सर्वेक्षणों के लिए ऑफ-साइट निगरानी प्रणाली (ओएमओएसवाईएस): गुणवत्ता आश्वासन के लिए एक भौगोलिक सूचना प्रणाली (जीआईएस) आधारित दृष्टिकोण

सुखबीर सिंह और विशाल मौर्य द्वारा ^

यह लेख व्यापक क्षेत्र सर्वेक्षणों में डेटा की गुणवत्ता सुनिश्चित करने के लिए सर्वेक्षणों के लिए एक नई ऑफ-साइट निगरानी प्रणाली (ओएमओएसवाईएस) का परिचय देता है। यह प्रणाली कंप्यूटर-सहायता प्राप्त व्यक्तिगत साक्षात्कार (सीएपीआई) उपकरणों से भौगोलिक सूचना प्रणाली (जीआईएस) डेटा का उपयोग करके संदिग्ध सर्वेक्षण प्रतिक्रियाओं की पहचान के लिए नए स्थानीय उपायों को नियोजित करती है। इस लेख में विकसित मॉडल संचालित संकेतक (एमडीआई) और निश्चित नियंत्रण संकेतक (एफ़सीआई) दृष्टिकोण बिना किसी मैनुअल हस्तक्षेप के लचीले तरीके से संदिग्ध मामलों की पहचान करने में मदद करते हैं। ओएमओएसवाईएस विभिन्न भौगोलिक क्षेत्रों में दक्षता को अधिकतम करने और सर्वेक्षण की गुणवत्ता बनाए रखने के लिए फील्ड विजिट्स की लक्षित ट्रैकिंग की सुविधा प्रदान करता है, जो समय-संवेदनशील और संसाधन-विवश सर्वेक्षणों के लिए महत्वपूर्ण हैं।

परिचय

सर्वेक्षण संगठित समाजों में एक महत्वपूर्ण भूमिका निभाते हैं जो निर्णय लेने और उसके बाद की निगरानी में इनपुट के रूप में सूचना पर निर्भर करते हैं। इस प्रकार सर्वेक्षण डेटा की सटीकता, विश्वसनीयता और समयबद्धता महत्वपूर्ण हो जाती है, क्योंकि विभिन्न नीति डोमेन में नीति-निर्माण प्रक्रिया में उनके दूरगामी निहितार्थ हैं। आर्थिक स्थितियों के कई क्षेत्रों में पूछताछ, व्यक्तिगत साक्षात्कार हालांकि अधिक संसाधन-गहन हैं, लेकिन लक्षित प्रतिक्रिया प्राप्त करने, मानवीय संपर्क के माध्यम से बेहतर समझ और उच्च प्रतिक्रिया दर के मामले में टेलीफ़ोनिक और ऑनलाइन सर्वेक्षणों की तुलना में अलग-अलग लाभ हैं। सर्वेक्षण के प्राथमिक डेटा की जाँच के अलावा, गुणवत्ता आश्वासन उपायों में अक्सर नमूना आधार पर सर्वेक्षण प्रतिक्रियाओं को मान्य करने के लिए गुणवत्ता नियंत्रकों/समन्वयकों द्वारा सर्वेक्षण किए गए स्थानों पर भौतिक अनुवर्ती दौरे शामिल होते हैं। ये ऑन-साइट दौरे न केवल जांचकर्ताओं द्वारा सर्वेक्षण के निर्देशों और डिज़ाइन के उचित अनुपालन को सुनिश्चित करते हैं, बल्कि पहले हाथ से समृद्ध अनुभव भी प्रदान करते हैं, जो जांचकर्ताओं के प्रशिक्षण को सुव्यवस्थित करने, सर्वेक्षण डिज़ाइन का अध्ययन करने और साथ ही सर्वेक्षण प्रश्नावली/स्पष्टीकरण को परिष्कृत करने में सहायक साबित होता है।

भारत जैसे विशाल और विविधतापूर्ण देश में, कठोर समय सीमा के साथ राष्ट्रव्यापी सर्वेक्षण करना सर्वेक्षण डिजाइन और कवरेज निर्देशों का जमीनी स्तर पर पालन सुनिश्चित करने के मामले में चुनौतीपूर्ण हो जाता है, क्योंकि जांचकर्ताओं द्वारा फील्ड विजिट की आवश्यकता वाला क्षेत्र बहुत बड़ा हो जाता है, और विशेष रूप से जांचकर्ताओं द्वारा निर्देशों के अनुपालन या उत्तरदाताओं की वास्तविकता के संबंध में समस्याएं उत्पन्न होती हैं। इससे प्रतिक्रिया सत्यापन के लिए फील्ड विजिट के दौरान स्थानों की पहचान करने की संभावना कम हो सकती है। यह चुनौती दूरदराज के क्षेत्रों में विशेष रूप से गंभीर होगी, जहां अनुवर्ती फील्ड सत्यापन की कम संभावना संभावित रूप से कुछ जांचकर्ताओं को लक्षित स्थान से दूर अधिक सुविधाजनक स्थानों पर सर्वेक्षण करने के लिए प्रेरित कर सकती है, जो इच्छित डिजाइन निर्देशों और सर्वेक्षण डेटा की वांछित गुणवत्ता से समझौता करेगा।

इन चुनौतियों का समाधान करने के लिए, सर्वेक्षण करने के लिए इस्तेमाल किए गए उपकरण द्वारा कैप्चर की गई उत्तरदाताओं की स्थान संबंधी जानकारी की क्षमता का लाभ उठाकर यहाँ अभिनव और व्यावहारिक तरीके विकसित किए गए हैं। भौगोलिक सूचना प्रणाली (जीआईएस) डेटा भू-निर्देशांक (अर्थात, अक्षांश और देशांतर) के रूप में सर्वेक्षणों के लिए एक ऑफ-साइट निगरानी प्रणाली (ओएमओसिस) विकसित करने के लिए उपयोग किया गया है, जो लगभग वास्तविक समय के आधार पर सर्वेक्षण निष्पादन को ट्रैक करता है। सांख्यिकीय विधियाँ विशेष रूप से इस उद्देश्य के लिए प्रक्रियाओं के साथ डिज़ाइन की गई हैं, जो इन विधियों को प्रभावी ढंग से लागू और संचालित करती हैं।

^ लेखक भारतीय रिज़र्व बैंक (आरबीआई) के सांख्यिकी और सूचना प्रबंधन विभाग (डीएसआईएम) के क्रमशः नई दिल्ली और चंडीगढ़ क्षेत्रीय कार्यालयों में सहायक परामर्शदाता हैं। वे डॉ. ओपीमॉल, कार्यपालक निदेशक, डॉ. एआर जोशी, प्रधान परामर्शदाता, डीएसआईएम, श्री रविशंकर, परामर्शदाता, डीएसआईएम और आरबीआई के 'वार्षिक सांख्यिकी सम्मेलन - 2023' के प्रतिभागियों के बहुमूल्य सुझावों के लिए आभारी हैं। यहाँ व्यक्त किए परामर्शदाता विचार लेखकों के हैं, आरबीआई के नहीं।

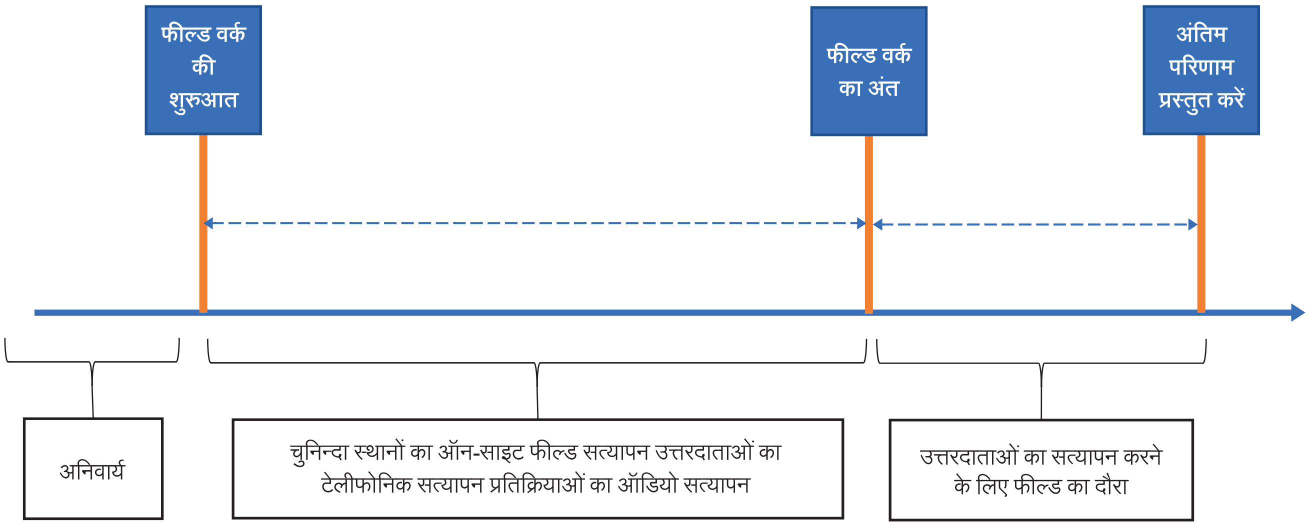
लेख में भारतीय रिजर्व बैंक (आरबीआई) द्वारा किए गए घरेलू सर्वेक्षणों में इसके उपयोग के उदाहरणों के साथ-साथ ओएमओएसवाईएस का विस्तृत विवरण दिया गया है। लेख को पाँच खंडों में संरचित किया गया है। परिचय के बाद, खंड II सर्वेक्षणों में नियोजित सामान्य डेटा गुणवत्ता आश्वासन (क्यूए) प्रक्रियाओं पर गहराई से चर्चा करता है और संबंधित चुनौतियों पर प्रकाश डालता है। खंड III जीआईएस सूचना का उपयोग करते हुए ओएमओएसवाईएस की व्यापक जानकारी प्रदान करता है, जो पहचाने गए मुद्दों का समाधान प्रस्तुत करता है। खंड IV वास्तविक मामलों और सिंथेटिक डेटा के माध्यम से ओएमओएसवाईएस के अनुप्रयोग को प्रदर्शित करता है, जो व्यावहारिक अंतर्दृष्टि प्रदान करता है। समापन खंड निष्कर्षों और सिफारिशों का सारांश प्रस्तुत करता है।

