

The Generalised Weight Share Method Applied to tourism surveys as an alternative to border surveys

Anne Coutière: Director of the Brittany Regional Observatory

Jean-Claude Deville: Head of the Survey Statistics Laboratory – CREST/INSEE

A number of methodological issues arise when carrying out a tourism survey of [individual holidaymakers](#) in an open environment and over territory of a given size, such as a French region. The first issue is the lack of sample frame. Indirect sampling of [consumption sites](#) can be used to overcome this problem. The sample of consumption sites (visits to regional tourist attractions) enables us to constitute a sample of tourists when they engage in [such consumption](#), as long as the links between the [tourist attractions](#) and the presence of tourists have been clearly established. The second issue that comes up is the matter of assigning a probability of selection, or an estimated weight, to the units surveyed from the target population [by distinguishing tourists from other visitors, so as to know the probability of their visits compared to the total number of visits observed](#). The Generalised Weight Share Method (GWSM) provides a theoretical framework for achieving this objective. This method enables us to ensure rigorous management of large samples over periods of about one year. A project for the application of the GWSM to a regional tourism survey was presented to the Seventh International Forum on Tourism Statistics in Stockholm[5]. Since then, the survey was carried out in one of France's most touristic regions, Brittany. [A technical committee oversaw the implementation of the project and the field work was done by TNS SOFRES. The processing and statistical work was done by the Brittany Regional Tourism Observatory \(ORTB\).](#)

1- Theoretical framework of the Generalised Weight Share Method applied to tourism

It is no easy task to collect data on tourists visiting a middle-ranking region such as Brittany. The characterisation of tourism relies [primarily](#) on accommodation, [leisure activities](#) and travel. Basing a survey on travel would lead to a conventional border survey. This technique will fail because the borders are open and cannot be policed. Furthermore, this technique fails to take account of tourism by residents of the region, which is a common in Brittany. Nor can we use accommodation as a basis for surveys, since there are no lists of holiday homes and because many tourists stay with relatives. Even collective accommodation establishments, such as hotels and campgrounds, fail to provide a sound basis for surveys when the contact with tourists is made [through the accommodation managers \(negative test results\)](#).

Indirect sampling, or the Generalised Weight Share Method (GWSM, [8][7][6][3][2] in order of importance) consists of capturing the target population, meaning tourists, or tourist households, or visits/trips using certain activities and certain forms of consumption that we can observe, in theory, and that we can sample with sufficient accuracy, in practice. However, in this simplified presentation we shall not attempt to go into detail on this subject We accept that every "tourist" (or tourist household) uses at least one such service during the observation period, as a result of which the tourist has a quantifiable non-zero probability of being selected for the sample and surveyed.

The services selected for the purposes of the MORGOAT survey, after a great deal of trial and error, are:

- visits to sixteen famous “tourist attractions”, which are either natural sites (Cap Fréhel, Carnac), or “museums” (Océanopolis). These produced a total of 9,476 useable questionnaires.
- purchases in bakeries. These produced a total of 5,719 useable questionnaires from 24 bakeries.

Entre février et décembre 2005, plus de 17 000 questionnaires ont été recueillis sur les différents points d'enquête représentatifs et emblématiques du tourisme en Bretagne :

- des sites de loisirs regroupés en 3 catégories : sites naturels, sites familiaux, sites patrimoniaux historiques,
- des boulangeries situées dans des communes touristiques, afin d'interroger les touristes ne visitant pas de sites de loisirs,
- le péage de La Gravelle (entrée de l'autoroute A81 vers Paris).



Les sites enquêtés



Les boulangeries enquêtées



10

Source : MORGOAT enquête Tourisme 2005 - TNS SOFRES - Traitement et réalisation ORTB

- vehicles leaving Brittany and passing through the La Gravelle highway toll station. These produced a total of 2,021 useable questionnaires.

The first is aimed primarily at conventional tourist activity for people from outside the region, the second aims at capturing **intra-regional** tourism and tourists who stay with relatives or in holiday homes, the third is a border survey that concerns about 44% of the visitors. The latter is aimed at enhancing the consistency of the observation system.

A specific survey system has been used for each “service”. The systems are generally based on fairly common and accepted practices. In each “sub-survey”, the survey days are selected at random.

Each survey day for “tourist attractions” gives rise to an estimate of the number of tourist visits N and n surveys. In practice, the sample selection is based on a specific system for each attraction, depending on its particular characteristics (paid admission, access mode, etc.) We let n/N be the probability of selection of each questionnaire and N/n be its extrapolation weight. The final weight is the product of this number multiplied by the extrapolation weight of the survey day. This weight is valid for the extrapolation of the population of *visits to the attraction* (population S).

