classement, les sites se situent pratiquement comme dans le classement correspondant basé sur les moyennes $\overline{MBE}$ de la variable écart moyen (voir tableau 24.7), sauf que les numéros d'ordre suivants sont intervertis : 2 et 3 d'une part, 4 et 5 d'autre part. Ce fait est logique puisque les variables t et MBE sont directement proportionnelles (l'une à l'autre) selon l'éq. (2.4.3).

b) Ordre de performance de ces formulations sur l'ensemble des six sites

Les données des tableaux 24.1 à 24.6 du paragraphe précédent permettent de déduire, pour l'ensemble des six sites, le tableau ci-après (tableau 24.8) des fréquences d'apparition en différentes positions (du classement par ordre croissant des valeurs de la variable t) de chacun des dix types de formulations obtenues.

Tableau 24.8 : Fréquences d'apparition des 10 formulations dans diverses positions de leur classement, sur l'ensemble des six sites, par ordre croissant des valeurs de la variable t

N^{os} des positions	1	2	3	4	5	6	7	8	9	10
N^{os} des formulations et leurs fréquences d'apparition en ces positions	(4.2) 2 (1.2) 2 (4.1) 2 - - -	(4.2) 3 (1.1) 1 (2.2) 1 (2.1) 1 - -	(2.2) 3 (2.1) 1 (3.2) 1 (5.2) 1 - -	(2.1) 2 (1.2) 1 (5.1) 1 (4.1) 1 (3.1) 1 -	(2.1) 2 (2.2) 1 (1.2) 1 (1.1) 1 (3.1) 1 -	(4.1) 2 (1.1) 2 (1.2) 1 (2.2) 1 - -	(3.2) 3 (5.2) 1 (3.1) 1 (5.1) 1 - -	(3.1) 3 (1.1) 1 (4.1) 1 (5.2) 1 - -	(5.2) 2 (1.1) 1 (5.1) 1 (4.2) 1 (1.2) 1 -	(5.1) 3 (3.2) 2 (5.2) 1 - - -

Une analyse de ces données, sur l'ensemble des six sites, permet de dégager l'ordre décroissant ci-après des performances (vis-à-vis de la variable t de Student) des dix types de formulations obtenues (réduites à leurs formes linéaires) : 1) fonction puissance $\bar{k}_t \; vs \; \overline{s'}_r$ (4.2) ; 2) fonction logarithmique $\bar{k}_t \; vs \; \overline{s'}_r$ (2.2) ; 3) fonction logarithmique $\bar{k}_t \; vs \; \overline{s}_r$ (2.1) ; 4) relation linéaire simple $\bar{k}_t \; vs \; \overline{s'}_r$ (1.2) ; 5) fonction puissance $\bar{k}_t \; vs \; \bar{s}_r$ (4.1) ; 6) relation linéaire simple $\bar{k}_t \; vs \; \bar{s}_r$ (1.1) ; 7) fonction exponentielle $\bar{k}_t \; vs \; \overline{s'}_r$ (3.2) ; 8) fonction exponentielle $\bar{k}_t \; vs \; \bar{s}_r$ (3.1) ; 9) polynôme du second degré $\bar{k}_t \; vs \; \overline{s'}_r$ (5.2) ; 10) polynôme du second degré $\bar{k}_t \; vs \; \bar{s}_r$ (5.1).

III.3.3. Compléments d'appréciation des formulations obtenues sur base des valeurs des variables statistiques considérées prises globalement

Sur base des tableaux synthétiques des résultats des sous-sections III.3.1 et III.3.2 précédentes, la suite du texte dans la présente sous-section vise la déduction d'éléments d'appréciation (complémentaires à ceux antérieurement formulés) des dix types de formulations obtenues.

III.3.3.1. Tableaux synthétiques

Les tableaux 25 et 26 ci-dessous font une synthèse des résultats des sous-sections susmentionnées.

Tableau 25 : Classement, sur l'ensemble des 10 types de formulations, des 6 sites par ordre des moyennes des 4 variables statistiques considérées

N°d'ordre / Variable	1	2	3	4	5	6
R_{moy} décroissant	Rwegura 0,9395	Ruvyironza 0,9385	Gitega-Zege 0,8935	Bujumbura-Aéroport 0,8934	Musasa 0,8862	Imbo-SEMS 0,8786
$\overline{MBE}$ croissant en valeur absolue	Gitega-Zege -0,0000175	Imbo-SEMS -0,000021	Bujumbura-Aéroport -0,0000435	Rwegura -0,0000535	Musasa -0,0002025	Ruvyironza 0,0003675
$\overline{RMSE}$ croissant	Rwegura 0,020	Bujumbura-Aéroport 0,024	Gitega-Zege 0,025	Ruvyironza 0,027	Imbo-SEMS 0,027	Musasa 0,039
$\bar{t}$ croissant	Gitega-Zege 0,0072	Bujumbura-Aéroport 0,0179	Imbo-SEMS 0,027	Musasa 0,0272	Rwegura 0,0320	Ruvyironza 0,0660

Tableau 26 : Classement, sur l'ensemble des 6 sites, des 10 types de formulations par ordre décroissant des performances, selon les valeurs des 4 types de variables statistiques statistiques