II. सर्वेक्षणों में गुणवत्ता आश्वासन प्रक्रिया

फील्ड सर्वेक्षण सबसे पहले सैंपलिंग फ्रेम डिजाइन करके डेटा इकट्ठा करते हैं, जहाँ पहले चरण की इकाइयों (एफ़एसयू) और दूसरे चरण की इकाइयों (एसएसयू) के चयन के लिए उपयुक्त सैंपलिंग योजना का उपयोग किया जाता है, और मजबूत डेटा गुणवत्ता आश्वासन (क्यूए) विधियों का उपयोग किया जाता है। उदाहरण के लिए, भारतीय रिजर्व बैंक (आरबीआई) द्वारा किए गए शहरी घरेलू सर्वेक्षणों के संदर्भ में, जैसे कि मुद्रास्फीति की उम्मीदें घरों का सर्वेक्षण (आईईएसएच) और उपभोक्ता विश्वास सर्वेक्षण (सीसीएस), मतदान केंद्र/वार्ड एफ़एसयू के रूप में कार्य करते हैं। ये एफ़एसयू बाद के घरेलू स्तर के जवाब प्राप्त करने का आधार बनते हैं, जहाँ पक्षपातपूर्ण परिणामों से बचने के लिए उत्तरदाताओं का यादृच्छिकीकरण महत्वपूर्ण है (आरबीआई, 2018, 2019)। एक-दूसरे के बहुत करीब से उत्तरदाताओं का चयन करने से अत्यधिक सहसंबद्ध प्रतिक्रियाएँ हो सकती हैं, जिससे संभावित रूप से पक्षपातपूर्ण और विषम परिणाम हो सकते हैं। इसके अतिरिक्त, यदि उत्तरदाता सैंपलिंग फ़ोकस क्षेत्र से दूर स्थित हैं, तो भौगोलिक प्रतिनिधित्व को विकृत करने की संभावना है। इस तरह की विकृति सर्वेक्षण प्रतिक्रियाओं को लक्षित क्षेत्र की विशेषताओं से जोड़ने के किसी भी बाद के प्रयास को प्रभावित कर सकती है। इस प्रकार, विभिन्न प्रश्नों के उत्तरों की अंतर-संगतता की निगरानी के अलावा, सर्वेक्षणों के लिए गुणवत्ता आश्वासन ढांचे में कार्यान्वयन पहलुओं को भी शामिल किया गया है, ताकि सर्वेक्षण के लिए नियुक्त अन्वेषकों द्वारा इकाइयों के चयन और इच्छित स्थानों पर वास्तविक दौरे के निर्देशों का अनुपालन सुनिश्चित किया जा सके।

चार्ट 1 सर्वेक्षणों के लिए एक व्यापक गुणवत्ता आश्वासन ढांचे को दर्शाता है, जिसमें जांचकर्ताओं को सर्वेक्षण-पूर्व प्रशिक्षण, चुनिंदा प्रतिक्रियाओं का ऑन-साइट/ऑडियो सत्यापन करना

चार्ट 1: सामान्य गुणवत्ता आश्वासन प्रक्रिया



स्रोत : लेखकों के स्पष्टीकरण

और नमूना आधार पर प्रतिक्रियाओं की प्रामाणिकता को मान्य करने के लिए सर्वेक्षण-पश्चात फ़ील्ड/टेलीफ़ोनिक सत्यापन लागू करना शामिल है। जबकि कंप्यूटर असिस्टेड पर्सनल इंटरव्यू (सीएपीआई) उपकरण के माध्यम से सर्वेक्षण के हिस्से के रूप में कैप्चर की गई ऑडियो रिकॉर्डिंग और संपर्क विवरण उत्तरदाताओं की प्रामाणिकता की जांच करने और जांचकर्ताओं की गुणवत्ता/दक्षता का मूल्यांकन करने में योगदान करते हैं, सर्वेक्षण के दौरान और उसके बाद फ़ील्ड विज़िट इस ढांचे के तहत आवश्यक घटक हैं (आरबीआई, 2009, 2010)।

सख्त समयसीमा के भीतर क्षेत्र सत्यापन के लिए सर्वेक्षण स्थानों की व्यापक कवरेज प्राप्त करना एक चुनौतीपूर्ण कार्य है, जिसमें पर्याप्त संख्या में स्थानों का दौरा करना, सर्वेक्षण डिजाइन विनिर्देशों का पालन करना और प्रतिक्रियाओं की गुणवत्ता बनाए रखना शामिल है। जब विशाल क्षेत्रों में एक साथ कई सर्वेक्षण किए जाते हैं तो जटिलता बढ़ जाती है।

सर्वेक्षण प्रक्रिया के संदर्भ में, सूचना प्रौद्योगिकी की उन्नति और समावेश के साथ हाल के वर्षों में डेटा संग्रह प्रक्रिया में तेजी से बदलाव आया है। पहले के पेन-एंड-पेपर इंटरव्यू (पीएपीआई) आधारित डेटा संग्रह विधियों को कई संस्थानों द्वारा सीएपीआई द्वारा प्रतिस्थापित किया गया है, जिसने डेटा संग्रह प्रयासों को कम कर दिया है और साथ ही डेटा की गुणवत्ता में सुधार करने में मदद की है (बेकर एट अल., 1995; कूपर, 2000; कैयर्स एट अल., 2010)। एक उप-उत्पाद के रूप में, सीएपीआई उत्तरदाताओं की स्थान संबंधी जानकारी और प्रतिक्रियाओं के समय टिकट भी प्रदान करता है, जो विश्लेषण और निगरानी उद्देश्यों के लिए मूल्यवान डेटा हैं। अगला खंड विशिष्ट सांख्यिकीय विधियों को प्रस्तुत करता है जो इस जानकारी को ऑफ-साइट निगरानी और यदि आवश्यक हो तो ऑन-कोर्स सुधार के लिए प्रभावी रूप से लाभ उठाने के लिए डिज़ाइन किए गए हैं।

III. सर्वेक्षण के लिए ऑफ-साइट मॉनिटरिंग सिस्टम (ओएमओएसवाईएस)

भौगोलिक जानकारी का उपयोग फील्ड विजिट के लिए स्थानों के चयन के लिए किया जा सकता है। पूरी तरह से यादृच्छिकीकरण दृष्टिकोण पर निर्भर रहने के बजाय, संदिग्ध स्थानों/मामलों की पहचान ऑफ-साइट निगरानी प्रणाली के माध्यम से की जा सकती है। इन पहचाने गए स्थानों को बाद के फील्ड विजिट के दौरान भौतिक रूप से सत्यापित किया जा सकता है, जिसके परिणामस्वरूप समय और संसाधन आवश्यकताओं के संदर्भ में अधिक दक्षता के साथ अधिक लक्षित निगरानी हो सकती है। ओएमओएसवाईएस को इस उद्देश्य को ध्यान में रखते हुए विकसित किया गया था और इसमें निम्नलिखित दो अभिन्न घटक शामिल हैं:

(i) स्थानिक उपाय, जो स्थान डेटा पर आधारित उपाय हैं और

(ii) स्थानिक उपायों का उपयोग करके संदिग्ध मामलों की पहचान करने के लिए संकेतक ढांचा।

(iii) उपर्युक्त प्रत्येक घटक की कार्यप्रणाली का विवरण निम्नलिखित उप-अनुभागों में दिया गया है।

III.1 *स्थानिक उपाय*

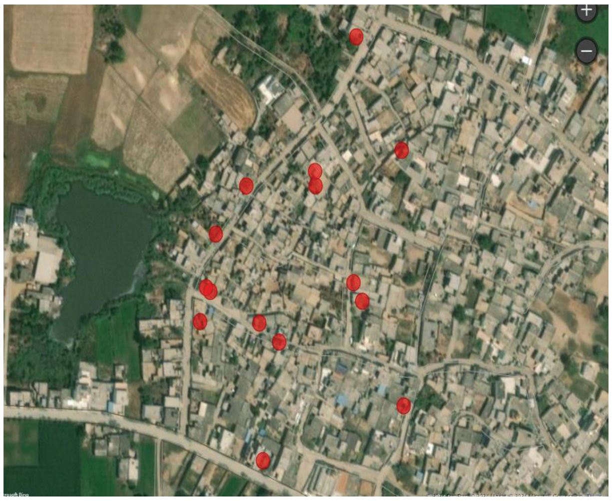
उत्तरदाताओं के अक्षांश और देशांतर द्वारा आरोपित सर्वेक्षण स्थान का मानचित्र एक उपयोगी प्रारंभिक खोजपूर्ण उपकरण के रूप में कार्य करता है, जिससे यह आकलन किया जा सकता है कि निर्देशों के साथ पर्याप्त प्रसार और अनुपालन प्राप्त हुआ है या नहीं (चार्ट 2)।[1] हालांकि, यह कच्चा दृष्टिकोण सैकड़ों मानचित्रों को नियमित रूप से तैयार करने और उनकी निगरानी करने की तार्किक असुविधा से भरा है, और यह क्षेत्र जांच के लिए स्थानों/उत्तरदाताओं के चयन में सीमित अंतर्दृष्टि प्रदान करता है। तदनुसार, प्रभावी निगरानी के लिए, मानचित्र में निहित जानकारी को सांख्यिकीय उपायों में बदल दिया जाता है, जैसा कि नीचे विस्तार से बताया गया है:

स्थान दूरी अंतर (एलडीजी)

जांचकर्ताओं के इच्छित स्थान पर क्षेत्र भ्रमण के सत्यापन के प्राथमिक लक्ष्य को ध्यान में रखते हुए, इच्छित सर्वेक्षण स्थान से उत्तरदाताओं की दूरी को एक उपयुक्त उपाय माना जाता है, जिसे एलडीजी कहा जाता है। अक्षांश और देशांतर को दर्शाते हुए

[1] कई ऑनलाइन सेवा प्रदाता हैं जो विभिन्न स्थानों के लिए ऑनलाइन मानचित्र उपलब्ध कराते हैं। यदि सर्वेक्षण स्थानों के लिए मानचित्र आसानी से उपलब्ध न हों तो इनका उपयोग किया जा सकता है।