In the case of bakeries, a sample (primary units) of bakeries is selected with probabilities that are a function of their estimated tourist clientele. Then, for each bakery and each survey day, a system that is analogous to the one used for tourist attractions is set up. In this case, an extrapolation weight can be assigned to each questionnaire with regard to *the population of purchases made in bakeries by tourists*. As we shall explain in more formal terms below, application of the Generalised Weight Share Method means that each "tourist household or member of a household making purchases in bakeries" (population B) is assigned a weight by dividing the extrapolation weight by the number of purchases during the trip. This number is assumed to be accurately determined for each questionnaire.

The survey at the toll station on the highway leading out of the region is a standard operation that produces extrapolation weights for the population P concerned (*people passing through the toll station*).

In sum, each of the sub-surveys has its own relatively ordinary methodology, its own theoretical and practical problems, but they enable us to obtain data that can be extrapolated for a given tourist sub-population. The Generalised Weight Share Method enables us to obtain a weighting of tourists in the population of interest T by the combination R of the disjointed services S, B and P, which appear to be sampling strata in technical terms.

The theoretical basis of the GWSM consists in using the links between R and T. These are described in a chart where arrows associate a unit of R with a unit of T, if the service i of R was used by the unit j of T. Each arc in the chart receives a strictly positive "charge" c_{ij} , which is 1 by default. If s_R denotes the sample drawn from R, weighted by w_i , we apply the weights $w_j = \frac{\sum_{i \in s_R} c_{ij} w_i}{\sum_{i \in s_R} c_{ij}}$ to the sample of tourists s_T with $w_i=0$ if i is not in the sample. See the references for more detail about the properties of these weights.

In practical terms, this means that we should obtain in each questionnaire, no matter where it was collected, the list of the attractions visited, the number of purchases in bakeries and any vehicle driving through the La Gravelle highway toll station. We can then calculate the attraction, bakery or toll station weights depending on the type of survey used to contact the tourist responding to the questionnaire. The methodology is complex and takes the specific characteristics of each survey into account. It is explained fully in [10].

All of this methodology, which includes the use of auxiliary data deemed to be reliable, makes it possible to assign a rigorously calculated weight to each questionnaire. These weights are used to obtain estimated totals and averages that are theoretically unbiased. We can also calculate an approximate confidence interval for each of these quantities under certain assumptions.

2 – Prerequisites for implementation in the [Brittany Region](#)

The Brittany regional authority was unable to continue the border surveys used in the past. It has found this new methodology to be a satisfactory alternative that meets its objectives, which are **to estimate and describe the core tourist numbers and consumption by individual holidaymakers in the region**. This new methodology can be applied to many types of territories, provided...

- that they are not so small that they require nothing more than a conventional on-site survey, or so large that they raise problems of cost and complexity owing to the structural diversity of types of tourism,

- that they have a number of sites that attract large numbers of visitors, including tourists (and not day trippers) of all types from the region, the rest of France and other countries, that these sites are evenly distributed over the territory under observation and representative of the tourist attractions of the region, and not visited only by tourists making their first trip to the region.

- that they are touristic enough to ensure that the number of visitors to the survey sites meets the needs of the sampling design, that they are likely to attract tourists regardless of the purpose of their trip (taking a beach holiday, discovering the region, staying in a holiday home, etc.)

Even though it is not easy for us to be impartial judges of our own work, we feel that **the objectives have been met very satisfactorily**. The methodology represents a major advance in terms of reliability and the production of new results in the area of tourist numbers and consumption, compared to our past production using the border survey method, and the results usually produced elsewhere.

3 – Main lessons drawn from the survey

What the methodology enabled us to produce: “In addition to the descriptive and economic dimensions of the results obtained, this approach makes possible indirect production of new territorial and sectoral marketing results that are immediately operational.”

Some examples:

- In addition to producing the conventional descriptive statistical indicators of the profiles of tourists visiting the region, the **volume and accuracy** of the data made it possible to estimate **the sizes of fairly fine market segments**. This made it possible for the players in the region to evaluate the impact of their promotion and communication actions for the first time and even to measure market penetration rates.