N° d'ordre / Variable	1	2	3	4	5	6	7	8	9	10
R ou R^2 décroissant	(5.2)	(1.2)	(2.1)	(2.2)	(5.1)	(1.1)	(4.2)	(4.1)	(3.1)	(3.2)
MBE croissant en valeur absolue	(4.1)	(4.2)	(2.2)	(1.2)	(2.1)	(1.1)	(5.2)	(5.1)	(3.2)	(3.1)
RMSE croissant	(5.2)	(5.1)	(1.2)	(2.2)	(1.1)	(2.1)	(3.2)	(4.2)	(4.1)	(3.1)
t croissant	(4.2)	(2.2)	(2.1)	(1.2)	(4.1)	(1.1)	(3.2)	(3.1)	(5.2)	(5.1)

III.3.3.2. Premiers compléments d'appréciation

Les données des tableaux 25 et 25, en plus de celles du tableau 20, permettent de formuler les faits complémentaires marquants ci-après :

1. Chacune des 60 formulations obtenues (réduite à sa forme linéaire) ajuste de façon très satisfaisante les données expérimentales $(x_i ; y_i)$ ou $(x'_i ; y_i)$ (où x = $\bar{s}_r$, x′ = $\bar{s}'_r$ et y = $\bar{k}_t$) car :

(i) les valeurs de R sont toutes proches de 1 ($R \in [0{,}8751 ; +0{,}9494]$) ;

ii) celles de MBE sont proches de (ou égales à) zéro (avec $\overline{MBE} \in [-0{,}0000175 ; +0{,}0003675]$) ;

iii) celles de RMSE sont très faibles p/r aux y_i (avec $\overline{RMSE} \in [0{,}020 ; +0{,}039]$) ;

iv) celles de t sont toutes très inférieures à leur valeur critique $t_c = 3{,}106$ ($avec\ \bar{t} \in [0{,}0072 ; 0{,}0660]$).

2. Même si le degré de performance de chacune des 60 formulations obtenues reste très satisfaisant vis-à-vis de chacune des variables statistiques (de test de cette performance), le

niveau de performance de l'ensemble des dix formulations diminue légèrement d'un site à l'autre selon les classements ci-après :

(i) Rwegura, Ruvyironza, Gitega-Zege, Bujumbura-Aéroport, Musasa et Imbo-SEMS, si on considère l'ordre décroissant de R_{moy} ;

(ii) Pratiquement le même classement que précédemment si on considère l'ordre croissant de $\overline{RMSE}$, sauf un interchangement d'emplacements pour Bujumbura-Aéroport et Ruvyironza, puis un autre pour Musasa et Imbo-SEMS ;

(iii) Gitega-Zege, Imbo-SEMS, Bujumbura-Aéroport, Rwegura, Musasa et Ruvyironza, si on considère l'ordre croissant de $\overline{MBE}$ en valeur absolue ;

(iv) Pratiquement le même classement que précédemment si on considère l'ordre croissant de $\bar{t}$, à part un interchangement d'emplacements pour Bujumbura-Aeroport et Imbo-SEMS, puis un autre pour Rwegura et Musasa.

3. Sur l'ensemble des six sites, dans le classement des dix types de formulations par ordre décroissant de R (ou R^2), comme par ordre croissant de RMSE, on note que :

(i) les formulations ($\bar{k}_t \, vs \, \bar{s}_r \; ou \; \bar{k}_t \, vs \, \overline{s'}_r$) de types : polynôme du second degré, relation linéaire simple et fonction logarithmique, occupent les six premières positions, tandis que celles de types : fonction exponentielle et fonction puissance occupent les quatre dernières positions ;

(ii) dans ces premières comme dans ces dernières positions, les formulations $\bar{k}_t \, vs \, \overline{s'}_r$ apparaissent plus fréquemment avant les formulations $\bar{k}_t \, vs \, \bar{s}_r$. Cela est un indicateur certain d'un léger avantage de l'utilisation de la variable indépendante $\overline{s'}_r = \frac{\bar{s}}{\bar{s}'_O}$ à la place $\bar{s}_r = \frac{\bar{s}}{\bar{s}_O}$ dans l'estimation de la variable dépendante $\bar{k}_t = \frac{\bar{G}}{\bar{G}_O}$ pour les sites et périodes de la présente étude.

4. De son côté, le classement (sur l'ensemble des six sites) des dix types de formulations dans l'ordre croissant de MBE en valeur absolue comme dans l'ordre croissant de t, montre que :

(i) les formulations de type fonction puissance sont dans le peloton des six premières positions, tandis que les formulations de type polynôme du second degré rejoignent le peloton des quatre dernières positions ;

(ii) dans ces premières comme dans ces dernières positions, la même apparition plus fréquente des formulations $\bar{k}_t$ vs $\overline{s'}_r$ avant les formulations $\bar{k}_t$ vs $\bar{s}_r$ (et donc le léger avantage de l'utilisation de la variable $\overline{s'}_r$ à $la\ place\ \ de\ \bar{s}_r$ stipulé au point n°3 (ii)), reste observée.

III.3.3.3. Autres compléments d'appréciation

Tels que présentés dans les tableaux 20, 25 et 26, les intervalles (ou o .d .g) des valeurs des variables statistiques utilisées pour tester les performances des formulations obtenues dans le présent travail, s'accordent très bien avec la gamme des valeurs publiées dans d'autres travaux sur les mêmes types de formulations. Les exemples de résultats proposés dans le tableau 27 ci-dessous témoignent de ce fait.