चार्ट 2: एक विशेष स्थान पर जांचकर्ता द्वारा कवर किए गए स्थानों को दर्शाने वाली नमूना छवि



स्रोत: बिंग मैप्स पर आधारित लेखक का चित्रण

l वें स्थान के लिए (ϕ_l, λ_l) और इस स्थान के r वें प्रतिवादी के लिए (ϕ_r, λ_r) के रूप में, LDG की गणना इस प्रकार की जा सकती है:[2]

$$LDG_r = R\sqrt{(\Delta\phi_r)^2 + (cos(\phi_{mr})\Delta\lambda_r)^2} \qquad (1)$$

जहाँ $\Delta\phi_r = \phi_r - \phi_l$; $\Delta\lambda_r = \lambda_r - \lambda_l$; $\phi_{mr} = (\phi_r + \phi_l)/2$ और R पृथ्वी की त्रिज्या है। यहां एलडीजी आर के समान इकाई में है।

समीकरण (1) स्थानिक दूरियों की गणना के लिए समतुल्य सन्निकटन सूत्र को दर्शाता है, जो विशेष रूप से छोटी दूरियों के लिए कुशल है (सिल्वा एट अल., 2014)। इस सूत्र का उपयोग करके, किसी स्थान से सभी उत्तरदाताओं के लिए एलडीजी की गणना की जा सकती है। एलडीजी जितना बड़ा होगा, उतनी ही अधिक संभावना है कि संबंधित उत्तरदाता का सर्वेक्षण इच्छित स्थान पर आयोजित नहीं किया गया था।

एलडीजी माप की गणना के लिए इच्छित सर्वेक्षण स्थान के भू-निर्देशांक का पूर्व ज्ञान आवश्यक है जो हमेशा आसानी से उपलब्ध नहीं हो सकता है। उदाहरण के लिए, भारतीय संदर्भ में, ग्रामीण सर्वेक्षणों के लिए नमूना फ्रेम में आम तौर पर गाँव शामिल होंगे, आधिकारिक एजेंसियों द्वारा प्रदान की गई गाँव की सीमा का केन्द्रक, जो सार्वजनिक डोमेन में उपलब्ध है, गाँव के स्थान के लिए एक प्रॉक्सी के रूप में काम कर सकता है।[3] इसके विपरीत, शहरी सर्वेक्षणों के लिए इस मुद्दे को संबोधित करना अधिक जटिल है यदि वे मतदान केंद्रों या वार्डों पर आधारित नमूना फ्रेम पर निर्भर करते हैं, जहां सीमा विवरण सार्वजनिक डोमेन में आसानी से उपलब्ध नहीं हैं। चयनित शहरी सर्वेक्षण स्थानों की स्थान संबंधी जानकारी के लिए प्रत्यक्ष प्रॉक्सी की अनुपस्थिति में, विभिन्न सेवा प्रदाताओं द्वारा प्रदान की गई जियो-कोडिंग सुविधा का उपयोग उनके अक्षांश और देशांतर प्राप्त करने के लिए किया जा सकता है।[4]

प्रत्युत्तरदाता दूरी अंतर (आरडीजी)

इच्छित स्थान पर अन्वेषक की यात्रा की पुष्टि करने के अलावा, क्षेत्र सत्यापन का दूसरा उद्देश्य सर्वेक्षण के डिजाइन के अनुसार घरों की एक विशिष्ट संख्या को छोड़ने के संबंध में निर्देशों के अनुपालन की पुष्टि करना है। आरडीजी, जिसकी गणना लगातार उत्तरदाताओं के बीच की दूरी के आधार पर की जाती है, ऐसे मामलों में एक मूल्यवान मीट्रिक के रूप में कार्य करती है। प्रथम प्रतिवादी और क्रमिक प्रतिवादी के लिए अक्षांश और देशांतर को $(\phi_{r1}, \lambda_{r1})$ के रूप में निरूपित करना। यानी... दूसरे प्रतिवादी के रूप में $(\phi_{r2}, \lambda_{r2})$, दूसरे प्रतिवादी के लिए आरडीजी की गणना इस प्रकार की जा सकती है:

$$RDG_{r2} = R\sqrt{(\Delta\phi_{r2})^2 + (cos(\phi_{mr2})\Delta\lambda_{r2})^2} \qquad (2)$$

जहाँ $\Delta\phi_{r2} = \phi_{r2} - \phi_{r1}$; $\Delta\lambda_{r2} = \lambda_{r2} - \lambda_{r1}$;

$\phi_{mr2} = (\phi_{r2} + \phi_{r1})/2.$

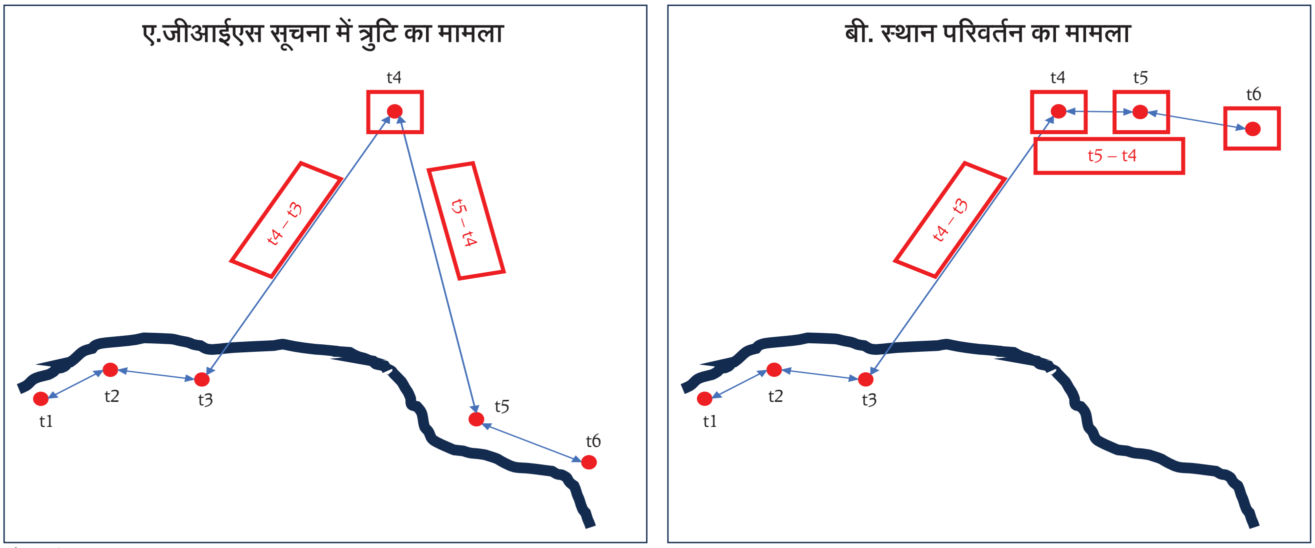
इस तरीके से, किसी स्थान पर सभी उत्तरदाताओं के लिए आरडीजी की गणना की जा सकती है। यह ध्यान रखना महत्वपूर्ण है कि गणना एक ही दिन में कवर किए गए सभी उत्तरदाताओं के लिए की जानी चाहिए क्योंकि उसी स्थान पर अगले दिन का शुरुआती बिंदु पिछले दिन के अंतिम स्थान से मेल नहीं खा सकता है। उच्च आरडीजी मान बड़े विचलन को इंगित करते हैं, जबकि

[2] ϕ और λ रेडियन में हैं। ϕ_m को $cos(\phi_{mr})$ निर्धारित करने के लिए उपयोग की जाने वाली विधि के साथ संगत इकाइयों में होना चाहिए।

[3] भारत में, गाँवों की सीमाएँ भारतीय सर्वेक्षण विभाग द्वारा जनता को उपलब्ध कराई जाती हैं, जिन्हें https://onlinemaps.surveyofindia.gov.in/ पर देखा जा सकता है।

[4] उदाहरण के लिए, गूगल जियो-कोडिंग सेवाएं प्रदान करता है, जहां सारणीबद्ध डेटा में स्थान के पते को भू-स्थानिक डेटा में परिवर्तित किया जा सकता है।

चार्ट 3: आरडीजी माप द्वारा इंगित वि चलन के उदाहरण



स्रोत: लेखक का स्पष्टीकरण

छोटे आरडीजी मान अवलोकनों या उत्तरदाताओं के एक करीबी समूह का सुझाव देते हैं। सर्वेक्षण स्थान के भीतर उत्तरदाताओं के एक समान प्रसार के लिए लक्ष्य बनाते समय दोनों परिदृश्य अवांछनीय हैं।

जबकि आरडीजी निगरानी के लिए एक उपयोगी उपाय के रूप में कार्य करता है, यह स्वीकार किया जाना चाहिए कि कभी-कभी, कमजोर इंटरनेट कनेक्टिविटी वाले स्थानों में, जीपीएस सुविधा गलत स्थान संबंधी डेटा को विशेष रूप से दूरदराज के क्षेत्रों में कैप्चर कर सकती है। इसके परिणामस्वरूप लगातार साक्षात्कारों के बीच बड़ी दूरी की गणना हो सकती है जो वास्तविक स्थिति को सटीक रूप से प्रतिबिंबित नहीं करती है, जिससे झूठे झंडे उत्पन्न होने से विश्लेषण की प्रभावशीलता कम हो जाती है। साथ ही, चूंकि यह ज्ञात है कि स्थान संबंधी जानकारी में त्रुटियां हो सकती हैं, इसलिए संभावना है कि निर्देश उल्लंघन के बारे में सही सकारात्मक संकेतों को गलत सकारात्मक के रूप में अनदेखा कर दिया जाए। जैसा कि चार्ट 3 ए में दर्शाया गया है , चौथा अवलोकन दूसरों से काफी अलग है, और चार्ट 3 बी में , चौथा, पांचवां और छठा अवलोकन, हालांकि क्लस्टर किया गया है,

प्रत्युत्तरदाता समय अंतराल (RTG)

