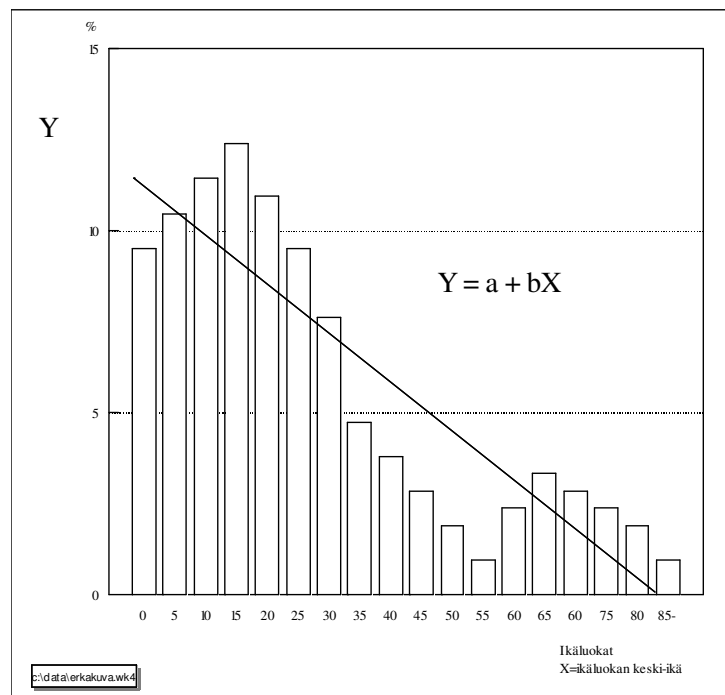


3. COULSON-INDEKSI

3.1. METODI

Michael Coulsonin (1968) mukaan ikärakenneindeksi tarjoaa kvantitatiivisen mittarin ikärakenteen mittaamiseen. Coulson havainnollisti ikäluokkien iän ja niiden osuuden välistä riippuvuutta regressiosuoralla. Tämän suoran kaltevuus ilmaisee ikärakenteen. Jyrkkä suora kuvaa nuorta väestöä ja loiva suora vanhaa väestöä. Suoran kulmakerroin ilmaisee suoran kaltevuuden. Sille tehdään lukumuunnos ja tätä saatua lukua Coulson kutsuu ikärakenneindeksiksi. Kulmakerroin havainnollistaa koko ikärakennetta. Ikärakenneindeksi lasketaan pienimmän neliösumman suoran (=PNS-suora) yhtälön avulla.

Coulsonin (1968) kehittämä indeksi saadaan laskemalla regressiosuoran yhtälö väestön ikäryhmien suhteellisen osuuden ja ikäryhmien keski-ikä välillä. Suhteellinen osuus ilmoitetaan prosenteissa ja ikäryhmät ovat 0-4, 5-9 ... yli 84-vuotiaat. Vastaavat ikäluokkien keski-ikäet ovat 2.5, 7.5, 12.5, 17.5 ... 82.5, 87.5 vuotta (kuva 7).



Kuva 7. Ikäjakauma ja sitä vastaava regressiosuora ($b=-0.14$).

$$Y = a + bX$$

Kaava 1

Yhtälössä (kaava 1) vakio b on pienimmän neliösumman menetelmällä lasketun regressiosuoran kulmakerroin, X on ikäryhmän keski-ikä eli luokkakeskus, vakio a on regressiosuoran leikkauspiste Y -akselin kanssa ja Y on regressiosuoran ilmoittama ko. ikäluokan osuus prosentteina. Coulson-indeksi ($=C_a$) saadaan kaavan 2 mukaisesti, kun ko. regressiosuoran kulmakertoimeen b lisätään luku -0.10000 ja luvun etumerkki poistetaan. Lopuksi luku kerrotaan vielä 100000 :lla. Kuvan 7 regressiosuoran kulmakerroin on -0.14 . Se voidaan muuntaa kaavan 2 mukaisesti, jolloin Coulson-indeksin C_a arvoksi tulee 24000 .