Tableau 27 : Valeurs de R^2, MBE, RMSE, et t obtenues dans quelques autres travaux sur les mêmes types de modèles $\overline{k_t}$ vs $\overline{s_r}$ qu'ici

N^{os}, sites (Régions)	Types de modèles	R^2	MBE	RMSE	T	Auteurs et réf.	Année
1. Constanta (Roumanie)	(1.1) ; (5.1)	[0,787 ; 0,794]	-	-	-	Paltineau, Cr. et al. [67]	2002
2. Six sites (Egypte)	(1.1) ; (2.1) ; (5.1)	-	0,001	0,06	-	El-Metwally, M. [56]	2005
3. Kuala Terengganu (Malaisie)	(1.1) ; (2.1) ; (3.1) ; (5.1)	0,71	[0,0000;0,0065]	[0,0337;0,0420]	[0,0015;0,9800]	Muzathik, A.M. et al. [68]	2011
4. Khatmandu (Himalaya)	(1.1)	0,71	0,055	0,071	-	Poudyal, K. et al. [54]	2012
5. Kebbi (Nigeria)	(1.1) ; (4.1) ; (5.1)	[0,9620;0,9880]	-	-	-	Gana, N.N. et al. [29]	2013
6. Cinq sites (Bangladesh)	(1.1) ; (2.1) ; (3.1) ; (4.1) ; (5.1)	[0,95165;0,98538]	[-0,00049 ;0,00425]	[0,08719;0,13805]	[0,01023; 0,11858]	Sarkar, M. N.I. et al. [44]	2016
7. Anantapur (Inde)	(1.1)	[0,70;0,86]	-	-	-	Reddy, O.S. et al. [39]	2018
8.Trois stations (Irak)	24 modèles dont (1.1) ; (2.1) ; (3.1) ; (4.1) ; (5.1)	[0,97;0,99]	-	-	-	Uçkan, I. et al. [69]	2022
9. Maiduguri (Nigeria)	9 modèles dont (1.1) ; (2.1) ; (3.1) ; (4.1) ; (5.1)	[0,592;0,777]	-	-	[0,0392;0,1912]	Akpootu, D.O. et al. [38]	2023
10. 85 stations (5 zones ; Chine)	21 modèles	[0,82;0,99]	-	-	-	Lan, K. et al. [31]	2023
11. Osogbo, Osun State (Nigeria)	(1.1) ; (5.1)	[0,677;0,760]	-	-	-	Ohunakin, O.S. et al. [40]	2013

CHAPITRE IV : SYNTHESE, CONCLUSION ET PERSPECTIVES

IV.1. Synthèse des étapes du travail

L'objectif visé dans ce travail était d'établir, pour des sites choisis au Burundi, différentes expressions analytiques (ou formulations) empiriques reliant, en termes de moyennes mensuelles (sur une assez longue période), l'irradiation solaire globale journalière G (via l'indice de clarté journalier k_t) d'une part, à la durée réelle d'ensoleillement journalier S (via chacune des durées relatives d'ensoleillement s_r et s'_r) d'autre part. A cette fin, en premier lieu (premier chapitre), divers concepts en rapport avec ce sujet ont été décrits, l'objectif précédent a été motivé, puis les cinq couples (c-à-d les dix types) suivants de formulations (k_t vs s_r et k_t vs s'_r) à établir ont été choisis : i) relations linéaires simples, ii) fonctions logarithmiques, iii) fonctions exponentielles, iv) fonctions puissances, v) polynômes du second degré.

En deuxième lieu (au deuxième chapitre), d'abord les moyennes mensuelles $\bar{S}$ (de S) et $\bar{G}$ (de G), comme premières données de base utilisées dans cette étude, ont été présentées.

Telles que recueillies auprès des services de l'IGEBU, ces données se rapportent à des périodes de 20 ans ou légèrement plus, puis aux sites ci-après : i) Bujumbura-Aéroport, ii)Gitega-Zege, iii) Imbo-SEMS, iv) Musasa, v) Ruvyironza, et vi) Rwegura. Ces périodes et sites ont été sélectionnés sur base des deux critères suivants : i) la disponibilité simultanée des deux types de moyennes ($\bar{S}$ et $\bar{G}$) sur une plus longue période possible et avec le moins d'interruptions possibles des mesures, ii) la couverture de plusieurs régions éco-climatiques du pays. Ensuite, pour chacun des six sites précédents et des 365 jours d'une année normale, les valeurs de chacune des grandeurs suivantes ont été calculées : i) durée théorique d'ensoleillement (S_o), ii) durée théorique d'ensoleillement modifiée (S'_o) tenant compte de l'horizon naturel du site (au lieu de son horizon astronomique), iii) irradiation solaire journalière hors atmosphère sur une surface horizontale au sol (G_o). De ces valeurs, les moyennes mensuelles ($\bar{S}_o$,$\bar{S'}_o$ et $\bar{G}_o$) ont été déduites. Puis, comme deuxièmes données de base utilisées dans cette étude, des échantillons de 12 valeurs mensuelles des deux couples de variables ($x = \bar{s}_r = \frac{\bar{S}}{\bar{S}_0}$, $y = \bar{k}_t = \frac{\bar{G}}{\bar{G}_0}$) et ($x' = \bar{s'}_r = \frac{\bar{S}}{\bar{S'}_0}$, $y = \bar{k}_t = \frac{\bar{G}}{\bar{G}_0}$) ont été établis pour chacun des six sites. Des commentaires sur les variations saisonnières et les moyennes annuelles, site par site, des huit grandeurs $\bar{S}, \bar{G}, \bar{S}_o , \bar{S'}_o, \bar{s}_r, \bar{s'}_r , \bar{G_0}$ et $\bar{k}_t$ ont même été émis. Une fois toutes ces données de base acquises, l'étape qui a suivi a été d'abord de noter

