



Paper

# Controle implementatie BISI

L.L. Soldaat  
M.J.M. Poot

December 2022

## Verklaring van tekens

Niets (blanco)	Een cijfer kan op logische gronden niet voorkomen
.	Het cijfer is onbekend, onvoldoende betrouwbaar of geheim
*	Voorlopige cijfers
**	Nader voorlopige cijfers
-	(indien voorkomend tussen twee getallen) tot en met
0 (0,0)	Het cijfer is kleiner dan de helft van de gekozen eenheid
2021-2022	2021 tot en met 2022
2021/2022	Het gemiddelde over de jaren 2021 tot en met 2022
2021/'22	Oogstjaar, boekjaar, schooljaar enz., beginnend in 2021 en eindigend in 2022
2019/'20-2021/'22	Oogstjaar, enz., 2019/'20 tot en met 2021/'22

In geval van afronding kan het voorkomen dat het weergegeven totaal niet overeenstemt met de som van de getallen.  
Verbeterde cijfers in de staten en tabellen zijn niet als zodanig gekenmerkt.

## Colofon

### *Uitgever*

Centraal Bureau voor de Statistiek  
Henri Faasdreef 312, 2492 JP Den Haag  
[www.cbs.nl](http://www.cbs.nl)

### *Prepress*

Centraal Bureau voor de Statistiek

### *Ontwerp*

Edenspiekermann

### *Inlichtingen*

Tel. 088 570 70 70

Via contactformulier: [www.cbs.nl/infoservice](http://www.cbs.nl/infoservice)

© Centraal Bureau voor de Statistiek, Den Haag/Heerlen/Bonaire, 2022.  
Verveelvoudigen is toegestaan, mits het CBS als bron wordt vermeld.

## **Inhoudsopgave**

1.	Inleiding	4
2.	De formule voor de gewogen BISI	5
3.	Toetsing van de BISI-formule aan de hand van een voorbeeld	6
4.	Controle van de implementatie in Excel	7
	Referenties	8
	Bijlage 1. Variantiebenadering voor een gewogen meetkundig gemiddelde van indices	9
	Bijlage 2. Hoe om te gaan met abundanties boven de referentiewaarde	10

# 1. Inleiding

De BISI is een zogenaamde MSI, een Multi-Species Indicator. In een MSI worden de jaarcijfers van meerdere soorten gecombineerd tot een indicator waarvan de ontwikkeling in de tijd indicatief is voor de ontwikkeling van (de kwaliteit van) een ecosysteem. Om statistische en wiskundige redenen (Buckland et al. 2005, Van Strien et al. 2012) worden de jaarcijfers van de soorten doorgaans meetkundig gemiddeld (d.w.z. dat de logaritme van de jaarcijfers wordt gemiddeld, waarna het gemiddelde weer wordt terug-getransformeerd). Dat betekent in de praktijk dat een verdubbeling van een soort een even zwaar telt als de halvering van een andere soort.

MSI's worden veel toegepast in de internationale biodiversiteitsmonitoring en de methoden zijn goed gedocumenteerd (Gregory et al. 2005, Soldaat et al. 2017).

## BISI

De BISI heeft een ontwikkeltraject doorgemaakt waarbij het CBS heeft geadviseerd om abundantie van soorten meetkundig te middelen. De BISI is daarmee een MSI geworden, met enkele kleine aanpassingen t.o.v. de meeste MSI's. De meest kenmerkende eigenschappen van de BISI zijn dat de abundantie van elke soort relatief wordt uitgedrukt ten opzichte van een ecologische referentiewaarde en dat soorten onderling gewogen worden. De laatste versie van de BISI kent deze formulering:

$$BISI = \exp\left(\frac{\sum (IV_i * \ln(\frac{O_i}{R_i}))}{S}\right) \quad \text{formule 1}$$