$$\text{Coulson-indeksi} = |-0.10000 + b| * 100000$$

Kaava 2

Coulson (1968) kokeili Coulson-indeksiä Kansas Cityn metropolin tilastoalueella (SMSA) Yhdysvalloissa. Tilastoalue oli jaettu 225 osa-alueeseen. Coulson-indeksin (C_a) arvo vaihteli välillä $00859-29026$ ($b=-0.19026-0.141$). Indeksillä avulla Coulson sai kuvatuksi kaupungin osa-alueiden ikärakenteita niin, että keskustan vanhat ja esikaupunkien nuoret ikärakenteet saatiin erottumaan (kuva 8). Hänen mukaansa indeksi kuvasi säännönmukaisesti alueiden ikärakenteita niin, että alueet, joilla oli vanha ikärakenne, saivat pieniä Coulson-indeksien arvoja ja alueet, joilla oli nuori ikärakenne, saivat suuria indeksien arvoja. Yhteenkasvaneiden kaupunkien Kansas City Missouri, Kansas City Kansas ja Independence keskustat erottuivat muista alueista selvästi pienemmillä indeksien arvoilla (kuva 8).

Silja Hoikkanen (1975) kokeili Coulson-indeksiä Suomen kaupunkien ikärakenteen havainnollistamiseen. Aineisto käsitti vuodet 1960 ja 1970. Kaupunkeja oli viisikymmentä. Coulson-indeksien (C_a) arvot vaihtelivat välillä $15515-25712$ ($b=-0.05515-0.15712$). Siddiqui (1984) käytti Coulson-indeksiä Intian Uddar Pradesh osavaltion alueiden kaupunkien ja maaseutuväestön ikärakenteiden vertailuun. Maaseutuväestön Coulson-indeksien (C_a) arvot vaihtelivat välillä $25180-40820$ ($b=-0.1518-0.3048$) ja kaupunkiväestön arvot välillä $31880-45000$ ($b=-0.2188-0.3500$).

Coulson-indeksien etuina voidaan pitää sitä, että se koskee koko ikärakennetta ja yksi luku ilmaisee koko ikärakenteen. Tämä mahdollistaa suorat vertailut eri väestöjen kesken riippumatta ajasta, paikasta ja populaation koosta (Birdsall 1980:467-470). Coulson (1968) tutki Coulson-indeksien luotettavuutta testaamalla t -testillä PNS-suoran kulmakertoimen merkitsevyyttä (kuva 9). Tällöin testisuure laskettiin selityssasteen (r^2) avulla. Testien riskitasona oli 0.01 . Tätä rajaa vastasi selityssasteen arvo 0.381 ja edelleen Coulson-indeksien arvo 14200 . Hänen mukaansa Coulson-indeksi oli epäluotettava arvoilla, jotka olivat alle 14200 . Coulson totesi edelleen, että ongelmat liittyivät vanhoihin ikärakenteisiin. Yli 15000 olevat indeksien arvot ovat hänen mukaansa luotettavia.

Kuva 8. Kansas Cityn metropolialueen tilastoalueiden osien ikärakenne Coulson-indeksillä ilmaistuna (Coulson 1968: kuva 6).

Kuva 9. Kansas Cityn metropolialueen tilastoalueiden osat regressioyhtälön selityssasteen (r^2) perusteella suuruusjärjestykseen lajiteltuna (Coulson 1968: kuva 5).

Birdsall (1980:467-471) kritisoi Coulson-indeksiä. Hänen mukaansa Coulson-indeksiä ei voida testata tilastollisin testein, koska testien asettamia tilastollisia vaatimuksia ei voida täyttää. Coulsonin käyttämässä menetelmässä on ongelmallista se, että luokkakeskusten arvot eivät ole sattumanvaraisesti valittuja lukuja, ikärakenteen sisäinen keskihajonta vaihtelee ikäluokittain sekä ikärakennefrekvenssit eivät ole jakaantuneet normaalisti. Birdsallin mukaan Coulson-indeksi sopii huonosti epälineaarisiin malleihin. Tällaisia malleja tarvitaan, jos "esiintyy häiriöitä, eli jos vanhusten osuus on suuri ko. alueella tai kyseessä on virkistyskeskus, hallituskeskus tai muuttoliikkeen vaivaama alue". Indeksien käyttöön liittyvät ongelmat tulevat Birdsallin mukaan alueellisesti laajenemaan, koska suuret ikäluokat ovat Yhdysvalloissa tulossa keski-ikään. Maan ikärakenne tulee yhä enemmän olemaan häiriintynyt, jolloin sitä on vaikea mitata Coulson-indeksillä.

Birdsall (1980) on ehdottanut Coulson-indeksin ongelmiin ratkaisuksi kehittämäänsä MSABI-indeksiä. MSABI-indeksi perustuu frekvenssijakauman vinoutta kuvaavaan kolmanteen origomomenttiin (kaava 3) (Smith, Duncan 1944:192; Mäkinen 1977:63; Ranta, Rita, Kouki 1991:41). Origomomentista saadaan normittamalla MSABI-indeksi (kaava 4).