chacune des dix formulations (à établir par site) sous forme d'une régression linéaire (simple ou double selon le cas). Puis, pour chacun de ces deux types de régression linéaire, un exposé succinct a été opéré sur la méthode (celle des moindres carrés) et les expressions analytiques (résultantes) de détermination des coefficients de régression, du coefficient de corrélation (R) et du coefficient de détermination (R^2). Une description sommaire a été ensuite effectuée (expressions analytiques de définition à l'appui) sur les trois variables statistiques complémentaires ci-après à utiliser pour évaluer le degré de performance de chacun des deux types précédents de régression linéaire : i) l'écart moyen (MBE) entre les valeurs calculées ($y_{i,calc}$) de la variable dépendante y (dans la régression linéaire considérée) et les valeurs mesurées ($y_{i,mes}$) de cette variable ; ii) l'écart quadratique moyen (RMSE) entre ces deux types de valeurs ; iii) la variable (t) de Student. Il a été également expliqué en quoi doit consister le test du degré de performance (de chacun des deux types de régression linéaire) vis-à-vis de chacune des cinq variables statistiques (R, R^2, MBE, RMSE et t).

En troisième lieu (troisième chapitre), il a été d'abord présenté, pour l'un des six sites (Musasa), des exemples des calculs effectués et des résultats obtenus sur des régressions linéaires (simples ou doubles). Ces calculs et résultats ont concerné la détermination des coefficients de régression et des cinq variables statistiques susmentionnées. Des calculs similaires aux précédents ayant été opérés pour les autres sites, il a été ensuite proposé un tableau synthétique des résultats. Celui- ci contient, pour les six sites, les soixante (10 x 6 = 60) formulations ($\bar{k}_t \; vs \; \bar{s}_r \; et \; \bar{k}_t \; vs \; \bar{s'}_r$) obtenues, ainsi que les valeurs des cinq variables statistiques (R, R^2, MBE, RMSE et t) pour chacune des 10 types de formulations. Une interprétation des résultats du tableau précédent a été enfin effectuée en les termes suivants :

i) appréciation du degré de performance de chacune des 10 types de formulations par site vis-à-vis des valeurs de chacune des 5 variables statistiques considérées ;

ii) classement, sur l'ensemble des 10 types de formulations, des six sites par ordre décroissant de performance, selon les valeurs moyennes de chacune des variables statistiques ;

iii) classement, sur l'ensemble des six sites, des 10 formulations également par ordre décroissant de performance selon les valeurs obtenues des variables statistiques,

iv) comparaison des intervalles des valeurs obtenues des variables statistiques des formulations de cette étude, avec les valeurs des mêmes variables proposées dans la littérature (dans divers travaux sur les mêmes types de formulations qu'ici).

IV.2. Synthèse des principaux faits dans les résultats

Des résultats du travail tels que présentés aux chapitres 2 et 3, il ressort les faits saillants ci-après pour les périodes et sites de cette étude :

1. Les trois mois les plus ensoleillés de l'année en chacun des six sites sont juin, juillet et août, ce qui correspond à la période la plus sèche (ou la moins pluvieuse). En effet, dépendamment du site, pour ces mois, $\bar{S}$ varie de 6,73 à 9,47 heures, $\bar{s}_r$ varie entre 0,57 et 0,80, $\overline{s'}_r$ varie de 0,61 et 0,86, puis $\bar{k}_t$ varie de 0,46 à 0,59.

2. Par ordre décroissant des moyennes annuelles de chacune des grandeurs $\overline{S}, \bar{G}, \bar{s}_r, \overline{s'}_r$ et $\bar{k}_t$, le classement des six sites est pratiquement le même, c-à-d le suivant : Musasa, Gitega-Zege, Bujumbura-Aéroport, Imbo-SEMS, Ruvyironza, Rwegura.

3. Les moyennes annuelles obtenues des deux types de durée relative d'ensoleillement journalier ($s_r \in [0{,}453; 0{,}586]\ et\ s'_r \in [0{,}481; 0{,}622]$) et de l'indice de clarté journalier ($\bar{k}_t \in [0{,}40; 0{,}52]$) caractérisent un ciel (ou une atmosphère) moyennement clair(e) ou partiellement nuageux (nuageuse) aux sites considérés.

4. Chacune des 60 formulations obtenues ajuste de façon très satisfaisante le nuage des points expérimentaux (x_i, y_i) ou (x'_i, y_i) de l'échantillon correspondant des valeurs des variables $(x = \bar{s}_r;\ y = \bar{k}_t)$ ou $(x' = \overline{s'}_r;\ y = \bar{k}_t)$. En effet, pour chacune de ces formulations (notée sous forme d'une régression linéaire), R est proche de 1 (R$\in [0{,}8751; 0{,}9494]$), MBE est très proche de (ou égale à) zéro ($\overline{MBE} \in [-0{,}0000175; 0{,}0003675]$), RMSE est très faible p/r aux valeurs $y_{i,mes}$ ($\overline{RMSE} \in [0{,}020; 0{,}039]$) et la variable t de Student est très inférieure à sa valeur critique t_c ($\bar{t} \in [0{,}0072; 0{,}0660]\ avec\ t_c$ =3,106).

