

Beslutsstödsverktyg för val av trafikprognosmodell

2021-07-01

Joakim Ahlberg mm, Ramboll

Gunnar Flötteröd, Swedish National Road and Transport Research Institute

Chengxi Liu, Swedish National Road and Transport Research Institute

Yusak Susilo, University of Natural Resources and Life Sciences, Austria

Introduktion och översikt

Detta projekt syftar till att stödja Trafikverket (TRV) i att bedöma framtida trafikprognosmodeller, möjligen aktivitetsbaserade, som en del av Trafikverkets verktygslåda för strategisk planering. Valet av modell innefattar så många osäkerheter (framtida TRV-analysbehov, datatillgänglighet, modellfunktioner och systemkonfigurationer) att en unik modell inte kan rekommenderas. Det valda tillvägagångssättet är således att tillhandahålla ett beslutsstödsverktyg som gör det möjligt för TRV att sortera relevant och tillgänglig information utifrån ett antal tolkningsbara parametrar.

Beslutsstödsverktyget är kodat i ett Excel-dokument, strukturerat i två huvuddelar. Strukturen motiveras av att den kompetens som krävs för att förutse framtida TRV-analysbehov och datatillgänglighet skiljer sig från den som krävs för att utvärdera lämpligheten av specifika modellfunktioner för att stödja dessa analyser, givet en viss mängd tillgängliga data. Beslutsstödsverktyget inkluderar därmed två typer av kalkylblad. Kalkylbladen INPUT/OUTPUT/OPTIONAL är anpassade för att användas av TRV och bedöma modelleringsbehovet på en icke-teknisk nivå. Endast gröna fält behöver fyllas i för detta ändamål. Relevanta resultat visas i gula fält. INTERNAL kalkylblad används för att specificera expertkunskap om förhållandet mellan applikationsbehov, modelleringsfunktioner och data. Dessa kalkylblads innehåll har fastställts under projektets gång och kan, bortsett från enstaka uppdateringar av experter inom domänen, anses vara fastställt när beslutsstödsverktyget används.

Logiken enligt vilken beslutsstödsverktyget utvärderar på, bygger på enkla men internt konsekventa kvantitativa argument. Detta kräver att begrepp som "betydelse" och "tillgänglighet" översätts till numeriska vikter, som sedan med Bayesiska argument, kombineras till kvantitativa värderingar om modellens lämplighet. När beslutsstödsverktyget används, är det tillräckligt att ha en klar förståelse för dess in- och utdata.

Modellbedömning

En tillämpning av beslutsstödsverktyget för modellbedömning kräver följande steg:

1. Ange TRV: s förväntade analysbehov i kalkylbladet INPUT: Analysis needs.
2. Ange TRV: s framtida datatillgänglighet i kalkylbladet INPUT: Data availability.
3. Utvärdering av relevanta modellfunktioner i kalkylbladet OUTPUT: Model assessment.
4. Jämför konkreta resultat för olika modeller i kalkylbladet OPTIONAL: Configuration comparison.

Nedanstående fyra avsnitt förklarar dessa steg i detalj. Detta följs av en teknisk specifikation av beslutsstödsverktyget, inklusive dess INTERNA kalkylblad för expertkonfiguration.

Kalkylblad INPUT: Analysis needs

Det här kalkylbladet gör det möjligt att specificera TRV: s framtida förväntade analysbehov. Detta är relevant eftersom olika analyser kan ha olika krav, vilket i sin tur bäst hanteras av olika trafikprognosmodeller. Att fylla i detta dokument kräver att man förutspår, eller antar olika hypoteser kring, TRV: s framtida transportpolitiska mandat.

De förväntade TRV-analysbehov är kategoriserade i fyra grupper (kolumn A), som var och en innehåller flera konkreta analyser (kolumn B). Dessa analyser får inte ändras men kan viktas av användaren. Vikter tilldelas i två steg: Först tilldelas en vikt till varje grupp ("per grupp"), sen tilldelas en vikt till varje analys inom gruppen ("inom grupp").