जैसा कि पहले चर्चा की गई है, आरडीजी की व्याख्या से जुड़ी चुनौतियों को कम करने के लिए, सीएपीआई उपकरण में दर्ज साक्षात्कार की शुरुआत और समाप्ति के समय के टिकट बेहद फायदेमंद पाए गए हैं। यह ध्यान रखना महत्वपूर्ण है कि जब तक जांचकर्ता पूरे दिन एक ही तरह के वाहन का उपयोग करता है (जैसे, पैदल, साइकिल, स्कूटर, मोटरबाइक), एक विशिष्ट दूरी तय करने में लगने वाले समय में बहुत अधिक बदलाव होने की संभावना नहीं है। इस धारणा का लाभ उठाते हुए, एक साक्षात्कार के अंत समय और बाद के साक्षात्कार के शुरू होने के समय के बीच के अंतर की गणना की जाती है और उसे RTG के रूप में दर्शाया जाता है।

पहले प्रत्युत्तरदाता के लिए साक्षात्कार के प्रारंभ समय और समाप्ति समय को (ts_{r1}, te_{r1}) और अगले प्रत्युत्तरदाता, यानी दूसरे प्रत्युत्तरदाता के लिए (ts_{r2}, te_{r2}) के रूप में दर्शाते हुए, दूसरे प्रत्युत्तरदाता के लिए आरटीजी की गणना इस प्रकार की जा सकती है:

$$RTG_{r2} = ts_{r2} - te_{r1} \qquad (3)$$

इस तरह, एक ही दिन में एक स्थान पर कवर किए गए सभी उत्तरदाताओं के लिए आरटीजी की गणना की जा सकती है। यदि आरडीजी उपाय द्वारा चिह्नित मामले के लिए आरटीजी, उसी स्थान पर गैर-चिह्नित मामलों के आरटीजी के साथ संरेखित होता है, तो यह अनुमान लगाया जा सकता है कि उच्च आरडीजी एक

गलत सकारात्मक होने की संभावना है और संकेत व्यवधान के कारण उत्पन्न हो सकता है। यह ध्यान रखना महत्वपूर्ण है कि संकेत मुद्दों से उत्पन्न गलत सकारात्मक कम से कम एक जोड़े में हो सकते हैं। यह चार्ट 3 ए में चित्रित किया गया है जहां 4 वें और 5 वें उत्तरदाता दोनों के लिए, पिछले उत्तरदाता के साथ दूरी मुख्य रूप से 4 वें उत्तरदाता की स्थान संबंधी जानकारी के मुद्दे के कारण बड़ी है। दूसरी ओर, यदि उच्च आरडीजी के साथ उच्च आरटीजी है, तो यह सुझाव देता है कि जांचकर्ता लक्षित नमूना इकाई/इकाइयों से विचलित हो सकता है।

III.2 *संकेतक ढांचा*

डिजाइन और लॉजिस्टिक्स विचारों की जटिलता के आधार पर पिछले उप-अनुभाग में चर्चा किए गए उपायों का उपयोग करके क्षेत्र कार्य की निगरानी करने के दो संभावित दृष्टिकोण हैं, अर्थात, 1) स्थान-पहले दृष्टिकोण, और 2) उत्तरदाता-पहले दृष्टिकोण। स्थान-पहले दृष्टिकोण में पहले संदिग्ध स्थानों की पहचान करना और फिर उन स्थानों के लिए उत्तरदाता डेटा की जांच करना शामिल है। इस दृष्टिकोण में, काफी ऊंचे औसत एलडीजी और काफी उच्च और निम्न आरडीजी दोनों को प्रदर्शित करने वाले क्षेत्रों को आगे की जांच के लिए चिह्नित किया जाता है। इन पहचाने गए स्थानों में उत्तरदाताओं के व्यक्तिगत एलडीजी और आरडीजी मूल्यों की उनके आरटीजी मूल्यों के साथ जांच की जाती है ताकि झूठे झंडों से बचा जा सके। इस प्रक्रिया से बाहर चिह्नित स्थानों/उत्तरदाताओं को क्षेत्र के दौरे के लिए चुना जा सकता है। दूसरा दृष्टिकोण, यानी उत्तरदाता-पहले दृष्टिकोण

हालाँकि दोनों दृष्टिकोण ऑफ-साइट निगरानी के लिए उपयोगी हैं, लेकिन जब एफ़एसयू की संख्या बड़ी हो और प्रत्येक एफ़एसयू के भीतर कम संख्या में प्रतिक्रियाएँ प्राप्त हों, तो स्थान-प्रथम दृष्टिकोण का उपयोग किया जा सकता है। दूसरी ओर, जब सीमित संख्या में स्थानों से बड़ी संख्या में प्रतिक्रियाएँ प्राप्त होती हैं, तो उत्तरदाता-प्रथम दृष्टिकोण का उपयोग किया जा सकता है।

क्षेत्र सत्यापन के लिए स्थान/उत्तरदाताओं के यादृच्छिक चयन की तुलना में, स्थानिक उपाय-आधारित दृष्टिकोण अधिक लाभ प्रदान करता है क्योंकि यह संदिग्ध मामलों के ऑफ-साइट चयन की अनुमति देता है। इसके अलावा, जबकि यादृच्छिक चयन दृष्टिकोण केवल स्थानों के एक अंश की जाँच करता है, प्रस्तावित विधि सभी स्थानों की निगरानी करती है, जिससे पूरे सर्वेक्षण डेटा के लिए नियंत्रण उपायों की प्रभावशीलता और पहुँच बढ़ जाती है।

छोटे पैमाने के सर्वेक्षणों के मामले में या जब पर्याप्त संसाधन हों, तो संदिग्ध स्थानों/मामलों को चिह्नित करने के लिए स्थान-प्रथम दृष्टिकोण या उत्तरदाता-प्रथम दृष्टिकोण को मैन्युअल रूप से निष्पादित किया जा सकता है, लेकिन यह विशेष रूप से तब मुश्किल हो जाता है जब स्थानों की संख्या व्यापक हो, संसाधन सीमित हों, और समयबद्धता एक बाध्यकारी बाधा हो। इसके अतिरिक्त, सर्वेक्षण डेटा सर्वेक्षणकर्ता को अनुक्रमिक रूप से या वास्तविक समय [5] (पूर्व अधिक सामान्य है) में प्रेषित किया जाता है, जिसे नियमित गुणवत्ता मूल्यांकन प्रक्रिया से गुजारा जा सकता है। चल रहे सर्वेक्षण क्षेत्र कार्य के दौरान, डिजाइन और निर्देशों से विचलन की प्रारंभिक पहचान बाद के सर्वेक्षण साक्षात्कारों की डेटा गुणवत्ता सुनिश्चित करने में महत्वपूर्ण हो जाती है। जबकि ऊपर चर्चा किए गए उपायों को प्रासंगिक स्थान संबंधी डेटा उपलब्ध होने के बाद आसानी से गणना की जा सकती है, मैन्युअल प्रक्रिया के माध्यम से संदिग्ध मामलों की पहचान सख्त समयसीमा के कारण चुनौतीपूर्ण हो सकती है, जिसे मैन्युअल हस्तक्षेप की आवश्यकता के बिना संदिग्ध मामलों की पहचान करने वाली प्रणाली को लागू करके संबोधित किया जा सकता है। ओएमओएसवाईएस में, इस क्षमता को संकेतक दृष्टिकोण का उपयोग करके एकीकृत किया गया है, जहां संदिग्ध मामलों को चिह्नित करने के लिए पहले चर्चा किए गए स्थान संबंधी उपायों के शीर्ष पर प्रासंगिक संकेतक डिज़ाइन किए गए हैं। दृष्टिकोण का विवरण नीचे दिया गया है:

मॉडल संचालित संकेतक (एमडीआई) दृष्टिकोण

यदि विगत समयावधियों से समान/समान/पायलट सर्वेक्षणों का डेटा उपलब्ध है, तो एमडीआई पद्धति का उपयोग उन संदिग्ध मामलों को चिह्नित करने के लिए किया जा सकता है जहां मॉडल विगत डेटा का उपयोग करके बनाए गए हैं।

[5] वास्तविक समय में फील्ड सर्वेक्षणों से डेटा साझा करना बहुत मुश्किल है, अगर असंभव नहीं है तो परिचालन संबंधी मुद्दों (जैसे दूरदराज के क्षेत्रों में असंगत सिग्नल) और लागत आदि के कारण। हालांकि, कई व्यावहारिक स्थितियों के लिए, डेटा का क्रमिक साझाकरण उद्देश्य की पूर्ति के लिए पर्याप्त हो सकता है। क्रमिक डेटा साझाकरण का उदाहरण तब होता है जब एक दिन में कवर किए गए सभी उत्तरदाताओं का डेटा एक या दो दिन के अंतराल के साथ साझा किया जाता है।

स्थान-प्रथम दृष्टिकोण के लिए, l स्थान के लिए मॉडल निम्नलिखित सामान्य रूप ले सकता है:

$$y_l = f(\beta_l) + \epsilon_l \quad (4)$$

कहाँ,

y_l , l^{th} स्थान के लिए रुचि का चर है;

β_l, l वें स्थान के लिए भौगोलिक वर्गीकरण का सदिश है;

f के बीच संबंध को मॉडल करने के लिए उपयोग किया जाने वाला कार्यात्मक रूप है y_l; तथा

ϵ_l मॉडल में अवशिष्ट पद का प्रतिनिधित्व करता है।

उत्तरदाता-प्रथम दृष्टिकोण के लिए, सामान्यता की हानि के बिना, उपरोक्त मॉडल को ऊपर परिभाषित स्थान-स्तरीय विशेषताओं के समान, व्यक्तिगत उत्तरदाता-स्तरीय विशेषताओं के आधार पर परिभाषित किया जा सकता है। मॉडल का चुनाव सर्वेक्षण के प्रकार पर निर्भर करता है। हालांकि रैखिक मॉडल का व्यापक रूप से उपयोग किया जाता है, लेकिन किसी अन्य मॉडलिंग दृष्टिकोण को भी एकीकृत किया जा सकता है। β_l स्थानों के समान क्लस्टरिंग पर निर्भर करता है; यानी, संयोजनों के प्रत्येक सेट के भीतर, उम्मीद है कि स्थान रुचि के चर के संदर्भ में समान होंगे। उदाहरण के लिए, ग्रामीण सर्वेक्षणों के लिए, β_l के घटकों में राज्य, जिला, जनसंख्या समूह आदि शामिल हो सकते हैं। शहरी सर्वेक्षणों के लिए, ये घटक शहर या वार्ड हो सकते हैं।