5. Même si le niveau de performance de chacune de ces 60 formulations reste très élevé vis-à-vis de chacune des cinq variables statistiques (R, R^2, MBE, RMSE et t) utilisées (pour tester ce niveau), le degré de performance de l'ensemble des 10 types de formulations (par site) diminue légèrement d'un site à l'autre comme suit si l'on considère l'ordre décroissant des moyennes R (ou R^2) ou l'ordre croissant de $\overline{RMSE}$: Rwegura, Ruvyironza, Gitega-Zege, Bujumbura-Aéroport, Musasa, Imbo-SEMS. N.B : pour le classement selon l'ordre croissant de $\overline{RMSE}$, il y a (dans le classement précédent) un interchangement d'emplacements pour Ruvyironza et Bujumbura-Aéroport, puis un autre pour Musasa et Imbo-SEMS.

6. Sur l'ensemble des six sites, le classement des 10 types de formulations (par site) par ordre décroissant de R (ou R^2), comme par ordre croissant de RMSE, indique que les formulations ($\bar{k}_t\ vs\ \bar{s}_r\ ou\ \bar{k}_t\ vs\ \overline{s'}_r$) de types : polynôme du second degré, relation linéaire simple et fonction logarithmique, occupent les six premières positions, alors que celles des types : fonction exponentielle et fonction puissance occupent les quatre dernières positions.

7. Toujours sur l'ensemble des six sites, le classement des 10 types de formulations (par site) par ordre décroissant de R (ou R^2), comme par ordre croissant de MBE (en valeur absolue), de RMSE et de t, montre que dans les six premières comme dans les quatre dernières positions, les formulations $\bar{k}_t\ vs\ \overline{s'}_r$ apparaissent plus fréquemment avant les formulations $\bar{k}_t\ vs\ \overline{s}_r$. Cela est un indicateur certain d'un léger avantage de l'utilisation de la variable indépendante $\overline{s'}_r = \frac{\bar{S}}{\bar{S'}_O}$ à la place de $\bar{s}_r = \frac{\bar{S}}{\bar{S}_O}$ dans l'estimation de la variable dépendante $\bar{k}_t = \frac{\bar{G}}{G_O}$ (via les modèles considérés pour les périodes et les sites de la présente étude). En d'autres termes, dans le cas d'étude en question ici, il y a un léger avantage à utiliser la durée théorique d'ensoleillement journalier S'_o modifiée en tenant compte de l'horizon naturel du site, à la place de la durée théorique d'ensoleillement journalier (non modifiée) S_o.

8. Enfin, un survol de la littérature sur divers modèles d'estimation de l'irradiation solaire globale journalière sur une surface horizontale à l'aide de la durée réelle (S) d'ensoleillement journalier (en termes de moyennes mensuelles sur une période de plusieurs années), montre que les intervalles des valeurs obtenues pour les cinq variables statistiques de la présente étude, s'accordent très bien avec la gamme des valeurs publiées dans d'autres travaux sur ces variables et pour les mêmes types de formulations qu'ici.

IV.3. Conclusion et perspectives

La présente étude a utilisé entre autres données de base, celles recueillies au sein des services de l'IGEBU et se rapportant à des périodes de vingt ans ou légèrement plus. Dix types de formulations ont été établies pour chacun des six sites choisis au Burundi, pour estimer (en termes de moyennes mensuelles) l'irradiation solaire globale journalière (G) sur une surface horizontale (via la variable dépendante « indice de clarté journalier », k_t) à l'aide de la durée réelle (S) d'ensoleillement journalier (via chacun des deux types de durée relative d'ensoleillement journalier, s_r et s'_r).

Par rapport aux valeurs obtenues de chacune des cinq variables statistiques (R, R^2, MBE, RMSE et t) utilisées, le niveau de performance de chacune des 60 formulations obtenues est trouvé très satisfaisant. De plus, ce niveau de performance est comparable à celui observé dans d'autres travaux sur les mêmes types de formulations.

Les expressions analytiques $\bar{k}_t$vs $\bar{s}_r$ et $\bar{k}_t$vs $\overline{s'}_r$ obtenues dans cette étude constituent un outil important dans la conception puis l'installation de systèmes héliotechniques aux sites considérés ou en des sites de même climat que ces derniers. Dans des travaux ultérieurs, il serait bon d'envisager, pour plusieurs sites du Burundi, l'établissement puis l'inter-comparaison, de divers modèles d'estimation de l'irradiation solaire globale journalière (G) sur une surface horizontale (ou inclinée) et même ses composantes directe (I) et diffuse (D), à l'aide des données d'autres (ou de diverses associations de) variables climatologiques (exemples dans la sous-section 1.3.2.2). Les résultats de telles études permettraient notamment, non seulement d'établir (puis de corriger périodiquement) des cartes nationales des grandeurs (comme S, G, I et D) liées au rayonnement solaire, mais aussi de définir (pour les sites du Burundi) une année météorologique typique (TMY) relativement à chacune des grandeurs précédentes.