För att tolka vikterna kan man föreställa sig att en viss mängd resurser (t.ex. skattepengar) görs tillgängliga för att realisera alla förväntade TRV-analysbehov. Dessa resurser fördelas till analyserna enligt ansatta vikter. Detta görs enligt följande två steg.

- Först fördelas alla tillgängliga resurser över grupper enligt deras vikter per grupp (kolumn C). Endast förhållanden mellan vikter per grupp är viktiga. Till exempel att tilldela en vikt per grupp på 4 för "Minskat privat resande och externa effekter" och en vikt per grupp på 2 för "Tillgänglighets- och lokaliseringsplanering" innebär att dubbelt så mycket resurser finns tillgängliga för analyser i gruppen "Minskat privat resande och externa effekter" än för analyser i gruppen "Tillgänglighets- och lokaliseringsplanering".
- I steg 2 fördelas sedan resurserna till varje grupp över motsvarande analyser inom varje grupp. För detta används viktning "inom grupp" (kolumn D). Dessa påverkar inte resursfördelningen mellan olika grupper. Endast förhållanden mellan vikter inom gruppen är viktiga. I ett exempel för gruppen "Minskat privat resande och externa effekter" ges analysen "Behovsoptimerade kollektivtrafik..." vikten 3 och analysen "Omfördela kapacitet: ..." ges en vikt på 1. Det betyder att tre gånger så mycket resurser är tillgängliga för "Behovsoptimerad kollektivtrafik..." än för "Omfördela kapacitet: ...".

Även om de förväntade TRV-analysbehov valdes för att vara så heterogena som rimligt möjligt, kan eventuella tvetydigheter uppstå när ett konkret analysbehov matchar mer än en från den tillgängliga listan över de femton identifierade analyser. Dessa fall löses med en subjektiv viktning mellan de förväntade TRV-analysbehov.

Många analyser i detta arbetsblad kräver en kostnads-nyttanalyt (CBA), till exempel i gruppen "tillgänglighet / lokaliseringsplanering". Eftersom en CBA inte är en aktivitet i sig men kopplad till någon analysfråga, finns den inte som ett separat analysbehov. INTERN: Spatial scales och CBA-arbetsblad förklarar hur CBA tas hänsyn till.

Kalkylblad INPUT: Data availability

Det här kalkylbladet gör det möjligt att ange en uppskattning av framtida datatillgänglighet för TRV. Detta är relevant eftersom olika trafikprognosmodeller har olika databehov vilket innebär att tillgängligheten till data påverkar genomförbarheten för olika prognosmodeller. Att infoga data i kalkylbladet kräver en uppskattning av TRV:s framtida tillgång till relevant data. Eftersom datatillgänglighet till stor del är ett resultat av investeringar (pengar och arbetstid) i datainsamling, förberedelse och lagring, kräver detta i sin tur en uppskattning av TRV: s framtida vilja och förmåga att investera resurser i dessa aktiviteter.

Kalkylbladet kategoriserar data i två grupper (kolumn A); Utbud och Efterfrågan. Utbudsdata är främst relaterat till de fysiska delarna i transportsystemet. Efterfråge-data är främst kopplat till resebetende i transportsystemet. För varje datagrupp definieras tre datakvalitetsnivåer: LÅG, MEDIUM och HÖG (kolumn B). Datagrupperna är internt nästlade vilket innebär att tillgängligheten av data på nivå MEDIUM också innefattar data på nivå LÅG och tillgängligheten av data på nivå HÖG även innefattar LÅG och MEDIUM-data. Data med kvalitetsnivå LÅG antas alltid vara tillgängliga. Datagrupperna och datakvalitetsnivåerna ändras inte av användaren.

Användaren kan ange uppskattningar av hur sannolikt det är att data på en viss kvalitetsnivå är tillgängliga (kolumn C). Dessa "tillgänglighetssannolikheter" ska tolkas på följande sätt: De två datagrupperna (utbud eller efterfrågan) betraktas separat. Det procentuella antalet som tilldelats den låga datakvalitetsnivån representerar sannolikheten att åtminstone data på nivå LÅG finns tillgängliga; detta antas alltid vara 100 % (vilket innebär att detta nummer inte får ändras). Det procentuella antalet som tilldelas data på MEDIUM-nivå representerar sannolikheten att åtminstone LÅG och MEDIUM-kvalitetsdata är tillgänglig. Det procenttal som tilldelats HÖG datakvalitetsnivå representerar sannolikheten för att data på alla tre kvalitetsnivåer är tillgängliga.