In woorden: de geobserveerde abundantie van soort  $i$  ( $O_i$ ) in een jaar wordt gedeeld door de abundantie in een referentiesituatie ( $R_i$ ). Die referentie is een soort streefwaarde voor de abundantie in een gezonde ecologische situatie.  $O_i/R_i$  is daarmee een relatief jaarcijfer (index). Vervolgens worden de  $O_i/R_i$  – indexen van soorten meetkundig gemiddeld: de gesommeerde natuurlijke logaritmes van de indexen  $\sum \ln(O_i/R_i)$  worden gedeeld door het aantal soorten ( $S$ ), en daarna terug-getransformeerd door de exponent te nemen. Anders dan bij de meeste MSI's worden de indexen in de BISI echter *gewogen* opgeteld. Iedere soort heeft bij de sommatie een gewicht  $IV_i$ , dat waarden tussen 0 en 1 kan aannemen. Deze gewichten zijn expert-schattingen. Als de gewichten voor iedere soort op 1 worden gesteld, wordt er niet gewogen. Hier dient wel opgemerkt te worden dat de  $IV_i$ -waarden die in het excelbestand worden gebruikt niet de oorspronkelijke  $IV_i$ -waarden zijn uit kolom R, maar de  $IV_i$ -waarden gedeeld door de gemiddelde  $IV$ -waarde van alle soorten.

## Methodologische inspectie

In deze notitie staan de bevindingen van een methodologische inspectie van de BISI. Het CBS heeft daarbij de volgende acties uitgevoerd:

1. Het wiskundig formuleren van de gewogen BISI en de bijbehorende variantie / standaarddeviatie.
2. Het toetsen van de formule aan de hand van een voorbeeld.
3. Het controleren van de juiste implementatie van de berekeningen in het door Ecoauthor geleverde excelbestand.

## 2. De formule voor de gewogen BISI

Voor het bepalen van de gewogen variant van de BISI hebben we het advies ingewonnen van de afdeling methodologie van het CBS. De theoretische afleiding is te vinden in Bijlage 1. In onderstaande formule is de BISI geschreven in de termen zoals die gebruikt worden in het Excel-bestand waarmee de BISI nu berekend wordt:

$$BISI = \exp\left(\frac{\sum\left(IV_i * \ln\left(\frac{O_i}{R_i}\right)\right)}{\sum(IV_i)}\right) \quad \text{formule 2}$$

In woorden: per soort (i) wordt de natuurlijke logaritme van de index ( $O_i/R_i$ ) vermenigvuldigd met de weegfactor van de soort ( $IV_i$ ). De som van deze factoren over alle soorten heen wordt gedeeld door de som van de weegfactoren. De exponent van de uitkomst is de BISI. Als alle weegfactoren 1 zouden zijn, wordt de som van de weegfactoren ( $\sum(IV_i)$ ) gelijk aan het aantal soorten (S) en kom je weer op de bekende formule voor een klassieke MSI, met dit verschil dat R in de klassieke MSI de referentie van een basisjaar is, maar in het geval van de BISI is R de referentiewaarde van een (gunstige) ecologische streefsituatie:

$$BISI = \exp\left(\frac{\sum\left(\ln\left(\frac{O_i}{R_i}\right)\right)}{S}\right) \quad \text{formule 3}$$

Voor de eenvoud (ook makkelijker te implementeren in excel of in R) is het handig om de gewichten te normaliseren zodat de som van de gewichten in de noemer van formule 2 gelijk wordt aan 1. De formule wordt dan:

$$BISI = \exp\left(\sum\left(IV_i * \ln\left(\frac{O_i}{R_i}\right)\right)\right) \quad \text{formule 4}$$

De bij deze formule horende variantie is:

$$VAR_{BISI} = BISI^2 * \sum\left(IV_i^2 * \left(\frac{R_i}{O_i}\right)^2 * VAR_i\right) \quad \text{formule 5}$$

Waarin  $VAR_i$  de variantie (het kwadraat van de standaarddeviatie) van de index ( $O_i/R_i$ ) van soort i is (Let op: de breuk  $R_i/O_i$  is de inverse van de index. Dat komt door de notatie:  $IV_i$  moet gedeeld worden door de index).

De wortel uit  $VAR_{BISI}$  is de standaarddeviatie van de BISI.

### 3. Toetsing van de BISI-formule aan de hand van een voorbeeld

Om de werking van de formule voor de gewogen-BISI te controleren en te demonstreren, is gewerkt met een testbestand met twee soorten. Soort A heeft een weegfactor 0,5 en soort B een weegfactor 1. De weegfactoren zijn genormaliseerd tot 0,33 en 0,67 (som = 1). Beide soorten zitten in jaar 1 op de referentiedichtheid, respectievelijk 10 en 20. Dat levert, zoals gewenst een BISI van 1,00 op. In jaar 2 is soort A verdubbeld en soort B gehalveerd. Bij een ongewogen BISI zouden deze twee elkaar compenseren en zou de BISI weer op 1,00 uitkomen, maar door de zwaardere weging van soort B telt de halvering nu zwaarder dan de verdubbeling van soort A. De BISI wordt 0,79.