REFERENCES

[1] Plan National de Développement (PND Burundi 2018-2027) (2018, August 31)

Annexe 2, 73-87.

https://www.presidence.gov.bi/wp-content/uploads/2018/08/PND-Burundi-2018-20

27-Version-Finale.pdf

[2] Ministère de l'Energie et des Mines, Etude Diagnostique du Secteur de l'Energie au Burundi dans le Cadre de l'Initiative du Secrétaire Général des Nations Unies sur l'Energie Durable pour Tous. (juin 2013) *Sustainable Energy for All*.

https://www.se4allafrica.org/fileadmin/uploads/se4all/Documents/Country_RAGAs/Burundi_Rapid_Assessment_Gap_Analysis__FR_.pdf

[3] International Rankings of Burundi-2017; Burundi Data Portal.

https://burundi.opendataforafrica.org/data#menu=topic

[4] Complete Vision Burundi 2025 (June 2011).

https://www.presidence.gov.bi/archives/IMG/pdf/Vision_Burundi_2025_complete_

FR.pdf

[5] République du Burundi, Cadre Stratégique de Croissance et de Lutte contre la Pauvreté,

CSLP II: 2012-2015, Bilan de Mise en Oeuvre.

https://www1.undp.org/content/dam/burundi/docs/publications/Rapport Bilan

global du CSLPII_final.pdf

[6] Bashahu, M. (1985) Techniques de Mesure et Modélisation des Composantes du Rayonnement Solaire, Applications au Site de Dakar. Thèse, Université de Dakar, Dakar.

[7] Robinson, N. (1966) Solar Radiation. *Elsevier Publishing Company,* Amsterdam, London, New York.

[8] Iqbal, M. (1983) An Introduction to Solar Radiation. *Academic Press*, Toronto, New York,…,Tokyo.

[9] Ratto, C.F. (1986) Sun-Earth Astronomical Relationships and the Extraterrestrial Solar Radiation. In : Guzzi, R. and Justus, C.G., Eds., Physical Climatology for Solar and Wind Energy, *World Scientific*, 56-147.

[10] Ntirandekura, J. (2017) Ajustement, à l'aide de FDP Béta Appropriées, des Distributions Mensuelles des Fréquences Relatives de l'Indice de Clarté et de la Fraction Diffuse Horaires en un Site Africain. Mémoire, Université du Burundi, Bujumbura.

[11] Ma, C.C.Y. and Iqbal, M. (1984) Statistical Comparison of Solar Radiation Correlations. *Solar Energy*, 33(2), 143-148.

[12] Bashahu, M. and Nkundabakura, P. (1994) Analysis of Global Irradiation Data for Five Sites in Rwanda and One in Senegal. *Renewable Energy*, 4(4), 425-435.

[13] https://energieplus-lesite.be/theories/climat8/ensoleillement-d8/ ; consulté le 22/4/2023

[14] Bashahu, M. (1982) Mesure de la Distribution Spectrale du Rayonnement Solaire. Mémoire, Université de Dakar, Dakar.

[15] Mpanzimana, D. (2005) Estimation du Rayonnement Solaire Global à l'Aide de Paramètres Météorologiques en Différents Sites du Burundi. Mémoire, Université du Burundi, Bujumbura.

[16] Iqbal, M. (1986) Spectral and Total Sun Radiance Under Cloudless Skies. In : Guzzi, R. and Justus, C.G., Eds., Physical Climatology for Solar and Wind Energy, *World Scientific*, 196-242.

[17] Bashahu, M. and Laplaze, D. (1994) An Atmospheric Model for Computing Solar Radiation. *Renewable Energy*, 4(4), 455-458.

[18] Ahmad, M. J. and Tiwari, G.N. (2009) Optimization of Tilt for Solar Collector to Receive Maximum Radiation. *The Open Renewable Energy Journal*, 2, 9-24.

[19] Bashahu, M. (2003) Statistical Comparison of Models for Estimating the Monthly Average Daily Diffuse Radiation at a Substropical African Site. *Solar Energy*, 75, 43-51.

[20] Ångström, A., (1924) Solar and Terrestrial Radiation. Q.J. Royal Met. Soc., 50, 121-126.

[21] Andretta, A., Bartoli, B., Coluzzi, B., Cuomo, v., Francesca, M. and Serio, C. (1981) Global Solar Radiation Estimation from Relative Sunshine Hours in Italy. *Journal of Applied Meteorology and Climatology*, 2(10), 1377-1384.

[22] Abdul-Aziz, J., Nogi, A.A. and Zumailan, A.A.R (1993) Global Solar Radiation Estimation from Relative Sunshine Hours in Yemen . *Renewable Energy*, 3 (6-7), 645-653.

[23] Cao, Q., Liu, Y., Lyu, K., Lu, Y., Li, D.H.W. and Yang, L. (2020) Solar Radiation Zoning and Daily Global Radiation Models for Regions with only Surface Meteorological Measurements in China. *Energy Conversion and Management*, 225, 113447.

[24] Govindasamy, T.R. and Chetty, N. (2018) Quantifying the Global Solar Radiation Received in Pietermaritzburg, Kwazulu-Natal to Motivate the Consumption of Solar Technologies. Open Phys., 16, 786-794.

[25] Brif, M. and Chikh, M. (2000) Estimation de l'Irradiation Solaire par la Logique Floue. *Rev.Energ. Ren.* : Chemss, 105-110.

[26] Khen, M.M. and Ahmad, M.J. (2012) Estimation of Global Solar Radiation Using Clear Sky Radiation in Yemen. *Journal of Engineering Science and Technology Review*, 5 (2), 12-19.

[27] Zhang, Y. (2006) Comparison of Measured and Estimated Monthly Average Daily Global Radiation in Shanghai, Nanjing and Hangzhou Areas in East China. Proceedings of the IXth World Renewable Energy Congress (WREC IX), 19-25 August, Florence (Italy), Paper ID: RTPC26 (CD-ROM).