Denna logik innebär att sekvensen av sannolikheter som tilldelats LÅG, MEDIUM och HÖG datakvalitet måste vara icke-ökande inom respektive grupp. En kontrollfunktionalitet ingår därför (kolumn D) som ger en varning om detta villkor inte är uppfyllt.

Kalkylblad OUTPUT: Model assessment

Detta kalkylblad presenterar en kvantitativ bedömning av olika modellfunktioners tillräcklighet och genomförbarhet, med avseende på indata från kalkylbladen Analysis needs och Data availability. Det görs via kompletterande kalkylblad som innehåller expertbedömningar och kombinerar tillgänglig information med hjälp av ett enkelt men internt statistiskt konsekvent ramverk. Avsnittet nedan förklarar den praktiska tolkningen av resultaten i kalkylbladet Model assessment.

En modellfunktion är synonymt med en modellegenskap. Kalkylbladet beaktar följande fyra modellfunktioner (kolumn A).

- Representation av efterfrågan på resor. Beskriver på vilka nivåer resor uppstår mellan start och målpunkter, och via vilka färdmedel dessa genomförs.
- Resekedjor. Beskriver hur episoder av resande mellan start och målpunkter kombineras.
- Dynamiskt resebeteende. Beskriver tidsdimensionen för resebeteendet, dvs val av start/sluttid och sekvensering av resor eller turer.
- Fordonsrepresentation och fysiska interaktioner. Beskriver resans fysiska verklighet, dvs hur resenärer/fordon rör sig och hur trängsel/förseningar uppstår.

En modellfunktion kan finnas på olika detaljeringsnivåer, här kallat funktionsnivåer. För varje modellfunktion räknas två eller flera funktionsnivåer i ordning efter ökande detaljeringsgrad (kolumn B). Tillgängligheten för en viss modellfunktion på en viss funktionsnivå innebär att den även är tillgänglig på alla lägre nivåer.

Sammantaget ger kalkylbladet, för alla funktionsnivåer och alla modellfunktioner, en uppskattning av

- i vilken utsträckning funktionsnivån är tillräcklig för att informera om de förväntade TRV-analysbehov (med beaktande av informationen i kalkylbladet Analysis needs),
- hur troligt att tillräckligt med data kan nås för att genomföra funktionsnivån (med hänsyn till informationen i kalkylbladet Data availability).

Tillräckligheten hos en modellegenskapsnivå (kolumnerna C till G) uttrycks som ett procentvärde mellan 0 och 100. Ett värde på X procent innebär att sannolikheten att den betraktade modellegenskapsnivån är tillräcklig för att stödja ett slumpmässigt valt framtida TRV-analysbehov är X procent. Sannolikheten för att välja ett analysbehov är proportionell mot dess vikter i kalkylbladet Analysis needs. Kalkylbladet presenterar (i kolumnerna C till F) tillräckliga värden differentierade per rumslig skala. Detta innebär att vart och ett av dessa endast tar hänsyn till framtida TRV-analysbehov i motsvarande skala. Ett övergripande tillräckligt värde anges också (kolumn G).

Genomförbarheten för en funktionsnivå (kolumn I), uttryckt som ett procentvärde mellan 0 - 100 och indikerar sannolikheten att denna funktionsnivå kan realiseras (dvs. att motsvarande modell kan byggas och kalibreras) baserat på antagandena i kalkylbladet Data availability.

Den kombinerade tillräckligheten och genomförbarheten för en funktionsnivå (kolumnerna K till O) representerar den kombinerade sannolikheten att en betraktad funktionsnivå är både tillräcklig och genomförbar. Den sista kolumnen (kolumn O) ger därefter ett totalt värde för varje funktionsnivå, där tillräcklighet över alla rumsliga skalor kombineras med genomförbarhet.