	Ri	IVi	IVi_norm	Oi		IVi*ln(Oi/Ri)	
				Jaar 1	Jaar 2	Jaar 1	Jaar 2
soort A	10	0,5	0,33	10	20	0,00	0,23
soort B	20	1,0	0,67	20	10	0,00	-0,46
BISI (gewogen)						1,00	0,79

Ter controle van deze waarde is de BISI berekend door soort B twee keer mee te nemen (omdat de weegfactor twee keer zo groot is als voor soort A), en vervolgens een ongewogen BISI ( $IV_i = 1$  voor beide soorten) te bepalen:

	Ri	IVi	IVi_norm	Oi		IVi*ln(Oi/Ri)	
				Jaar 1	Jaar 2	Jaar 1	Jaar 2
soort A	10	1,0		10	20	0,00	0,69
soort B	20	1,0		20	10	0,00	-0,69
soort B	20	1,0		20	10	0,00	-0,69
BISI (ongewogen)						1,00	0,79

Een tweede controle is uitgevoerd door beide soorten te laten verdubbelen. Dit zou moeten leiden tot een verdubbeling van de BISI, waarmee wordt voldaan aan de wenselijke eigenschap *proportionality* van een MSI (Van Strien et al. 2011). Inderdaad verdubbelt de BISI wanneer alle soorten verdubbelen:

	Ri	IVi	IVi_norm	Oi		IVi*ln(Oi/Ri)	
				Jaar 1	Jaar 2	Jaar 1	Jaar 2
soort A	10	0,5	0,33	10	20	0,00	0,23
soort B	20	1,0	0,67	20	40	0,00	0,46
BISI (gewogen)						1,00	2,00

Ook leidt een halvering van beide soorten tot een halvering van de BISI. De grootte van weegfactoren maakt hierbij niets uit (getest, maar resultaat niet getoond).

## 4. Controle van de implementatie in Excel

De implementatie van de berekening is getest aan de hand van de “General quality” – indicator in het door Ecoauthor geleverde excel-bestand. Niet alle details in het excelbestand zijn gecontroleerd, vanwege de complexiteit van de berekening door de vele verwijzingen naar andere cellen. De belangrijkste controle betreft het vergelijken van de BISI-berekening conform bovenstaande formules met de resultaten uit het excelbestand.

### 3.1.

De berekening van de BISI in cel AB85 en de SD in cel AL85 worden correct uitgevoerd. Wel adviseren we de berekening te vereenvoudigen conform formule 4 en 5, wat de inzichtelijkheid van de berekeningen vergroot en de kans op fouten verkleint. De correcte berekening van de BISI houdt stand wanneer de weegfactoren in kolom R (die voor de “General quality”- indicator voor iedere soort op 1 staan) met de hand worden veranderd.

### 3.2.

In cel P34 en P35 is in de berekening van  $O_i/R_i$  voor *Echinocardium flavescens* en *Echinocyamus pusillus* een verschuiving opgetreden van één rij, waardoor de BISI een beetje en de SD flink verandert. Dit illustreert het risico van berekeningen met Excel. Het CBS adviseert de berekeningen te automatiseren in R. De correcte berekening van de SD houdt stand wanneer de weegfactoren in kolom R (die voor de “General quality”- indicator voor iedere soort op 1 staan) met de hand worden veranderd.

### 3.3

Afkappen:

1. De twee ‘afkapformules’ in kolom P en Q voor de index  $O_i/R_i$  en de SD kappen indexen tussen 0 en 0,01 af, maar bijbehorende SD’s niet. Dat is bijvoorbeeld het geval voor *Donax vittatus*. Logischer is om de SD te herschalen naar een SD die hoort bij de index van 0,01.
2. Indexen van 0 worden afgekapt op 0,01 en de SD van de index wordt ook op 0,01 gezet. Logischer is het om de SD op 0 te zetten, omdat een niet-aanwezige soort altijd  $SD=0$  heeft, ongeacht de afkapwaarde.