[28] Almorox, J. and Hontoria, C. (2004) Global Solar Radiation Estimation Using Sunshine Duration in Spain. *Energy conversion and Management*, 45 (9-10).

[29] Gana, N.N. and Akpootu, D.O (2013) Estimation of Global Solar Radiation Using Four Sunshine Based Models in Kebbi, North-Western, Nigeria. *Advances in Applied Science Research,* 4 (5), 409-421.

[30] Er, Z., Rouabah, Z. and Turna, I.B. (2017) A Modal Study on the Estimation of Daily Global Solar Radiation from Sunshine Duration. *Acta Physica Polonica A*, 132 (3-II), 1070-1075.

[31]. Lan, K., Wang, L., Zhou, Y., Zhang, Z., Fang, S. and Cao, P. (2023) The Applicability of Sunshine-Based Global Solar Radiation Models Modified with Meteorological Factors for Different Climate Zones of China. *Frontiers in Energy Research*, 10: 1010745.

doi : 10.3389/fenrg.2022.1010745.

[32] Kotu, T.B., Ancha,V.R., Abulkadir, A., Hassen, A. A. and Fanta, S.W. (2023) Prediction of Global Solar Radiation Based on Sunshine Duration Hours in Ethiopia's Amhara Region. In : Woldegiorgis, B.H., Mequanint, K., Getie, M.Z., Mulat, E.G. and Alemayehu Assegie, A., Eds., *Advancement of Science and Technology, Green Energy and Technology, Springer,* Cham.

https://doi.org/10.1007/978-3-031-33610-1

[33] Bashahu, M. and Mpanzimana, D. (2009) Estimation of the Monthly Average Daily Global Solar Irradiation with Climatological Parameters at some Burundian Stations.

International Review of Physics, 3 (4), 237-243.

[34] Bristow, K.L. and Campbell, G.S. (1994) On the Relationship Between Incoming Solar Radiation and Daily Maximum and Minimum Temperature. *Agricultural and Forest Meteorology,* 31 (2), 159-166.

[35] Grossi Gallegos, H., Righini, R. and Raichijk, C. (2006) Analysis of Alternative for the Assessment of the Solar Resource in Argentina. Proceedings of the IX th World Renewable Energy Congress (WREC IX), 19-25 August, Florence (Italy), Paper ID: RT7 (CD-ROM).

[36] Fan, J., Chen, B., Wu, L., Zhang, F., Lu, X. and Xiang, Y. (2018) Evaluation and Development of Temperature-Based Empirical Models for Estimating Daily Global Solar Radiation in Humid Regions. *Energy*, 144, 903-914.

[37] Chen, R., Ersi, K., Yang, J., Lu, S., and Zhao, W. (2004) Validation of Five Global Radiation Models with Measured Daily Data in China. *Energy Conversion and Management*, 45 (11-12), 1759-1769.

[38] Akpootu, D.O. et al. (2023) Sunshine and Temperature Based Models for Estimating Global Solar Radiation in Maiduguri, Nigeria. *Saudi Journal of Engineering and Technology,* 8 (5), 82-90.

Doi : 10.36348/sjet. 2023.v08i05.001

[39] Reddy, M. V., Suvarana, R.P., Kumari, S.P. and Copal, K.R. (2018) Estimation of Global Solar Radiation Using Sunshine and Temperature Based Models over Anantapur Region. *International Journal of Creative Research Thoughts (IJCRT),* 6 (2), 607-618.

[40] Ohunakin, O.S., Adaramola, M.S., Ogewola, O.M. and Fagenle, R.O. (2013) Correlations for Estimating Solar Radiation Using Sunshine Hours and Temperature Measurement in Osogbo, Osun State, Nigeria. Front. Energ. , 7 (2), 214-222.

https://doi.org/10.1007/s11708-013-0241-2

[41] Mohamed, N. (2010) Etude Comparative des Méthodes d'Estimation du Rayonnement Solaire. Mémoire de Magister, Université Ferhat Abbas, Sétif.

[42] Al-Salihi, A.M, Kadun, M.M. and Mohammed, A.J. (2010) Estimation of Global Solar Radiation on Horizontal Surface Using Routine Meteorological Measurements for Different Cities in Iran. *Asian Journal of Scientific Research,* 3 (4), 240-248.

[43] Fan, J., Wang, X., Wu, L., Lu, X. and Xiang, Y. (2018) New Combined Models for Estimating Daily Global Solar Radiation Based on Sunshine Duration in Humid Regions: A Case Study in South China. *Energy Conversion and Management*, 156, 618-625.

[44] Sarkar, Md.N.I. and Sifat, A.I. (2016) Global Solar Radiation Estimation from Commonly Available Meteorological Data for Bangladesh. Renewables: Wind, Water and Solar, 3, Article number : 6.

[45] Ettoumi, F. Y., Mefti, A., Adane, A. and Bouroubi, M.Y. (2002) Statistical Analysis of Solar Measurements in Algeria Using Beta Distributions. Renewable Energy, 26, 47-67

[46] Pashiardis, S., Palengaris, A. and Kalogurou, S.A. (2023) Geographical Distribution of Global Radiation and Sunshine Duration over the Island of Cyprus. *Applied Sciences*, 13 (9), 5422.

https:// doi.org/10.3390/app13095422

[47] Luis, M.A., Felipe, D.R. and Pilar, N. R. (2008) Estimation of Global Solar Radiation by Means of Sunshine Duration. In: Goswami, D.Y. and Zhao, Y., Eds., Proceedings of ISES World Congress 2007 (vol. I-vol. II). Springer, Berlin, Heidelberg.

https://doi.org/10.1007/978-3-540-75957-3-530

[48] Anis, Md.S., Jamil, B., Ansari, Md.A. and Bellos, E (2019) Generalized Models for Estimation of Global Solar Radiation Based on Sunshine Duration and Detailed Comparison with the Existing: A case Study for India. *Sustainable Energy Technologies and Assessments*, 31,179-198.

https://www.researchgate.net/profile/Basharat-Jamil-2/...