TAILLE DES MARCHÉS EN BRETAGNE

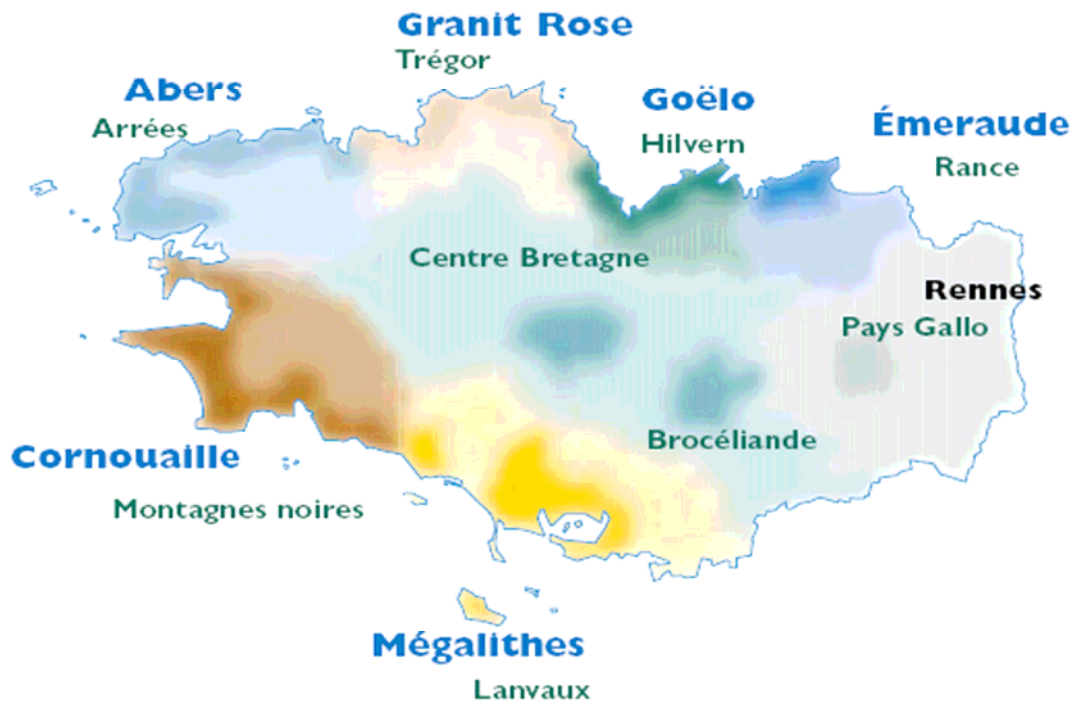
Ensemble de la clientèle

Volumes en millions	Au global			
	Séjours %	Séjours Volume	Touristes Volume	Nuitées Volume
Global	100	2,6	8,7	99
MODES D'HEBERGEMENTS				
Hébergement marchand	58,1	1,6	5,0	48,8
Camping	15,4	0,4	1,4	16,1
Hôtel	15,3	0,4	1,1	5,2
Meublé	10,9	0,3	1,2	13,4
Gîte rural	6,2	0,2	0,7	6,9
Chambre d'hôtes	4,2	0,1	0,3	1,5
Camping-car	3,9	0,1	0,3	1,3
Résidence de tourisme	3,4	0,1	0,3	3,3
Village de vacances	2,2	0,1	0,2	2,1
Hébergement non marchand	39,1	0,9	3,6	46,4
Résidence principale de parents et amis	13,9	0,4	1,2	9,1
Résidence secondaire de parents et amis	11,8	0,3	1,3	13,3
Résidence secondaire personnelle	11,0	0,3	1,1	18,0

- It is difficult to collect reliable data on consumption, especially using self-administered surveys during holiday trips. Nevertheless, this problem was factored into the survey design and the answers to consumption questions were far better than we imagined. They enabled us to produce consistent data and **estimate overall tourist consumption in the region**. However, the results were still fragile when broken down into expenditure items.

6. CONSOMMATION TOURISTIQUE ET DURÉE DE SÉJOUR	Indicateurs de la consommation touristique en 2005			
<p><i>La dépense moyenne par jour par personne en 2005 est globalement de 28,0 euros. Les touristes de proximité ne dépensent en moyenne que 19,3 euros par jour par personne, alors que les touristes étrangers consomment plus du double, soit 42,8 euros de dépense moyenne par jour par personne.</i></p> <p><i>On observe par ailleurs une durée de séjour pour les étrangers nettement supérieure à celle de la clientèle de proximité (+ 3 jours en moyenne).</i></p> <p><i>Au final, pour un budget moyen de séjour d'un foyer touristique de 1100 euros, le budget de la clientèle étrangère (1800 euros) est près de 3 fois supérieur à celui de la clientèle de proximité (650 euros).</i></p>		Global	Proximité	Étranger
	Dépense moyenne/jour/personne (en euros)	28,0	19,3	42,8
	Durée moyenne de séjour/personne (en jours)	11,5	9,5	12,5
	Taille moyenne du foyer touristique (en personnes)	3,5	3,5	3,4
	Budget moyen du séjour du foyer touristique (en euros)	1100	650	1800