(4) में मॉडल के लिए सामान्य रूप को पिछले डेटा का उपयोग करके अनुमानित किया जा सकता है और एक ऐसी प्रणाली के लिए लागू किया जा सकता है जिसके माध्यम से क्रमिक रूप से साझा किए गए डेटा को आउटलेर्स के लिए जांचा जा सकता है। इस उद्देश्य के लिए, अनुमानित ϵ_l को मानकीकृत किया जा सकता है, और इन मानकीकृत मूल्यों की z_p के साथ तुलना करने के बाद चरम मूल्यों को चिह्नित किया जा सकता है, जहां z_p मानक सामान्य वितरण का महत्वपूर्ण मूल्य है जैसे कि $Prob(|Z| > z_p) = p$ या $Prob(Z > z_p) = p$, जैसा भी मामला हो। p का मूल्य परीक्षण के लिए महत्व का स्तर है, जिसे उपलब्ध संसाधनों और समय की कमी को देखते हुए जांचे जा सकने वाले संदिग्ध मामलों की संख्या के आधार पर निर्धारित किया जा सकता है। पहचाने गए संदिग्ध मामलों को फॉलो-अप फील्ड विजिट के लिए चुना जा सकता है,

निश्चित नियंत्रण संकेतक (एफसीआई) दृष्टिकोण

ऐसी परिस्थितियाँ होती हैं जहाँ सर्वेक्षण के लिए पिछला डेटा उपलब्ध नहीं होता (जैसे, एक बिल्कुल नया सर्वेक्षण शुरू करते समय या किसी मौजूदा नियमित सर्वेक्षण के कवरेज का विस्तार करते समय), जहाँ पहले चर्चा किए गए एमडीआई दृष्टिकोण को सीधे लागू नहीं किया जा सकता। एक वैकल्पिक दृष्टिकोण मॉडल को क्रमिक रूप से बनाना है, यानी, जैसे-जैसे डेटा बैचों में आता है, मॉडल का अनुमान लगाया जा सकता है और बाद के बैचों में आउटलेयर की पहचान करने के लिए इसका उपयोग किया जा सकता है। इसके बाद, उस बिंदु तक प्राप्त संयुक्त डेटा के साथ मॉडल का फिर से अनुमान लगाया जा सकता है, और यह प्रक्रिया तब तक जारी रहती है जब तक अनुमान स्थिर नहीं हो जाते।

हालांकि, गैर-नियमित सर्वेक्षण के मामले में, मॉडल स्थिरीकरण के लिए व्यापक संसाधन आवंटित करना अनावश्यक हो सकता है। सर्वेक्षण के उद्देश्य के साथ संरेखण में सर्वेक्षण के आरंभकर्ता के अनुभव के आधार पर, स्थानिक उपायों पर निश्चित नियंत्रण सीमाएँ लगाना एक स्वाभाविक दृष्टिकोण हो सकता है। ऐसी स्थिति में, डेटा पर निश्चित-नियंत्रण निचली सीमा (एफ़सीएलएल) और/या निश्चित-नियंत्रण ऊपरी सीमा (एफ़सीयूएल) लगाई जा सकती है। इन सीमाओं के बाहर किसी भी सर्वेक्षण प्रतिक्रिया को संदिग्ध के रूप में चिह्नित किया जा सकता है, जिसके लिए आगे की जांच की आवश्यकता होती है। हालांकि सरल, इस दृष्टिकोण के परिणामस्वरूप संदिग्ध मामलों की पर्याप्त संख्या हो सकती है यदि अत्यधिक रूढ़िवादी सीमाएँ लगाई जाती हैं, या यह वास्तविक संदिग्ध मामलों को अनदेखा कर सकता है यदि लगाई गई सीमाएँ बहुत अधिक उदार हैं। एक और चुनौती स्थानों का प्राकृतिक समूहन है, जहाँ रुचि के चर समूहों में व्यापक रूप से भिन्न हो सकते हैं: यदि सभी समूहों पर समान रूप से समान सीमाएँ लगाई जाती हैं, तो संदिग्ध मामले केवल कुछ समूहों में देखे जा सकते हैं, जबकि यदि सीमाएँ क्लस्टर-वार लगाई जानी हैं, तो बहुत अधिक सीमाएँ उच्च मैन्युअल हस्तक्षेप के साथ लगाने की आवश्यकता होगी।

इस प्रकार, जबकि एमडीआई और एफसीआई दृष्टिकोण अलग-अलग और गैर-अतिव्यापी उपयोग के मामले हैं, किसी विशेष दृष्टिकोण का चयन उद्देश्य, प्रकार और संसाधन उपलब्धता

पर भी निर्भर करता है। अगला खंड इन दृष्टिकोणों के विवरण में गहराई से जाता है और सिंथेटिक डेटा के आधार पर विभिन्न स्थितियों में उनकी प्रभावशीलता की जांच करता है।

IV. उपयोग मामला: रिज़र्व बैंक के घरेलू सर्वेक्षण

भारतीय रिजर्व बैंक वर्ष 2005 और 2010 में शुरू किए गए द्विमासिक [6] शहरी परिवार सर्वेक्षणों अर्थात् परिवारों का मुद्रास्फीति प्रत्याशा सर्वेक्षण (आईईएसएच) और उपभोक्ता विश्वास सर्वेक्षण (सीसीएस) के माध्यम से प्रमुख आथक मानदंडों पर उपभोक्ता भावनाओं में उतार-चढ़ाव की नियमित रूप से निगरानी करता रहा। शहरी क्षेत्रों में उनके सैंपलिंग फ्रेम में मतदान केंद्रों को एफ़एसयू के रूप में उपयोग किया जाता है, जिन्हें सर्वेक्षण स्थान माना जाता है, जहाँ से एसएसयू (यानी, आमतौर पर प्रत्येक एफ़एसयू में 15-20 घर) प्रतिक्रिया प्राप्त करने के लिए चुने जाते हैं। चयनित सर्वेक्षण स्थान के भीतर, एक यादृच्छिक शुरुआत और पहले सफल साक्षात्कार के बाद, साक्षात्कार के लिए अगले घर का चयन करने के लिए दाएं हाथ के नियम का पालन करते हुए एक निश्चित संख्या में घरों को छोड़ दिया जाता है (नमूना अंतराल)।

ये दोनों सर्वेक्षण 19 शहरों को कवर करते हैं, जिनमें से प्रत्येक इन शहरों के 400 से अधिक स्थानों का सर्वेक्षण करता है। आरबीआई के मौद्रिक नीति चक्र के अनुरूप, ये सर्वेक्षण द्वि-मासिक आवृत्ति पर आयोजित किए जाते हैं, जिसमें सर्वेक्षण फ़ील्डवर्क 10 दिनों के लिए निर्धारित होता है। जैसा कि खंड II में चर्चा की गई है, ऑन-साइट/फ़ॉलो-अप फ़ील्ड विज़िट इन सर्वेक्षणों के लिए गुणवत्ता आश्वासन ढांचे का एक अभिन्न अंग हैं। इन सर्वेक्षणों की लगातार प्रकृति और सख्त समयसीमा के कारण, सर्वेक्षण स्थानों का व्यापक भौगोलिक कवरेज मौजूदा गुणवत्ता आश्वासन ढांचे के तहत डेटा की गुणवत्ता सुनिश्चित करने में चुनौतियां पेश करता है।

ओएमओएसवाईएस की उपयोगिता आरबीआई के घरेलू सर्वेक्षणों की सेटिंग्स और डिज़ाइन, जैसे प्रत्येक साक्षात्कार के बाद 10-12 घरों को छोड़ना और दाएं हाथ के नियम का पालन करना, का उपयोग करके उत्पन्न सिंथेटिक डेटा के माध्यम से ऐसे परिदृश्यों में प्रदर्शित की जाती है। [7] सिंथेटिक डेटा इस तरह से उत्पन्न होता है कि इसमें सर्वेक्षण की आवश्यक विशेषताएं शामिल होती हैं, जिसमें चित्रण के लिए चिंता के कुछ उदाहरण शामिल होते हैं। स्थानों के लिए डेटा विभिन्न राज्यों, जिलों और जनसंख्या समूह संयोजनों में एलडीजी, आरडीजी और आरटीजी में देखे गए औसत और प्रसरण का उपयोग करके उत्पन्न होता है। चित्रण के उद्देश्य से, क्रमशः सूत्र (1) और (2) का उपयोग करके एलडीजी और आरडीजी की गणना के लिए आवश्यक पृथ्वी की त्रिज्या का मान $R = 6371.0$ किमी लिया गया है।

IV.1 *कंप्यूटिंग उपाय*

अनुभाग III में पहले चर्चा किए गए उपायों को यहां वास्तविक डेटा के साथ प्रदर्शित किया गया है। उदाहरण के लिए निम्नलिखित मामलों का चयन किया गया है:

- **केस 1:** एक सर्वेक्षण स्थान जहां निर्देशों का उचित रूप से पालन किया जाता है और इंटरनेट सिग्नल समस्या के कारण कोई गलत अलर्ट नहीं होता है (चार्ट 4)
- **केस 2:** एक सर्वेक्षण स्थान जहां निर्देशों का सही ढंग से पालन किया गया है, लेकिन इंटरनेट सिग्नल समस्याओं के कारण कुछ गलत अलर्ट उत्पन्न होते हैं (चार्ट 5)

---

[6] मार्च 2014 तक, आरबीआई ने तिमाही आधार पर शहरी घरेलू सर्वेक्षण, अर्थात् 'घरों की मुद्रास्फीति अपेक्षा सर्वेक्षण' और 'उपभोक्ता विश्वास सर्वेक्षण' आयोजित किए। इसके बाद, सर्वेक्षणों को द्वि-मासिक मौद्रिक नीति ढांचे के साथ संरेखित करने के लिए, दो अतिरिक्त दौर (मई और नवंबर) शुरू किए गए।

[7] सिंथेटिक डेटासेट का निर्माण लागत और समय प्रभावी तरीके से मॉडलों को प्रशिक्षित/परीक्षण/चित्रित करने का उपयोगी तरीका है, जहां डेटासेट को वास्तविक जीवन परिदृश्यों को दर्शाते हुए सिम्युलेट किया जाता है (सर्गेई, 2021)।