[49] Carnegie Mellon University, Chapter 9: Simple Linear Regression (28 pages).

https ://www.stat.cmu.edu>c…

[50] Montgomery, D.C., Peck, E.A. and Vining, G.G. (2012) Introduction to Linear Regression Analysis, Fifth Edition. LCWU Online Directory, Wiley (872 pages).

https://ocd.lcwu.edu.pk>cfiles

[51] Luhunga, P., Boltai, J., and Kahimba, F. (2016). Evaluation of the performance of CORDEX Regional Climate Models in Simulating Present Climate Conditions of Tanzania. *Journal of Southern Hemisphere Earth Systems Science, 66* (1), 32-54.

[52] Betti, T., Zulim, I., Brkic, S., and Tuka, B. (2020) A Comparison of Models for Estimating Solar Radiation from Sunshine Duration in Croatia. *International Journal of Photoenergy, 2020,* Article ID : 9605950.

https://doi.org/10.1155/2020/9605950

[53] Razika, I., Nabila, I., and Marouane, M. (2015, December). Estimation of Global Solar Radiation Using Sunshine Duration for M'sila Region (Algeria). In *2015 3rd International Renewable and Sustainable Energy Conference (IRSEC 2015)*, Marrakesh (Morocco), 1-6.

doi :10/1109/IRSEC.2015.7455004

[54] Poudyal, K. N., Bhattarai, B. K., Sapkota, B., and Kjeldstad, B. (2012) Estimation of Global Solar Radiation Using Sunshine Duration in Himalaya Region. Res. J. Chem. Sci., 2, 20-25.

[55] Bashahu, M., and Butisi, B. (2010) Statistical Analysis of Daily Solar global Irradiation Data using Beta Distributions at Some Burundian Stations. *International Journal of Physical Science, 2* (3), 35-41.

[56] El-Metwally, M. (2005) Sunshine and Global Solar Radiation Estimation at Different Sites in Egypt. *Journal of Atmospheric and Solar-Terrestrial Physics*, *67* (14), 1331-1342.

[57] Bashahu, M., and Ntirandekura, J. (2018) Statistical Analysis of Hourly Clearness Index and Diffuse Fraction Data Using Beta Probability Density Functions. *Modern Environmental Science and Engineering*, *4* (4), 350-357.

[58] Chubaka Bashige, A. et al. (2022) Estimation et Modélisation de l'Irradiation Solaire Globale Journalière Perçue sur un Plan Horizontal à Bukavu. *IOSR Journal of Applied Physics (IOSR-JAP)*, 14 (01), 13-24.

[59] Bashahu, M., and Buseke, M. (2016) Statistical Analysis of Hourly Wind Speed Data from Some Burundian Stations Using Beta Probability Density Functions. *Modern Environmental Science and Engineering*, *2* (11), 740-746.

[60] Benkaciali, S. and Gairaa, K. (2014) Modélisation de l'Irradiation solaire Globale Incidente sur un Plan Incliné. *Journal of Renewable Energies*, 17 (2), 245-252.

[61] Bensaha, A., Benkouider, F., and Bekkouche, S. M. A. (2017) Estimation du Rayonnement Solaire en Ciel Clair par des Modèles Empiriques: Application au Site de Ghardaïa (Algérie). In: *1st International Seminar on the Apport of the Simulation in Technological Innovation, Ghardaïa (Algeria)*, March 2017.

[62] Albarelo, T. (2016) Estimation de l'Irradiation Solaire sur le Plateau des Guyanes: Apport de la Télédétection Satellite. Doctoral Dissertation, Université de Guyanne.

[63] Honerkamp, J. (1999) Stochastic Dynamical System: Concepts, Numerical Methods, Data Analysis. VCH Publishers, New York, Weinheim, Cambridge.

[64] Critical Values of Student's t Distribution with ν Degrees of Freedom ($t_{1-\alpha'}$, ν).

https://www.itl.nist.gov/div898/handbook/eda/section3

[65] Tossi, P. (1989) Méthodes Statistiques. Economica, Paris.

[66] Harerimana, C. (2005) Contribution à la Détermination de la Turbidité Atmosphérique et à la Modélisation des Rayonnements Solaires Global et Diffus. Mémoire, Université du Burundi, Bujumbura.

[67] Paltineau, G., Mihailescu, I.F., Torica,V. and Albu, A.N. (2002) Correlation Between Sunshine Duration and Global Solar Radiation in South-Eastern Romania. *International Agrophysics, 16 (2), 139-145.*

[68] Muzathik, A.M. et al. (2011) Daily Global Solar Radiation Estimate Based on Sunshine Hours. *International Journal of Mechanical and Materials Engineering* (IJMME), 6 (1), 75-80.

[69] Uçkan, I. and Khudhur, K.M. (2022) Improving of Global Solar Radiation Forecast by Comparing Other Meteorological Parameter Models With Sunshine Duration Models. *Environmental Science and Pollution Research.*

https://doi.1007/s11356-022-18781-3