- In addition, the survey enabled us to plot “**demand geography**” for the first time. This geography is **made up of tourist areas** based on the consumption sites (links between accommodation sites and tourist attractions) **at the regional level**, whereas, for many years, we worked with “supply geography”, using administrative entities that often had little to do with the geographical realities of tourist consumption. This is certainly one of the major indirect benefits of this method. It provides fabulous material for tourism developers from different regional entities.



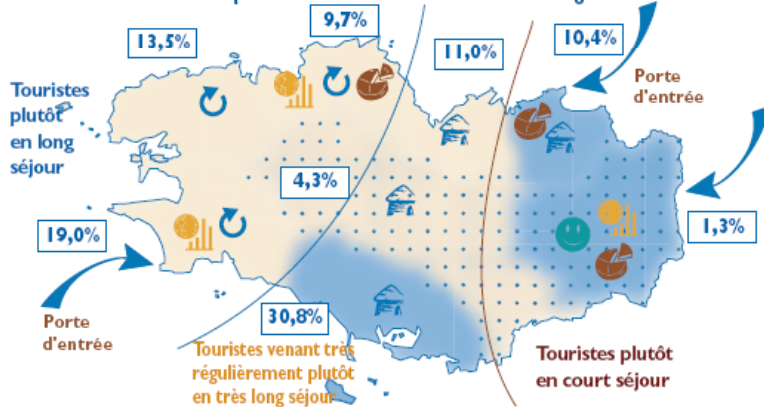
*En foncé, les coeurs de bassins organisés en pôles ou en rubans côtiers.
En clair, les arrières pays des bassins*

MORGOAT enquête Tourisme 2005 - TNS SOFRES - Traitement et réalisation ORTB

- We were able to estimate **intra-regional tourism volumes and profiles**. The scale of such tourism is probably underestimated since the people concerned did not always agree to respond to the survey, because they did not consider themselves to be tourists. This is bound to have biased the resulting profiles. On the other hand, we were able to make better estimates and descriptions of the profiles of the local tourist clientele (from Brittany and neighbouring regions), which is equivalent in terms of tourist nights to the foreign tourist clientele in Brittany. This local clientele is particularly attractive, since it could make it possible to attenuate the seasonal nature of tourist activity.

Synthèse

Rôles des bassins de fréquentation dans le tourisme en Bretagne en 2005



Rôles des bassins	Désaisonnalisée	Déspatialisée	Diversifie les origines	Ségmente la clientèle	Fidélise	Attire les 1 ^{er} séjours	Résidentialisée	Rajeunit
Émeraude - Rance	oui			CS, CO		oui	RS	
Goëlo - Hilvern							retraités	
Granit Rose - Trégor			oui	CO	oui			
Abers - Arrées					oui			
Comouaille - Montagnes Noires			oui		oui	oui		
Mégalithes - Larvaux	oui						retraités	
Centre Bretagne - Brocéliande		oui					Britanniques	
Rennes - Pays Gallo	oui	oui	oui	CS, CO		oui		oui

CS : Court séjour, CO : couple, RS : résidence secondaire

Source : MORGOAT enquête Tourisme 2005 - TNS SOFRES - Traitement et réalisation ORTB

6. SEGMENTATION

LÉGENDE :

- Fidélise
- Attire des 1^{er} séjours
- Résidentiale
- Diversifie géographiquement
- Rajeunit
- Segmente
- Désaisonnalisée
- Déspatialisée
- % Part des nuitées régionales des bassins

4 – Requirements for transferring such an approach to other types of territories and environments

What the methodology requires: “the methodologist steering the different stages of the survey, the clients and the partners concerned need be involved, responsive and cooperative as problems come up, and the solutions found must not create bias or entail excessive costs. These are the absolute requirements for the first implementation of the methodology”.

Such a project requires:

- **very detailed knowledge of the region’s tourism**, so as to choose the most appropriate tourist attractions for meeting the survey objectives. These sites will be the foundation for the survey and for the “mirror” for the tourist numbers obtained. Obviously, this net will fail to capture some very specific categories of tourists. The appropriate number of survey sites is ultimately a matter of balancing representativeness, the levels of detail at which the statistics are to be used and budget concerns.