[8] इस चित्रण में पृथ्वी की औसत त्रिज्या का उपयोग किया गया है क्योंकि पृथ्वी एक पूर्ण गोलाकार नहीं है; यह एक चपटा गोलाकार है, जिसका अर्थ है कि यह ध्रुवों पर थोड़ा चपटा है और भूमध्य रेखा पर उभरा हुआ है। पृथ्वी की औसत त्रिज्या पृथ्वी के आकार के कारण त्रिज्या में इस मामूली बदलाव को ध्यान में रखती है। हालाँकि, सूत्र (1) और (2) की प्रयोज्यता इस विशिष्ट विकल्प तक सीमित नहीं है, और वैकल्पिक पृथ्वी त्रिज्या, जैसे भूमध्यरेखीय या ध्रुवीय त्रिज्या (वांग एट अल., 2021), का उपयोग विशिष्ट आवश्यकताओं के अनुसार किया जा सकता है।

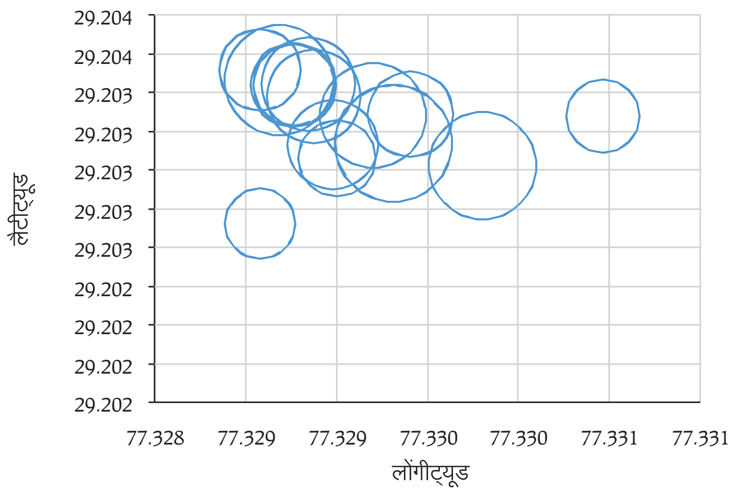
[9] सीआईएसबीआई वाणिज्यिक और सहकारी बैंकों के साथ-साथ अन्य चुनिंदा वित्तीय संस्थानों के स्थान संबंधी विवरण के साथ-साथ बुनियादी सांख्यिकीय रिटर्न (बीएसआर) कोड के आवंटन और रखरखाव के लिए रिज़र्व बैंक का ऑनलाइन पोर्टल है। इनमें शाखाएँ, कार्यालय, गैर-प्रशासनिक रूप से स्वतंत्र कार्यालय (एनएआईओ), स्वचालित टेलर मशीनें (एटीएम), निश्चित स्थान के व्यवसाय संवाददाता और अन्य निश्चित स्थान के ग्राहक सेवा केंद्र शामिल हैं। सारणीबद्ध रूप में डेटा https://dbie.rbi.org.in/#/banking-outlet पर भी जनता के लिए उपलब्ध कराया गया है।

चार्ट 4: निर्देशों से कोई विचलन किए बिना स्थान का सर्वेक्षण करें

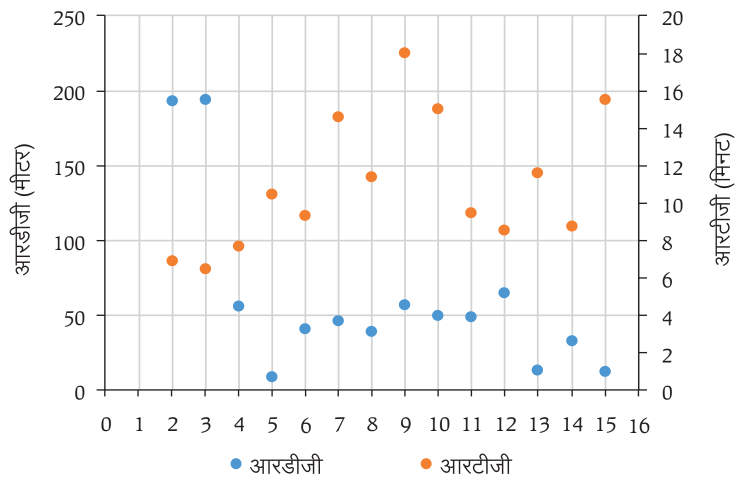
ए: स्थान के लिए सभी उत्तरदाताओं का डेटा

प्रतिक्रिया	अवधि (मिनट)	आरटीजी (मिनट)	आरडीजी (मीटर)
1	13		
2	12	7	193
3	13	6	194
4	16	8	56
5	10	10	8
6	10	9	40
7	12	15	46
8	10	11	39
9	15	18	57
10	9	15	49
11	11	9	49
12	11	9	65
13	11	12	13
14	8	9	33
15	11	15	12

बी: वृत्त की त्रिज्या के रूप में आरटीजी के साथ स्थान चार्ट



सी: अनुक्रमिक क्रम में आरडीजी और आरटीजी



स्रोत : लेखक की गणना

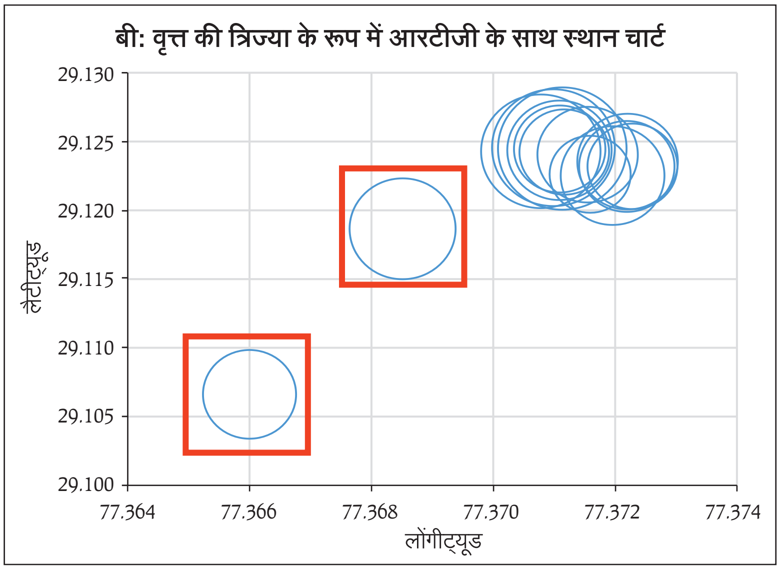
- **केस 3:** एक सर्वेक्षण स्थान जहां निर्देशों से विचलन है, सांख्यिकीय उपायों द्वारा उठाए गए अलर्ट द्वारा संकेतित (चार्ट 6)

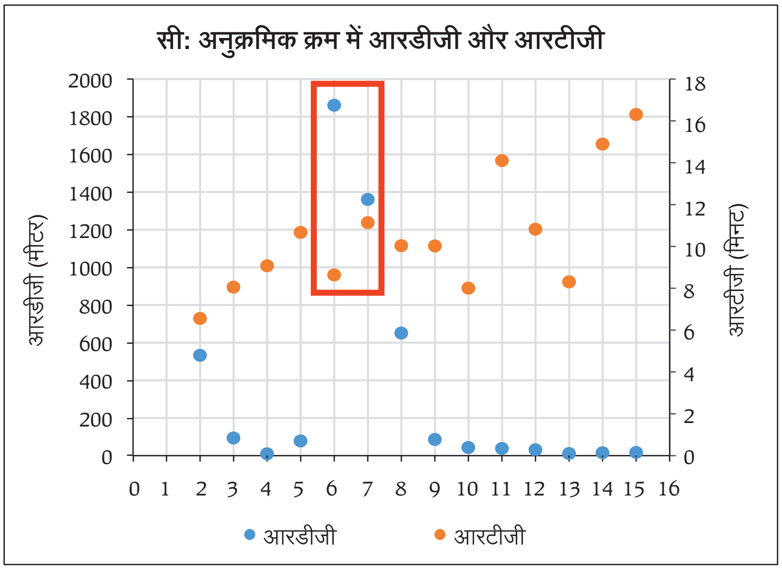
केस 1 के लिए, उपायों की गणना की गई है और चार्ट 4 में प्रस्तुत किया गया है। यहां, 'अवधि' प्रत्येक साक्षात्कार की अवधि का प्रतिनिधित्व करती है। चार्ट 4 बी से यह स्पष्ट है कि कोई महत्वपूर्ण विचलन नहीं देखा गया है और दूरी अंतराल (आरडीजी) और समय अंतराल (आरटीजी) उत्तरदाताओं में स्थिरता प्रदर्शित करते हैं, जैसा कि चार्ट 4 ए और चार्ट 4 सी में दिखाया गया है। दूसरी ओर, केस 2 के लिए, दो अवलोकनों को आरडीजी उपाय (चार्ट 5 ए) का उपयोग करके लाल झंडा लगाया गया है। चार्ट 5 बी में उत्तरदाताओं का सर्कल प्लॉट इंगित करता है कि ये दो अवलोकन वास्तव में अन्य अवलोकनों के समूह से विचलित हुए। हालांकि, इन दो अवलोकनों के लिए आरटीजी अन्य आरटीजी अवलोकनों के अनुरूप है, जैसा कि सर्कल की त्रिज्या और डेटा तालिका में रीडिंग द्वारा संकेत दिया गया है। इससे पता चलता है कि आरडीजी द्वारा चिह्नित विचलन गलत

चार्ट 5 : आरडीजी माप से फाल्स पॉज़िटिव के साथ सर्वेक्षण स्थान

ए: स्थान के लिए सभी उत्तरदाताओं का डेटा

प्रतिक्रिया	अवधि (मिनट)	आरटीजी (मिनट)	आरडीजी (मीटर)
1	9		
2	11	7	533
3	11	8	93
4	10	9	9
5	10	11	77
6	12	**9**	1861
7	14	**11**	1360
8	13	10	651
9	11	10	85
10	12	8	43
11	10	14	38
12	16	11	31
13	12	8	11
14	13	15	15
15	10	16	16