MSABI-indeksi on jatkuva muuttuja. Sen ilmaisemaa lukuarvoa ei enää tarvitse testata tilastollisin testein. Sitä käytettäessä ikäluokan frekvenssi voi olla myös nolla.

$$Jakauman\ vinous =_3 \sqrt{\frac{\sum f_i X_i^3}{N}} \quad \text{Kaava 3}$$

MSABI-indeksin kaava Birdsallin (1980:469) mukaan

$$MSABI =_3 \sqrt{\frac{\sum f_i \left(\frac{(X_i - X_m + 1)}{s} \right)^3}{\sum f_i}} \quad \text{Kaava 4}$$

X_i = ikäluokan luokkakeskus

X_m = ikäluokkien luokkakeskusten jakauman keskiarvo

s = ikäluokkien luokkakeskusten jakauman keskihajonta

f_i = ikäluokan frekvenssi

Birdsallin käyttämä menetelmä täyttää tilastolliset vaatimukset paremmin kuin Coulson-indeksi. MSABI-indeksiä on kuitenkin hankala tulkita niin graafisesti kuin tilastollisestikin. Se ei pysty kuvaamaan luotettavasti kaikkia ikäjakaumia. Birdsallin mukaan keski-ikäisiin kohdistuneet ikärakenteen häiriöt aiheuttavat ongelmia indeksin luotettavuudelle.

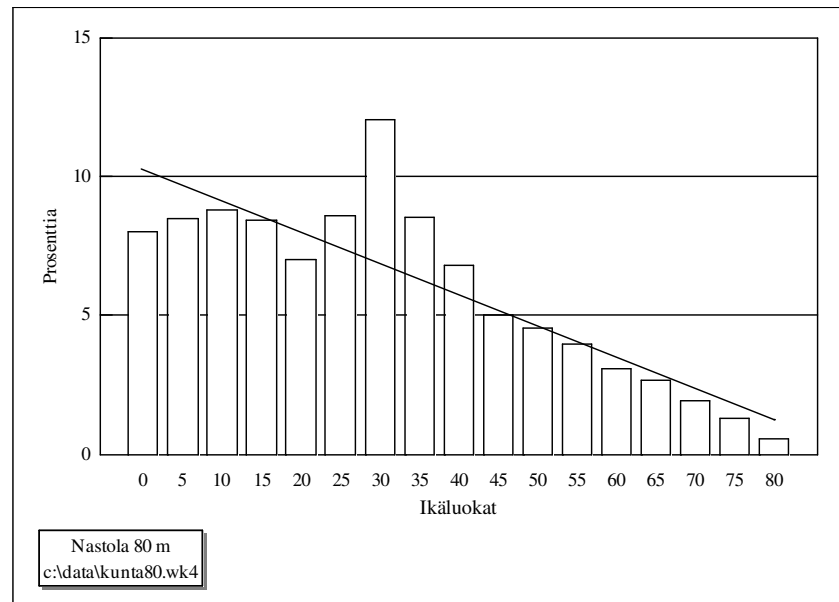
3.2. COULSON-INDEKSIIN KEHITTÄMINEN

Tässä tutkimuksessa on kehitetty Coulson-indeksin laskumenetelmää siitä, mitä Coulson (1968) ja Hoikkanen (1975) käyttivät. Alkuperäisessä muodossaan Coulson-indeksi (C_a) lasketaan tekemällä regressiosuoran kulmakertoimelle kaavan 2 mukainen muunnos. Tässä tutkimuksessa muunnosta ei tehdä, koska Coulson-indeksin ja regressiosuoran kulmakertoimen välinen yhteys jää liian hankalaksi päätellä. Regressiosuoran matemaattinen ominaisuus on se, että kulmakerrointa nolla vastaava suora on vaakasuora. Negatiivisiä kulmakertoimia vastaavat suorat ovat laskevia ja vastaavasti positiivisiä kulmakertoimen arvoja saavat suorat ovat nousevia. Useimmiten ikärakenteista lasketut regressiosuorat ovat laskevia. Uusi Coulson-indeksi (C_b) on siis sama kuin pelkkä regressiosuoran kulmakerroin b (kaava 1). Tässä tutkimuksessa kehitetyssä ikärakenteen luokittelumenetelmässä ei ole myöskään käytetty kaavan 2 mukaisesti muunnettuja kulmakertoimia Coulson-indeksin arvoina, joten uusi menetelmä eroaa Coulsonin (1968) ja Hoikkasen (1975) menetelmästä lähinnä lukujen