Kalkylblad OPTIONAL: Configuration comparison

En konfiguration är ett val på exakt en nivå för varje möjlig modellfunktion. Detta kalkylblad ger en enkel visualisering av skillnaden mellan två konfigurationer med avseende på tillräcklighet eller

genomförbarhet (eller båda dessa) av respektive funktionsnivå. All terminologi är identisk med den som används i kalkylbladet Model assessment.

För varje konfiguration väljs (i kolumnerna D och E) vilken nivå konfigurationen implementerar varje modellfunktion på. Exakt en nivå per konfiguration och modellfunktion måste väljas. Man väljer vidare (i kolumn I) exakt ett jämförelsekriterium: tillräcklighet, genomförbarhet eller båda.

Kalkylbladet visar sedan (i kolumner L och M), för varje konfiguration och modell, motsvarande värden för tillräcklighet (eller genomförbarhet, eller båda). Dessa siffror kopieras från motsvarande kolumner i arbetsbladet Model assessment. En grafisk presentation av samma siffror finns också i ett stapeldiagram (ovanpå kolumnerna N till V).

Detta kalkylblad ger inte någon ny information jämfört med kalkylbladet för Model assessment. Det är dock användbart om man vill jämföra olika konfigurationer (t.ex. som erbjuds i olika planeringsprogram) med ett minimum av in- och utdata. Kalkylbladet gör det relativt enkelt att testa igenom många konfigurationsvarianter genom att jämföra funktionsnivåer av intresse.

Domänkännedom

Domänkännedom omfattar kunskap som en ren användare av verktyget inte kan förväntas tillhandahålla. Avsnittet domänkännedom är indelat i tre kompletterande ("INTERNAL") kalkylblad:

- Kalkylbladet INTERNAL: Spatial scales and CBA identifierar vid vilka rumsliga skalor som förväntade TRV-analysbehov uppstår, och i vilken utsträckning dessa kräver en kostnadsnyttoanalys.
- Kalkylbladet INTERNAL: Model features identifierar vilka funktionsnivåer som krävs för att hantera olika förväntade TRV-analysbehov.
- Kalkylbladet INTERNAL: Data needs identifierar vilken data som behövs för att förverkliga de olika funktionsnivåerna.

Dessa kalkylblad innehåller dolda fält som används för slutsatsberäkningar. Dessa dolda fält kan ignoreras vid uppdatering av domänkännedomen.

Kalkylblad INTERNAL: Spacial scales and CBA

Detta kalkylblad upprepar (i kolumnerna A och C) de (grupper av) förväntade TRV-analysbehov som angetts i kalkylbladet Analysis needs. Kalkylbladet adderar en kod (i kolumn B) till varje aktivitet som behövs senare.

Fyra rumsliga skalor beaktas (kolumnerna H till K):

- Långdistans/(inter)nationell: Avser hela Sverige och även internationellt.
- Regionalt: Spänner över kommuner, län, upp till Sampers analysregioner.

- Urban, zon: Avser städer, med trafikanalyzoner som den mest detaljerade nivån för geografisk analys.
- Urban, subzon: Avser städer, inklusive geografisk analys inom trafikanalyzoner (t.ex. gång inom zon, hushållstilldelning till enskilda byggnader).

Användaren infogar för varje kombination av förväntat TRV-analysbehov och rumslig skala en uppskattning av hur sannolikt det är att den rumsliga skalan kan beaktas i den analysen. Detta uttrycks som ett procenttal, där 0 % betyder "den betraktade rumsliga skalan uppstår aldrig i den betraktade analysen" och 100 % betyder "den betraktade rumsliga skalan uppstår alltid i den betraktade analysen". De infogade värdena behöver dock inte summeras till 100 % över alla rumsliga skalor för varje förväntat TRV-analysbehov. Detta för att möjliggöra analyser som spänner över flera rumsliga skalor.