- **tests**, which are essential and helpful for refining the method for the shift from theory to practice. Tests taught us many lessons and helped us to avoid many pitfalls.

- **strong partnerships with the entities responsible for the survey sites, which are not all in the public sector** (unlike border surveys in most cases). This required entering agreements with private-sector entities to receive **daily visitor counts without fail**, with the installation of visitor-counting systems, even at major natural attractions used as survey sites in exchange for producing visitor numbers for those sites.

- **having access to auxiliary data** to verify the consistency of the results obtained and to improve the extrapolation methods.

What we have learned from the first implementation of this methodology: “**transposition of the methodology cannot be automatic**, transposition calls for adjustments to adapt to the specific functional structure of tourism in the territory under observation”, especially when first implemented, and even as a matter of precaution when it is reproduced, but to a lesser degree.

In conclusion:

- The main advantage of this survey is **the accuracy and consistency of the statistical results produced**, as verified by cross-checking against auxiliary data.

- The main limitation is that **some forms of tourism are not properly captured**. The holes in the net are too big and the survey sites are not specific enough to tourism. Adding more sites (to include business travellers, yachters, seawater cure clients and sports tourists) would increase the complexity and cost of the survey in order to capture small market segments in very small areas with their own specific travel and consumption patterns.

- Such a survey is mainly an opportunity to diversify methods and improve observation systems. **It enables us to improve our ability to capture the reality of the region’s tourism using other tools**. Each methodology makes its own fundamental contribution and introduces its own significant distortions, which enable us to calibrate observation systems.

- The risk for this type of a survey lies in the **lack of methodological certification**, which penalises the conduct of real statistical surveys and benefits pseudo-statistical surveys, which can be done more rapidly and more cheaply. This means that the latter may prevail, given the current choices concerning communications budgets. The extreme lack of uniformity in statistical production leads to discordant results. This undermines the credibility of tourism statistics for both institutional users and the tourist industry by limiting the possibility and validity of comparisons between territories and, more especially, by hampering the economies of scale that could be achieved by dealing with certain observation issues at the appropriate territorial level.

The publications resulting from this statistical survey can be downloaded from the Brittany Regional Tourism Observatory website under the name **collection M.O.R.G.O.A.T enquête tourisme en Bretagne en 2005**.

References:

[1] Collection MORGOAT enquête tourisme en 2005 en Bretagne Partenariat Etat-Région, CDT et CCI de Bretagne ORTB. *These publications can be downloaded from the ORTB website or ordered on CDROM.*

[2] Deville, J.C., (1998). Comment attraper une population en se servant d’une autre. *Insee méthodes*, n°84-85-86, *Actes des Journées de méthodologie statistique des 17-18 mars 1998*, pp 63-82.

[3] Deville, J.C., and P. Lavallée (2006). Indirect Sampling: The Foundations of the Generalized Weight Share Method, *Techniques d’enquête*, December 2006 Vol. 32, No 2, pp. 185-196

[4] Deville, J.C., and M. Maumy (2006) Extensions de la méthode d’échantillonnage indirect et son application aux enquêtes dans le tourisme, *Survey Methodology*, December 2006, Vol. 32, No. 2, pp. 177-185

[5] Deville, J.C., and M. Maumy (2004) A new survey methodology for describing tourism activities and expanses, *Seventh International Forum on Tourism Statistics*

[6] Ernst, L. (1989). Weighting issues for longitudinal household and family estimates. In *Panel Surveys* (Kasprzyk, D., G. Duncan, G. Kalton, M.P. Singh, Editors), John Wiley and Sons, New York, pp. 135-159.

[7] Lavallée, P. (1995) « Pondération transversale des enquêtes longitudinales menées auprès des individus et des ménages à l'aide de la méthode du partage des poids », *Techniques d'enquête* vol. 21, pp.27-35.

[8] Lavallée, P. (2002) *Le sondage indirect, ou la méthode généralisée du partage des poids*, Editions de l'Université de Bruxelles, Editions Ellipses, Bruxelles.

[9] Torres Manzanera, E., I. Sustacha Melijosa, J.M. Menéndez Estébanez, L. Valdés Pelaàez (2002) « A solution to problems and disadvantages in statistical operations of surveys of visitors at accommodation establishments and at popular visitors places », Akos Probàld (Ed.), *Proceedings Of The Sixth International Forum On Tourism Statistics. Hungarian Central Statistical Office*, Budapest.

[10] Viscart, J. (2008) La petite histoire de MORGOAT enquête Tourisme, *Internal Document, Observatoire Régional du Tourisme de Bretagne*.