स्रोत : लेखक की गणना

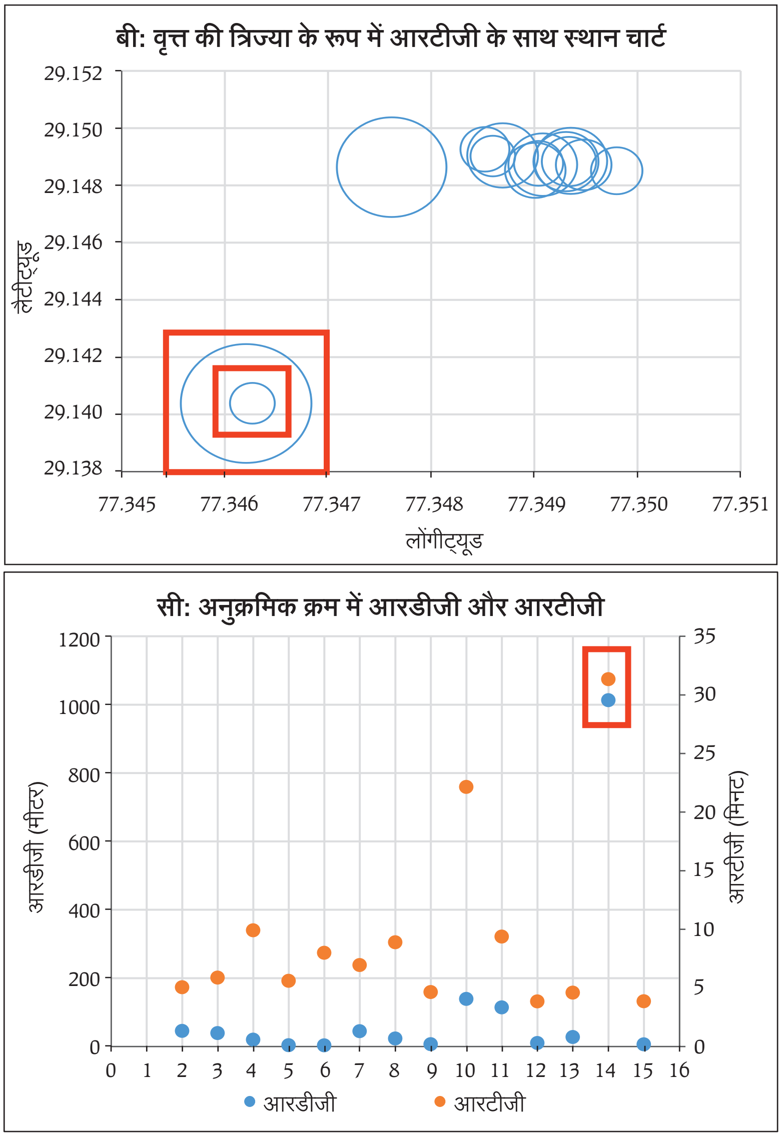
सकारात्मक हो सकते हैं चार्ट 6 केस 3 को दर्शाता है, जो पहले दो मामलों से अलग है। चार्ट 6 बी में, दो अवलोकन अन्य अवलोकनों के समूह से दृष्टिगत रूप से अलग हैं। पहले अवलोकन के लिए आरटीजी का उच्च मूल्य, जिसे आरडीजी माप द्वारा चिह्नित किया गया है, उच्च संभावना को दर्शाता है कि अन्वेषक शेष समूह या इच्छित सर्वेक्षण स्थान से दूर चला गया हो सकता है। बाद के उत्तरदाताओं के लिए आरडीजी और आरटीजी की रीडिंग से संकेत मिलता है कि अन्वेषक ने पहले चिह्नित उत्तरदाता का साक्षात्कार अन्य उत्तरदाताओं से बहुत दूर करने के बाद, उसी दूर के स्थान पर बाद के उत्तरदाताओं का साक्षात्कार करना जारी रखा। ये उत्तरदाता क्षेत्र सत्यापन के लिए संभावित मामले हैं।

उपरोक्त उदाहरणों ने ऑफ-साइट निगरानी के माध्यम से क्षेत्र सर्वेक्षणों में समस्याग्रस्त मामलों की पहचान करने में स्थानिक उपायों की क्षमता को दर्शाया है। हालाँकि, इसकी प्रभावशीलता सभी सर्वेक्षण स्थानों के लिए स्थानिक उपायों की मैन्युअल निगरानी की आवश्यकता से कम हो जाती है, जो श्रमसाध्य और त्रुटियों और चूक के लिए अतिसंवेदनशील दोनों है। इन सीमाओं

चार्ट 6 : आरडीजी माप से वास्तविक सकारात्मकता के साथ सर्वेक्षण स्थान

ए: स्थान के लिए सभी उत्तरदाताओं का डेटा

प्रतिक्रिया	अवधि (मिनट)	आरटीजी (मिनट)	आरडीजी (मीटर)
1	12		
2	10	5	45
3	15	6	39
4	14	10	20
5	14	6	3
6	12	8	3
7	20	7	44
8	11	9	23
9	12	5	6
10	12	22	139
11	16	9	114
12	10	4	10
13	11	5	27
14	13	31	1012
15	9	4	6



स्रोत : लेखक की गणना

को संबोधित करने के लिए अनुभाग III.2 में प्रस्तावित रणनीतियों को नीचे प्रदर्शित किया गया है।

IV.2 *संदिग्ध मामलों को चिह्नित करना*

आरबीआई के सर्वेक्षणों के नमूनाकरण डिज़ाइन को देखते हुए, जिसमें प्रत्येक स्थान से केवल थोड़ी संख्या में उत्तरदाताओं (आमतौर पर 15-20) के साथ बड़ी संख्या में स्थान शामिल हैं, इस संदर्भ में अधिक उपयुक्त दृष्टिकोण स्थान-प्रथम दृष्टिकोण होगा, जिसकी चर्चा अनुभाग III.2 में की गई है। इस प्रकार, यहाँ प्रदर्शन में स्थान-प्रथम विधि शामिल होगी, अर्थात, पहले संदिग्ध स्थानों की पहचान करना और उसके बाद उन स्थानों के लिए सर्वेक्षण प्रतिक्रियाओं की जाँच करना।

सबसे पहले, हम एमडीआई ढांचे पर चर्चा करते हैं, जिसमें नियंत्रण सीमाएँ सिंथेटिक डेटा से प्राप्त की जाती हैं। उदाहरण के लिए, मॉडल (4) के लिए एक रैखिक कार्यात्मक रूप पर विचार किया जाता है, जिसमें राज्य, जिला और जनसंख्या समूह जैसे भौगोलिक वर्गीकरण शामिल होते हैं। भारत में विशाल भौगोलिक और सांस्कृतिक विविधता को देखते हुए, भौगोलिक वर्गीकरण महत्वपूर्ण हैं क्योंकि महत्वपूर्ण विचार, जैसे कि जनसंख्या घनत्व,

क्षेत्र आयाम और घरों का फैलाव सभी राज्यों में समान नहीं हैं। यहां तक कि एक राज्य के भीतर भी, सभी जिले एक समान नहीं हैं। चूंकि ये सभी कारक देखे गए एलडीजी, आरडीजी और आरटीजी को प्रभावित करते हैं, इसलिए उन्हें मॉडल में समायोजित करने की आवश्यकता है ताकि रुचि के चर में उनके द्वारा प्रेरित भिन्नता को ध्यान में रखा जा सके।

इन सेटिंग्स के बाद, आरडीजी के लिए रैखिक स्थान-प्रथम मॉडल [10] और l^{th} स्थान के लिए एलडीजी को इस प्रकार लिखा जा सकता है:

$$\overline{RDG}_l = \alpha_1 + \beta_{1i} + \gamma_{1j} + \xi_{1k} + \epsilon_{1ijkl} \qquad (5)$$

$$\overline{LDG}_l = \alpha_2 + \beta_{2i} + \gamma_{2j} + \xi_{2k} + \epsilon_{2ijkl} \qquad (6)$$

जहाँ,

आरडीजी, और एलडीजी, इस राज्य में वें जिले के जनसंख्या समूह के स्थान l के लिए सभी आरडीजी और एलडीजी के के औसत हैं:

$(\alpha_1, \beta_{1i}, \gamma_{1j}, \xi_{1k})$ और $(\alpha_2, \beta_{2i}, \gamma_{2j}, \xi_{2k})$ क्रमशः मॉडल (5) और (6) के लिए राज्य, जिले और जनसंख्या समूह के अवरोधन और निश्चित प्रभाव हैं।

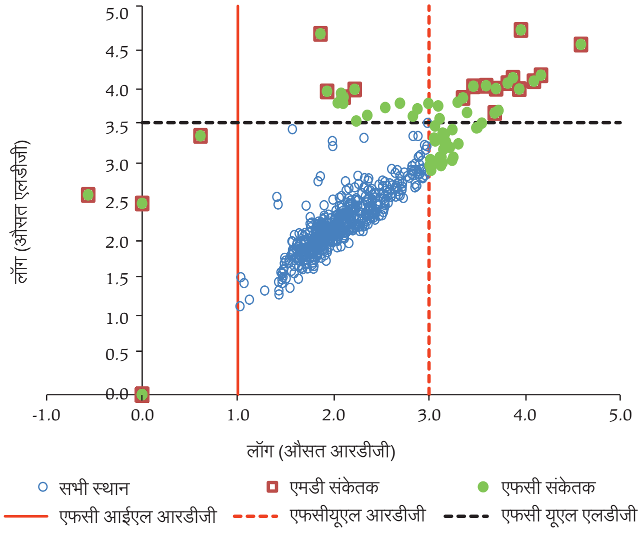
ϵ_{1ijkl} और ϵ_{2ijkl} क्रमशः आरडीजी और एलडीजी मॉडल के लिए अवशिष्ट त्रुटि पद का प्रतिनिधित्व करते हैं।