Användaren infogar (i kolumn S) för varje förväntat TRV-analysbehov en uppskattning av sannolikheten att (ett slumpmässigt valt projekt som bedriver) denna analys kräver en kostnadsnyttoanalys. Dessa siffror behandlas emellertid inte vidare i kalkylbladet, utan underhålls endast för eventuell framtida användning. Modellfunktionsnivåer har redan bedömts utifrån TRV:s förväntade analysbehov, varav vissa kräver en CBA, vilket gör en separat bedömning av deras lämplighet för CBA överflödigt; en modellfunktionsnivå som är otillräcklig för att ta itu med en CBA är också otillräcklig för att stödja de analyser som kräver en CBA.

För att exemplifiera hur CBA kommer in i bilden, beakta analysbehov A 3.2 (Analysera parkeringsefterfrågan och utbud: Parkeringskostnader). För att genomföra den CBA som är nödvändig för att möta detta behov krävs en tidsdimension inom dagen såvida inte parkeringsbehovet fördelas uniformt över hela dagen. Följaktligen kommer en statisk representation av efterfrågan knappast att räcka för att hantera detta analysbehov, medan en dynamisk representation kommer att göra det. Detta kan läsas ur raderna 16–18 (Dynamiskt resbeteende), kolumn M (analysbehov A 3.2): En modell utan dynamik adresserar endast 20% av detta analysbehov (rad 16), medan en förenklad resp. helt dynamisk modell adresserar 90% resp. 100% (rad 17 och 18).

De ursprungliga siffrorna i detta arbetsblad är resultatet av en diskussion mellan TRV (Leonid Engelson) och projektgruppen.

Kalkylblad INTERNAL: Model features

Detta kalkylblad utvärderar tillräckligheten för varje möjlig funktionsnivå (som definieras i kalkylbladet för modellbedömning) för att stödja respektive förväntat TRV-analysbehov. Den innehåller en tabell där modellfunktionerna (och deras nivåer) är sorterade i raderna och de förväntade TRV-analysbehov i kolumnerna. För läsbarhet identifieras analysbehov endast med sina koder, vilka definieras i kalkylbladet Spatial scales and CBA. Modellens funktionsnivåer ges också koder (i kolumn B) för senare användning.

För varje cell i denna tabell infogar användaren sannolikheten att funktionsnivån i den betraktade raden, med avseende på motsvarande modellfunktion, är tillräcklig för att stödja det förväntade

TRV-analysbehovet i den betraktade kolumnen. I denna kontext betraktas 0 % som "otillräckligt i samtliga fall" och 100 % betyder "alltid tillräckligt". Med tanke på ett förväntat TRV-analysbehov bör dessa sannolikheter därför öka med ökande nivåer per modellfunktion.

De ursprungliga siffrorna i detta arbetsblad är resultatet av en diskussion mellan TRV (Leonid Engelson) och projektgruppen.

Kalkylblad INTERNAL: Data needs

Detta kalkylblad bedömer tillräckligheten för en viss kombination av efterfråge- eller utbudsdata på olika nivåer (som definieras i kalkylbladet Data availability) för att implementera en viss funktionsnivå. Datanivåerna är sorterade i raderna i denna tabell, modellens funktionsnivåer är sorterade i dess kolumner. För läsbarhet identifieras funktionsnivåerna endast med sina koder (enligt definitionen i kalkylbladet Model features).

För varje cell i denna tabell infogar användaren sannolikheten (ett tal mellan 0% och 100%) att den betraktade funktionsnivån kan realiseras baserat på den specifika kombinationen av efterfråge-/utbudsdata. Begreppet sannolikhet är svårare att tolka här än i tidigare fall. Användning av andra värden än 0% eller 100% kan användas för att uttrycka möjligheten att åtminstone en ofullkomlig approximation av en viss funktionsnivå kan uppnås baserat på en viss datanivå.

Siffrorna som infogas i detta kalkylblad är endast resultatet av en preliminär bedömning av projektgruppen och kan behöva ses över av TRV.

Tekniska egenskaper

Nedanstående beskrivning inkluderar en något oprecis användning av matematisk terminologi och notation för att minska formalia.