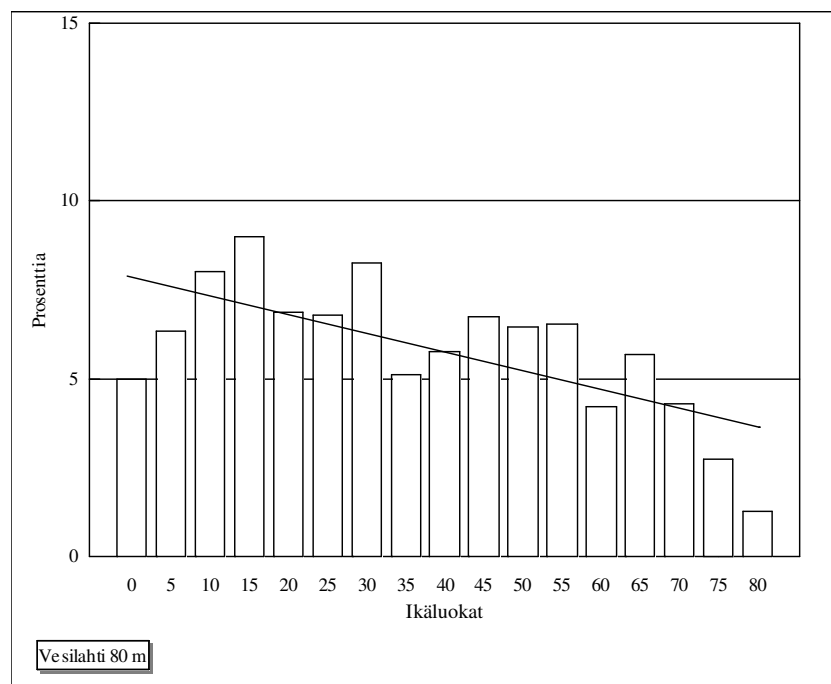
esitystavaltaan. Kulmakertoimen b arvoon ei lisätä kaavan 2 mukaisesti lukua -0.10000 , eikä poisteta saadun luvun etumerkkiä. Lisäksi tätä lukua ei kerrota 100000 :lla. C_b on sama kuin kulmakerroin b . Liitteessä 1 on muunnostaulukko näiden kahden laskumenetelmän välille. Liitteen taulukossa on kaavan 2 mukaisesti C_a ja sitä vastaavan C_b :n arvot. Siddiqui (1980) teki lukumuunnoksen hieman eri tavalla kertoen saadun luvun 10000 :lla 100000 :n sijaan. Coulson käytti 18 ikäluokkaa (kuva 7). Tässä tutkimuksessa on ikäluokat 80-84 ja yli 85 vuotiaat summattu yhteen, jolloin ikäluokkia jää 17 kappaletta (kuva 13). Tämän ikäluokan luokkakeskus on 82.5. Myös Hoikkanen (1975) käytti 17 ikäluokkaa. Tässä tutkimuksessa ilmoitetut Coulson-indeksin arvot ovat tyyppiä C_b , ellei toisin ilmoiteta.

Coulson-indeksi (C_b) saa sitä suuremman arvon mitä vanhempi ikärakenne on. Tämä perustuu siihen, että ikärakennetta kuvaavan regressiosuoran (kaava 1) kulmakertoimen (b) arvo kasvaa. Ikärakenteeltaan vanhimmat väestöt saavat nollaa suurempia positiivisia C_b -arvoja. Vastaavasti Coulson-indeksi saa sitä pienemmän arvon mitä nuorempi ikärakenne on. Mitä nuorempi ikärakenne on, sitä enemmän regressiosuora on laskeva ja kulmakerroin sitä negatiivisempi. Kansas Cityn tilastoalueilla Coulson-indeksi C_a :n pienin arvo oli 29000 (Coulson 1968:157). Se vastaa C_b :n arvoa -0.19 (kaava 2). Hoikkanen sai Suomen kaupunkien C_a :n pienimmäksi arvoksi 26000. Se vastaa C_b :n arvoa -0.16 (kaava 2). Tässä tutkimuksessa käytetään ilmaisia nuori, keski-ikäinen tai vanha ikärakenne oman subjektiivisen arvion perusteella sen mukaan, mitenkä Coulson-indeksin arvot jakaantuvat tai mitkä ovat nuorten, keski-ikäisten ja vanhusten osuudet. Jos esimerkiksi intialaisten maakuntien väestötilastoja olisi käytetty vertailuarvoina, kaikki tutkimusalueiden kunnat olisivat olleet ikärakenteeltaan vanhoja. "Nuoret väestöt" on siis varsin suhteellinen käsite.