### 3.4

De vermenigvuldiging met het aantal monsterpunten bij de berekening van de standaarddeviatie, zoals in een vorige versie van het excelbestand, is niet correct. De bedoeling hierachter was om de standaarddeviaties van soorten die op meer monsterpunten zijn gemeten een zwaarder gewicht te geven. Maar het effect van een betere bemonstering zit indirect al in de standaarddeviatie van de soort: die is lager naarmate de bemonstering beter is.

## Referenties

- Buckland, S.T., A.E. Magurran, R.E. Green & R.M. Fewster. 2005. Monitoring change in biodiversity through composite indices. *Phil. Trans. R. Soc. B* 360: 243-254.
- Gregory, R.D., A. van Strien, P. Voříšek, A.W. Gmelig-Meyling, D.G. Noble, R.P.B. Foppen, D.W. Gibbons 2005. Developing indicators for European birds. *Philos. T. Roy. Soc. B*, 360: 269-288.
- Soldaat, L.L., J. Pannekoek, R.J.T. Verweij, C.A.M. van Turnhout & A.J. van Strien. 2017. A Monte Carlo method to account for sampling error in multi-species indicators. *Ecological Indicators* 81: 340-347.
- Strien, A.J. van, L.L. Soldaat & R.D. Gregory. 2012. Desirable mathematical properties of indicators for biodiversity change. *Ecological Indicators* 14: 202-208.



# Bijlage 1. Variantiebenadering voor een gewogen meetkundig gemiddelde van indices

Sander Scholtus, CBS, 2021

Gegeven zijn  $T$  waargenomen indices  $I_1, \dots, I_T$  met bijbehorende gewichten  $w_1, \dots, w_T$ . Het gewogen meetkundig gemiddelde van deze indices is te schrijven als

$$\bar{I}^{(w)} = \sqrt[\sum_{t=1}^T w_t]{\prod_{t=1}^T I_t^{w_t}} = \left( \prod_{t=1}^T I_t^{w_t} \right)^{1/\sum_{t=1}^T w_t} = \exp\left(\frac{\sum_{t=1}^T w_t \ln I_t}{\sum_{t=1}^T w_t}\right).$$

We gaan ervan uit dat elke index  $I_t$ , onafhankelijk van de andere indices, is gemeten met een bepaalde onzekerheid, uitgedrukt in zijn variantie  $\text{var}(I_t)$ . De gewichten  $w_1, \dots, w_T$  beschouwen we als vast. Definieer kortheidshalve  $W = \sum_{t=1}^T w_t$ , zodat  $\bar{I}^{(w)} = \exp\left(\frac{1}{W} \sum_{t=1}^T w_t \ln I_t\right)$ .

De index  $\bar{I}^{(w)}$  is een niet-lineaire functie van de afzonderlijke indices. De variantie van deze index kan worden benaderd door de variantie van zijn lineaire Taylorreeks-benadering:

$$\text{var}(\bar{I}^{(w)}) \approx \sum_{t=1}^T \left( \frac{\partial \bar{I}^{(w)}}{\partial I_t} \right)^2 \text{var}(I_t),$$

waarbij  $\partial \bar{I}^{(w)} / \partial I_t$  de partiële afgeleide is van  $\bar{I}^{(w)}$  naar de index  $I_t$ . Om deze variantie te schatten worden deze partiële afgeleiden geëvalueerd gebruikmakend van de waargenomen indices  $I_1, \dots, I_T$ . We vinden dan:

$$\frac{\partial \bar{I}^{(w)}}{\partial I_t} = \frac{\bar{I}^{(w)}}{W} \frac{w_t}{I_t}$$

en dus

$$\text{var}(\bar{I}^{(w)}) \approx \left( \frac{\bar{I}^{(w)}}{W} \right)^2 \sum_{t=1}^T \frac{w_t^2}{I_t^2} \text{var}(I_t).$$

In het speciale geval dat de gewichten zijn genormaliseerd zodat  $W = \sum_{t=1}^T w_t = 1$  vereenvoudigen deze formules tot:

$$\bar{I}^{(w)} = \exp\left(\sum_{t=1}^T w_t \ln I_t\right),$$

$$\text{var}(\bar{I}^{(w)}) \approx \left(\bar{I}^{(w)}\right)^2 \sum_{t=1}^T \frac{w_t^2}{I_t^2} \text{var}(I_t).$$

En in het speciale geval dat alle gewichten gelijk zijn aan 1 ( $w_1 = w_2 = \dots = w_T = 1$ , zodat  $W = \sum_{t=1}^T w_t = T$ ) gaan deze formules over in:

$$\bar{I}^{(w)} = \bar{I} = \exp\left(\frac{1}{T} \sum_{t=1}^T \ln I_t\right),$$

$$\text{var}(\bar{I}^{(w)}) = \text{var}(\bar{I}) \approx \left(\frac{\bar{I}}{T}\right)^2 \sum_{t=1}^T \frac{\text{var}(I_t)}{I_t^2}.$$

Voor dit speciale geval is deze variantiebenadering eerder afgeleid door Jeroen Pannekoek in de appendix van Gregory et al. (2005).

