Document de travail · Vérification interne | v1.0 | 2026-11-20

Une référence pour prévoir les arrivées mensuelles de touristes en séjour à Sainte-Lucie

Persistance saisonnière et tendance estimée sur six années de test

Adrian Dunkley | info@caribbeanaiassociation.com

Résumé

Cette étude examine si une tendance estimée améliore la prévision des arrivées de touristes qui passent la nuit à Sainte-Lucie. À la fin de chaque année, nous avons estimé les douze mois suivants de deux façons. La première reprend les arrivées du même mois de l'année précédente. La seconde utilise une régression avec un effet propre à chaque mois et une tendance linéaire commune. Nous avons évalué 60 prévisions par méthode sur 2015–2019, puis examiné 2020 séparément. Avant 2020, l'erreur absolue moyenne mensuelle était de 1 867 arrivées pour la première méthode et de 1 885 pour la régression. En 2020 elle atteignait 24 666 et 24 443. La somme des mois diffère du total publié pour deux années. Nous avons conservé les observations mensuelles et précisons les deux écarts. Ce test offre une référence reproductible pour évaluer des modèles plus complexes.

Question de recherche

Une tendance estimée fait-elle mieux que le nombre observé un an plus tôt ? Comment les deux méthodes se comportent-elles en cas de rupture de la série ?

Données et méthodes

Nous avons extrait les douze valeurs mensuelles de chaque année de 2010 à 2020 dans le tableau des arrivées de touristes du Bureau central de statistique, mis à jour le 16 mai 2024 [1]. Sa série annuelle identifie ces visiteurs comme des touristes en séjour [2]. Nous avons conservé les zéros d'avril, mai et juin 2020. Les douze valeurs de 2021 sont absentes du tableau mensuel. Nous avons donc arrêté la série en 2020.

À la fin de chaque année de 2014 à 2019, nous avons produit douze prévisions. La première méthode reprend la valeur du même mois de l'année précédente. La seconde ajuste une régression par moindres carrés ordinaires à toutes les observations mensuelles depuis 2010 jusqu'à la date de prévision. Chaque mois possède sa propre constante. Tous ont une même tendance annuelle. Aucune méthode n'utilise d'observation de l'année à prévoir. L'évaluation respecte l'ordre chronologique des données [3].

Nous avons mesuré la valeur absolue de l'écart entre prévision et observation. Nous avons donné le même poids aux 60 mois de 2015–2019 et évalué séparément les douze mois de 2020. Le calcul d'une erreur relative mensuelle imposerait une division par zéro pour trois mois. Nous avons aussi comparé la somme des mois de chaque année avec le total indiqué dans le même tableau [1].

Résultats

En 2015–2019, l’erreur absolue moyenne était de 1 867 arrivées par mois pour la règle de l’année précédente et de 1 885 pour la régression, soit un écart de 18. La régression a fait mieux en 2016 et 2017, moins bien en 2015, 2018 et 2019. Les deux méthodes ont généralement sous-estimé les arrivées. L’erreur signée moyenne, soit la prévision moins l’observation, était respectivement de −1 426 et −1 295.

En 2020, l’erreur absolue moyenne par mois s’est élevée à 24 666 et 24 443 arrivées respectivement. Pour avril 2020, la règle de l’année précédente prévoyait 36 828 arrivées. Le Bureau en a enregistré zéro [1]. Aucun des deux modèles ne disposait d’une variable signalant une rupture soudaine des conditions de voyage.

En 2017, les mois totalisent 386 327 arrivées, soit 200 de plus que le total publié de 386 127. En 2019, ils totalisent 423 741, cinq de plus que les 423 736 annoncées. Les sommes concordent pour les neuf autres années. La série annuelle donne aussi 386 127 pour 2017 et 423 736 pour 2019 [2]. Les sources ne précisent aucun mois à corriger. Nous avons donc conservé les valeurs mensuelles.

Erreur absolue moyenne des prévisions mensuelles

arrivées par mois



Chaque barre présente l’erreur absolue moyenne des 60 mois de 2015–2019 ou des douze mois de 2020. Une valeur basse correspond à une erreur plus faible. Calculs d’après [1].