अनुमान लगाने के लिए, सिंथेटिक डेटा को दो भागों में विभाजित किया गया था, पहले का इस्तेमाल मॉडल (5) और (6) का अनुमान लगाने के लिए किया गया था, और दूसरे भाग को ताजा डेटा आगमन के रूप में माना गया था, जहां अनुमानित मॉडल का इस्तेमाल संदिग्ध मामलों की पहचान करने के लिए किया गया था। अनुमान लगाने के दौरान, आरोपित संदिग्ध मामलों को डेटासेट से बाहर रखा गया था ताकि यह सुनिश्चित किया जा सके कि मॉडल अनुमान सर्वेक्षण में केवल अंतर्निहित विविधताओं को ही पकड़ता है और चरम टिप्पणियों से प्रभावित नहीं होता है। डेटा के दूसरे भाग में संदिग्ध मामलों या आउटलेर्स की पहचान पहले से चर्चा की गई मानकीकृत अनुमानित अवशिष्ट विधियों का उपयोग करके की जाती है। [11] जैसा कि उप-अनुभाग III.1 में चर्चा की गई है, आरडीजी के उच्च और निम्न दोनों मूल्य अवांछनीय हैं; इस प्रकार, मॉडल (5) के लिए दो-तरफा तुलना का उपयोग $p_1 = 20$ प्रतिशत के महत्व के स्तर के साथ किया जाता है।

परिणाम चार्ट 7 में प्रस्तुत किए गए हैं , जहाँ स्पष्टता के लिए, औसत आरडीजी और औसत एलडीजी मान लघुगणकीय शब्दों में प्रस्तुत किए गए हैं। एमडीआई ढांचे का उपयोग करके पहचाने गए स्थानों को लाल वर्गों के रूप में हाइलाइट किया गया है। हालाँकि, इन संदिग्ध स्थानों की अधिक बारीकी से जाँच आवश्यक है, विशेष रूप से, प्रत्येक उत्तरदाता के आरडीजी की इन उत्तरदाताओं के आरटीजी के साथ-साथ जाँच की जानी चाहिए ताकि यह निर्धारित किया जा सके कि पहचाने गए स्थान पर जमीनी दौरा करना उचित है या नहीं।

एफसीआई दृष्टिकोण को स्पष्ट करने के लिए, उसी डेटासेट का उपयोग किया जाता है। इस उद्देश्य के लिए, आरडीजी के लिए एफसीएलएल और एफसीयूएल को क्रमशः 10 मीटर (लॉग (10) = 1) और 1000 मीटर (लॉग (1000) = 3) माना जाता है। इसी तरह, एलडीजी के लिए एफसीयूएल को 3000 मीटर (लॉग (3000) = 3.5) के रूप में लगाया जाता है। चार्ट 7 में चिह्नित मामलों को हरे रंग के बिंदुओं के रूप में हाइलाइट किया गया है । ऊपर दिया गया चित्रण दर्शाता है कि, जब सीमाएँ ठीक से चुनी जाती हैं, तो दोनों दृष्टिकोण समान परिणाम दे सकते हैं (एफसीआई दृष्टिकोण में झंडों की थोड़ी अधिक संख्या के साथ)।

चार्ट 7: संदिग्ध स्थान की पहचान



स्रोत : लेखकों की गणना.

[10] 10 मॉडल को स्थिति और आवश्यकताओं के अनुसार अनुकूलित किया जा सकता है।

[11] हालाँकि, सामान्यता की हानि के बिना किसी अन्य आउटलाइयर पहचान दृष्टिकोण का भी उपयोग किया जा सकता है।

V. निष्कर्ष

घरेलू सर्वेक्षणों के कई क्षेत्रों में, व्यक्तिगत साक्षात्कारों में लक्षित प्रतिक्रियाएँ प्राप्त करने, बेहतर समझ और बेहतर सर्वेक्षण प्रतिक्रिया दर के मामले में टेलीफोनिक/ऑनलाइन सर्वेक्षणों की तुलना में अलग-अलग लाभ हैं। अच्छे नीतिगत निर्णय और उचित निगरानी के लिए विश्वसनीय और समय पर सर्वेक्षण डेटा की आवश्यकता होती है, जहाँ सावधानीपूर्वक तैयार किए गए सर्वेक्षण डिज़ाइन के अनुपालन को सुनिश्चित करना, विशेष रूप से भौगोलिक रूप से व्यापक सर्वेक्षणों में, चुनौतियाँ पेश करता है। पारंपरिक अनुवर्ती सत्यापन क्षेत्र के दौरे तार्किक रूप से चुनौतीपूर्ण हो जाते हैं, खासकर दूरदराज के क्षेत्रों में।

इस लेख में प्रस्तावित ओएमओएसवाईएस इन चिंताओं को व्यावहारिक तरीके से संबोधित करने के लिए सांख्यिकीय उपायों को विकसित करने के लिए सीएपीआई उपकरणों के माध्यम से प्राप्त जीआईएस डेटा का उपयोग करता है। यह ऑफ-साइट निगरानी के माध्यम से संदिग्ध मामलों की पहचान करने के लिए स्थानिक उपायों (जैसे, एलडीजी , आरडीजी और आरटीजी) को नियोजित करता है। एलडीजी उत्तरदाताओं की इच्छित स्थान से दूरी को मापता है, जबकि आरडीजी और आरटीजी साक्षात्कारों के बीच निर्देशों को छोड़ने और समय अंतराल के अनुपालन का आकलन करते हैं। प्रणाली कुशल कार्यान्वयन के लिए सांख्यिकीय तरीकों को शामिल करती है। रिज़र्व बैंक के घरेलू सर्वेक्षणों की सेटिंग्स और डिज़ाइन का उपयोग करके उत्पन्न सिंथेटिक डेटासेट का उपयोग ओएमओएसवाईएस की प्रभावशीलता को दर्शाने के लिए किया जाता है। एमडीआई दृष्टिकोण, जो नियंत्रण सीमाएँ निर्धारित करने के लिए ऐतिहासिक डेटा का उपयोग करता है, और एफसीआई दृष्टिकोण, जो निश्चित सीमाओं को नियोजित करता है, संदिग्ध मामले की पहचान के लिए प्रदर्शित किए जाते हैं। दोनों दृष्टिकोण विभिन्न सर्वेक्षण स्थितियों के लिए उपयुक्त हैं और परिणाम संदिग्ध मामलों की पहचान करने में उनकी प्रभावशीलता को प्रदर्शित करते हैं।

यह ओएमओएसवाईएस व्यापक सर्वेक्षण निगरानी के लिए एक व्यवहार्य समाधान प्रदान करता है, विशेष रूप से सीमित संसाधनों और सख्त समयसीमा वाले परिदृश्यों में। यादृच्छिक चयन दृष्टिकोण के विपरीत, जो केवल कुछ स्थानों की जाँच करता है, यह समाधान सभी स्थानों की निगरानी करता है, जिससे सर्वेक्षण नियंत्रण उपायों की प्रभावशीलता और पहुँच में तेजी से वृद्धि होती है। यह क्षेत्र के दौरे के लिए एक लक्षित दृष्टिकोण सुनिश्चित करता है, दक्षता को अधिकतम करता है और विविध भौगोलिक क्षेत्रों में सर्वेक्षण की गुणवत्ता बनाए रखता है। निष्कर्ष विभिन्न क्षेत्रों में सीएपीआई-आधारित सर्वेक्षणों में डेटा गुणवत्ता में सुधार के लिए व्यापक अनुप्रयोगों का सुझाव देते हैं।

संदर्भ:

Baker, R.P., Bradburn, N.M., and Johnson, R.A. (1995). Computer-assisted personal interviewing: An experimental evaluation of data quality and cost. *Journal of Official Statistics*, Vol. 11, No. 4, 413–431. Retrieved from https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/computer-assisted-personal-interviewing-an-experimental-evaluation-of-data-quality-and-cost.pdf

Caeyers, B., Chalmers, N., and Weerdt, J.D. (2010). A comparison of CAPI and PAPI through a randomised field experiment. Retrieved from https://openknowledge.worldbank.org/server/api/core/bitstreams/7aad9cf4-06b4-5332-8ff2-9dc51ec1f016/content

Couper, M.P. (2000). Usability evaluation of computer-assisted survey instruments. *Social Science Computer Review*, 18(4), 384–396. https://doi.org/10.1177/089443930001800402

Reserve Bank of India. (2009). *Report of the Working Group on Surveys*, Available at https://rbi.org.in/scripts/PublicationReportDetails.aspx?UrlPage=&ID=557

Reserve Bank of India. (2010). *Report of The Technical Advisory Committee on Surveys, Sep-2009.* RBI Bulletin, May. Available at https://rbi.org.in/Scripts/BS_ViewBulletin.aspx?Id=11209

Reserve Bank of India (2018). Inflation Expectations Survey of Households: 2017-18. RBI Bulletin October, 105-116, available at https://m.rbi.org.in/Scripts/BS_ViewBulletin.aspx?Id=17820

Reserve Bank of India (2019). The Consumer Rules! Some Recent Survey-based Evidence. RBI Bulletin April, 117-131, available at https://m.rbi.org.in/scripts/BS_ViewBulletin.aspx?Id=18173

Reserve Bank of India. *Consumer Confidence Survey*, Various Issues, Available at https://www.rbi.org.in/scripts/BimonthlyPublications.aspx?head=Consumer%20Confidence%20Survey%20-%20Bi-monthly

Reserve Bank of India. *Inflation Expectation Survey on Households*, Various Issues, Available at https://www.rbi.org.in/scripts/BimonthlyPublications.aspx?head=Inflation%20Expectations%20Survey%20of%20Households%20-%20Bi-monthly

Sergey, I.N. (2021). *Synthetic Data for Deep Learning.* Springer Optimisation and Its Applications. Vol. 174. doi:10.1007/978-3-030-75178-4, https://link.springer.com/book/10.1007/978-3-030-75178-4

Silva, Y.N., Reed, J.M., Tsosie, L.M. and Matti, T.A. (2014). Similarity Join for Big Geographical Data, In E. Pourabbas (Ed.), *Geographical Information Systems – Trends and Technologies* (pp 20-49), Boca Raton: CRC Press – Taylor & Francis Group. DOI https://doi.org/10.1201/b16871

Survey of India. *Online Maps Portal*, https://onlinemaps.surveyofindia.gov.in/

Wang D., Li M., Huang X., and Zhang X. (2021). Spacecraft Autonomous Navigation Technologies Based on Multisource Information Fusion. (pp 311) Springer. https://doi.org/10.1007/978-981-15-4879-6