## Referentie

R.D. Gregory, A. van Strien, P. Vorisek, A.W. Gmelig Meyling, D.G. Noble, R.P.B. Foppen & D.W. Gibbons (2005), Developing Indicators for European Birds. *Philosophical Transactions of the Royal Society B* **360**, pp. 269–288.

## Bijlage 2. Hoe om te gaan met abundanties boven de referentiewaarde

Het werken met een referentiewaarde geeft de BISI een eigenschap die ecologisch gezien aandacht vraagt. Want wat gebeurt er wanneer de abundantie van een soort hoger wordt dan de referentie? De BISI kan dan hoger worden dan onder referentie-omstandigheden ( $BISI=1$ ), en de vraag is hoe de BISI dan geïnterpreteerd moet worden. In tabel 3 wordt dit gedemonstreerd aan de hand van een voorbeeld waarbij soort A in jaar 2 een abundantie van twee keer de referentiewaarde heeft, terwijl soort B is gehalveerd t.o.v. de referentiewaarde. Het netto-effect op de BISI is nul: de BISI blijft 1, terwijl er nog best wat te verbeteren valt, namelijk voor soort B.

Tabel 3. BISI bij abundanties boven de referentiewaarde.

	Ri	IVi	IVi_norm	Oi		IVi*ln(Oi/Ri)	
				Jaar 1	Jaar 2	Jaar 1	Jaar 2
soort A	10	1,0	0,50	10	20	0,00	0,35
soort B	20	1,0	0,50	20	10	0,00	-0,35
BISI (gewogen)						1,00	1,00

Óf hier iets mee gedaan moet worden en wát, hangt af van het doel van de BISI. De BISI is gericht op het beschrijven van de ecologische toestand van een ecosysteem of een gebied, en het evalueren van het effect van maatregelen. De referentiewaarden van soorten zijn daarbij een soort streefwaarde: als die bereikt is, is het voor die soort goed. De BISI vat het bereiken van de streefwaarde voor alle soorten samen. Vanuit dat oogpunt zou de abundantie  $O_i$  van iedere soort 'afgekapt' kunnen worden op de referentiewaarde wanneer de abundantie daar bovenuit komt. Dat is gedaan in tabel 4.

Tabel 4. De BISI uit tabel 3, maar nu met hogere abundanties dan de referentiewaarde afgekapt op de referentiewaarde (d.w.z. dat de 20 voor soort A in jaar 2 op referentiewaarde 10 wordt gezet).

AFGEKAPT							
	Ri	IVi	IVi_norm	Oi		IVi*ln(Oi/Ri)	
				Jaar 1	Jaar 2	Jaar 1	Jaar 2
soort A	10	1,0	0,50	10	20	0,00	0,00
soort B	20	1,0	0,50	20	10	0,00	-0,35
BISI (gewogen)						1,00	0,71

De BISI komt dan nooit boven de 1 en geeft een redelijk goed beeld van de afstand tot het doel. In de huidige aanpak wordt niet afgekapt op de referentiewaarde, maar op 0,01 en 100 om de invloed van extreme abundanties van een soort op de BISI te reduceren.

Hoewel er statistisch/wiskundig niets mis is met de huidige aanpak, geeft het CBS ter overweging om hoge abundanties af te kappen op de referentiewaarde om het bereiken van de doelen beter in het vizier te houden (in het voorbeeld van tabel 3 zou de BISI dan op 0,707 uitkomen). Lage abundanties (bijvoorbeeld  $<0,01$  t.o.v. van de referentie) kunnen afgekapt

worden om de invloed van extreme waarden te beperken. Dat geeft meteen ook een oplossing voor nulwaarden, die vanwege de  $\ln$ -transformatie moeten worden vervangen door een klein getal.