Beakta kalkylbladet Analysis needs. Låt G vara en av de fyra analysgrupperna, och låt v_G vara vikten per grupp kopplad till grupp G . Låt w_A vara vikten inom gruppen kopplad till framtida TRV-analysbehov A . Uttrycket $G=G(A)$ uttrycker att analysen A tillhör grupp G . Sannolikheten $\Pr(A)$ för att en krona av skattepengar tilldelas den framtida TRV-analysen A beräknas enligt följande:

$$(1) \quad \Pr(A) = v_{G(A)} / \sum_G v_G \cdot w_A / \sum_{A': G(A')=G(A)} w_{A'}$$

Denna ekvation är kodad i kolumner D-F i kalkylbladet Spatial scales and CBA. Den första termen på höger sida är förhållandet mellan analys A 's per gruppvikt över summan av alla vikter per grupp. Det innebär att den representerar sannolikheten att en slumpmässigt vald krona används för att adressera ett analysbehov inom analys A 's grupp. Den andra termen på höger sida är förhållandet mellan analys A 's inom gruppvikt över summan av alla inom gruppvikter. Detta

innebär att den representerar sannolikheten att en slumpmässigt vald krona används för att adressera analysbehov A med tanke på att denna krona bara användas inom analys A's grupp.

Beakta kalkylbladet Spatial scales and CBA. Låt S vara en rumslig skala och låt x_{AS} vara sammanfallet av rumslig skala och analysbehov för analys A och skala S; detta är direkt tillgängligt från kalkylbladet. Sannolikheten $\Pr(S|A)$ för att rumslig skala S uppstår i analys A uppnås genom normalisering:

$$(2) \quad \Pr(S|A) = x_{AS} / \sum_S x_{AS}$$

Denna ekvation är kodad i kolumner N-Q, rader 3-17 i kalkylbladet Spatial scales and CBA. Från detta erhålls den totala sannolikheten $\Pr(S)$ för att rumslig skala S uppstår:

$$(3) \quad \Pr(S) = \sum_A \Pr(S|A) \cdot \Pr(A)$$

Denna ekvation är kodad i kolumner N-Q, rad 19 i kalkylbladet Spatial scales and CBA. På samma sätt, låt $(CBA|A)$ vara sannolikheten för att analys A kräver en CBA; detta är direkt tillgängligt från kalkylbladet. Den totala sannolikheten $(CBA|S)$ för att en CBA behövs i rumslig skala S approximeras sedan enligt följande (ett villkorligt antagande om oberoende indikeras av symbolen "≈"):

$$(4) \quad \Pr(CBA|S) = \frac{\Pr(CBA,S)}{\Pr(S)} = \frac{1}{\Pr(S)} \sum_A \Pr(CBA,S|A) \cdot \Pr(A) \approx \frac{1}{\Pr(S)} \sum_A \Pr(CBA|A) \cdot \Pr(S|A) \cdot \Pr(A)$$

$\Pr(A)$

Denna ekvation är kodad i kolumner N-Q, rad 20 i Spatial scales and CBA.

Beakta kalkylbladet för modellfunktioner. Låt $\Pr(F|A)$ vara sannolikheten att funktionsnivå F är tillräcklig för att ta itu med analysbehov A; detta är direkt tillgängligt från kalkylbladet. Sannolikheten $\Pr(F|S)$ för att funktionsnivå F är tillräcklig för analyser som uppstår i rumslig skala S härleds enligt följande (vilket åter innebär ett villkorligt antagande om oberoende):

$$(5) \quad \Pr(F|S) = \frac{\Pr(F,S)}{\Pr(S)} = \frac{1}{\Pr(S)} \sum_A \Pr(F,S|A) \cdot \Pr(A) \approx \frac{1}{\Pr(S)} \sum_A \Pr(F|A) \cdot \Pr(S|A) \cdot \Pr(A)$$

Denna ekvation är kodad i kolumnerna T-W, raderna 11-21 i kalkylbladet Model features. Därifrån kopieras det över till kalkylbladet för Model assessment. Sannolikheten $\Pr(F)$ för att funktionsnivån F är tillräcklig över alla rumsliga skalor är

$$(6) \quad \Pr(F) = \sum_S \Pr(F|S) \cdot \Pr(S)$$

Denna ekvation är kodad i kolumn X, rad 11-21 i kalkylbladet Model features och därifrån kopieras över till kalkylbladet Model assessment.