Discussion

Avant 2020, l’écart moyen entre méthodes est faible par rapport à leurs erreurs mensuelles. Cet échantillon ne suffit pas pour privilégier une méthode. Une équipe peut tester un modèle d’apprentissage automatique contre ces deux références pour les mêmes douze horizons mensuels. Des résultats par année montrent si le gain dépend d’une seule année atypique.

L’année 2020 montre les limites d’un modèle qui ne connaît que le calendrier. Une équipe disposant de données d’exploitation datées pourrait tester les horaires de vol et les restrictions de voyage annoncées. Elle devrait utiliser les informations disponibles au moment du pronostic. Des révisions ultérieures donneraient au modèle un avantage sur les planificateurs de l’époque.

Limites

Cette étude rétrospective porte sur une île, cinq années ordinaires et une année de rupture. Nous ne fournissons ni test de significativité ni intervalle de probabilité. Les erreurs de mois successifs peuvent être liées. Tous les horizons ont le même poids.

La source est postérieure aux dates de prévision. Des publications antérieures ont pu donner des valeurs différentes. Les deux écarts de somme limitent les comparaisons annuelles. Les chiffres ne désignent ni les établissements, ni les liaisons, ni les travailleurs touchés par un changement. Cette étude ne mesure l’effet causal d’aucune politique de voyage.

Reproduire l’analyse

Utiliser les douze lignes mensuelles. Réserver la ligne Total au contrôle des sommes. Pour chaque année cible de 2015 à 2020, ajuster le modèle sur les mois de janvier 2010 à décembre de l’année précédente.

Reprendre l’observation du même mois de l’année précédente pour la première prévision. Pour la seconde, ajuster une régression du nombre d’arrivées sur douze indicateurs mensuels et l’année, sans constante supplémentaire. Prévoir les douze mois cibles.

Calculer les erreurs absolues à partir des observations mensuelles publiées. Faire la moyenne sur 2015–2019 et séparément sur 2020. Pour le biais, conserver le signe de la prévision moins l’observation. Arrondir uniquement les résultats.

Tableau de données

Mois	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Janv.	26083	26993	25605	25899	27643	31541	30661	31649	32755	34546	33675
Févr.	27867	26142	28947	27853	30135	32083	31637	33402	35536	38166	39638
Mars	29580	29536	30885	33842	34538	34595	35069	35427	41741	41674	18086
Avr.	25984	29122	27399	27772	30757	31200	30104	34322	32555	36828	0
Mai	30349	24786	24257	26679	27676	29658	29813	31436	32277	34354	0
Juin	22993	22404	21151	24071	25268	26347	25855	31312	31318	34040	0
Juil.	34186	32835	29416	28428	32100	32481	33381	38291	37844	42778	1955
Août	29589	28429	27866	27536	28646	27822	29991	33400	34168	36342	3261
Sept.	17393	16844	16687	18391	18247	17468	18169	20249	18745	21608	4122
Oct.	20624	22431	22248	22385	22805	23745	24251	27967	25991	26923	7808
Nov.	14741	22536	23709	25167	26933	25821	25254	29962	30422	32913	7766
Déc.	26548	30346	28631	30603	33410	32147	33687	38910	41428	43569	14384
Total	305937	312404	306801	318626	338158	344908	347872	386127	394780	423736	130695

Références

[1] Central Statistical Office of Saint Lucia. Tourist Arrivals Monthly 2010 to 2023. Updated 16 May 2024.

[2] Central Statistical Office of Saint Lucia. Selected Visitor Statistics, 2012 to 2023. Stay-over arrivals series.

[3] Hyndman, R. J., and Athanasopoulos, G. Forecasting: Principles and Practice, third edition. Chapter 5, The forecaster’s toolbox.

Crédits des données

This product was adapted from the information of The Central Statistical Office of Saint Lucia, which is licensed under the Open Licence Agreement of The Central Statistical Office of Saint Lucia. Contributors: Saint Lucia Tourist Board, Saint Lucia Air and Sea Ports Authority, Central Statistical Office.