Beakta kalkylbladet Data availability. Låt D resp. SU vara den kvalitetsnivå vid vilken framtida efterfråge- resp. utbudsdata kommer att finnas tillgänglig. Tre konfigurationer beaktas för D (symmetriskt för SU):

- D_{100} innebär att endast data på LÅG nivå är tillgängliga och att inga mer detaljerade data finns tillgängliga.
- D_{110} betyder att data för LÅG och MEDIUM är tillgängliga och att inga mer detaljerade data finns tillgängliga.

- D_{111} betyder att data på LÅG, MEDIUM och HÖG nivå är tillgängliga.

Dessa tre konfigurationer hänvisar till överlappande mängder data, men som händelser är de inbördes unika. Det förefaller inte realistiskt att förvänta sig att en användare av detta kalkylblad specificerar datatillgänglighets sannolikheterna $\Pr(D_{100}), \Pr(D_{110}), \Pr(D_{111})$ eftersom detta skulle kräva en gemensam bedömning av sannolikheten för att data upp till en viss nivå är tillgängliga och av data på högre nivå inte tillgänglig. Det bedöms enklare att enbart ange sannolikheten för att data upp till en viss nivå är tillgängliga, oberoende av att mer detaljerad data är tillgänglig eller inte.

Användaren infogar följaktligen sannolikheterna $\Pr(D_{1**}), \Pr(D_{11*}), \Pr(D_{111})$ för data på åtminstone LÅG, MEDIUM och HÖG nivå. Följande relationer gäller:

- Sannolikheten $\Pr(D_{111})$ för att data på alla nivåer är tillgängliga kräver ingen ytterligare analys.
- Sannolikheten $\Pr(D_{11*})$ för att åtminstone data om LÅG och MEDIUM-nivå är tillgänglig ges av $\Pr(D_{11*}) = \Pr(D_{110}) + \Pr(D_{111})$.
- Sannolikheten $\Pr(D_{1**})$ för att åtminstone LÅG nivådata är tillgänglig ges av $\Pr(D_{1**}) = \Pr(D_{100}) + \Pr(D_{110}) + \Pr(D_{111})$.

Via omorganisering konstateras att

$$(7) \quad \Pr(D_{110}) = \Pr(D_{11*}) - \Pr(D_{111})$$

$$(8) \quad \Pr(D_{100}) = \Pr(D_{1**}) - \Pr(D_{110}) - \Pr(D_{111}) = \Pr(D_{1**}) - \Pr(D_{11*})$$

Dessa ekvationer är kodade i kolumnerna B, C och E, F i kalkylbladet Data needs.

Beakta kalkylbladet Data needs. Låt $\Pr(\text{FF}|D, \text{SU})$ vara sannolikheten för att funktionsnivå FF kan realiseras med tanke på efterfråge-data på nivå D och utbudsdata på nivå SU; detta är direkt tillgängligt från kalkylbladet. Den totala sannolikheten $\Pr(\text{FF})$ för att modellen har nivån FF kan beräknas enligt följande (gör ett antagande om oberoende):

$$(9) \quad \Pr(\text{FF}) \approx \sum_{D \in \{D_{100}, D_{110}, D_{111}\}} \sum_{\text{SU} \in \{\text{SU}_{100}, \text{SU}_{110}, \text{SU}_{111}\}} \Pr(\text{FF}|D, \text{SU}) \cdot \Pr(D) \cdot \Pr(\text{SU})$$

Denna ekvation är kodad i rad 13 i kalkylbladet Data needs och kopieras därifrån till Model assessment kalkylbladet. Eftersom det inte finns något separat fall där "ingen (inte ens på LÅG nivå) data är tillgänglig", måste det antas att LÅG nivådata alltid är tillgänglig, vilket betyder $\Pr(D_{1**}) = 1$, vilket också krävs i kalkylbladet Data availability.