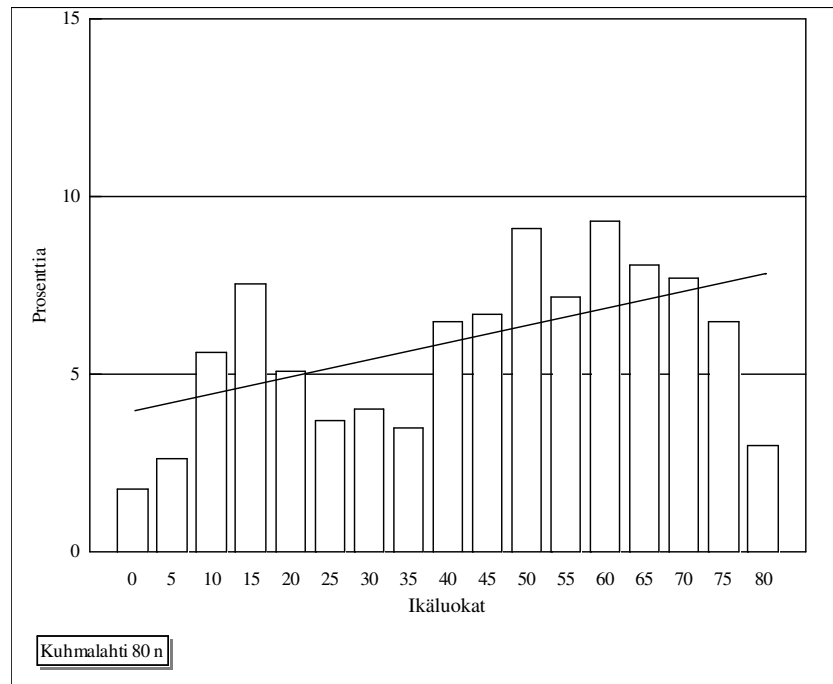
Kuvissa 10, 11 ja 12 on esimerkkejä Suomen kuntien erilaisista ikärakenteista ja niitä vastaavista Coulson-indeksin C_b arvoista. Nastolan miesten ikärakenne vuonna 1980 oli nuori (kuva 10). Nuoria ja keski-ikäisiä oli paljon, mutta vanhuksia vähän. Ikäjakauma sai pienen Coulson-indeksin arvon -0.10 . Vesilahden miesten ikärakenne vuonna 1980 oli keski-ikäinen (kuva 11). Ikäjakauma oli varsin tasainen ja ikäluokkien osuuksissa ei ollut suuria eroja. Kuhmalahden naisten ikärakenne vuonna 1980 oli vanha (kuva 12). Vanhojen ikäluokkien osuus oli hyvin korostunut, nuorten ja keski-ikäisten osuus oli sen sijaan pieni. Coulson-indeksi sai arvon 0.05 . Tällöin ikärakennetta kuvaava regressiosuora oli nouseva ja suoran kulmakerroin oli positiivinen. Coulson-indeksi C_b sai arvon -0.05 .



Kuva 10. Nuori ikärakenne. Coulson-indeksi C_b on -0.10.



Kuva 11. Keski-ikäinen ikärakenne. Coulson-indeksi C_b on -0.05.



Kuva 12. Vanha ikärakenne. Coulson-indeksi C_b on 0.05.

3.3. IKÄJAKAUMAN PNS-POIKKEAMA

Väestön ikäjakaumia voidaan verrata todellisiin tai teoreettisiin ikäjakaumiin. Ikärakennehäiriöillä tarkoitetaan todellisen ikäjakauman poikkeamia vertailuikäjakaumasta. Ei ole olemassa yleistä väestön ihanneikärakennetta, mutta jonkin tiettyjen väestön ominaisuuksien perusteella voidaan määritellä keinotekoinen vertailuikärakenne. Tällaisia ominaisuuksia ovat esimerkiksi syntyneiden ja kuolleiden määrä, ikäryhmittäinen kuolleisuus ja muuttoliike. Ikärakenne voisi olla tietyssä mielessä ihanteellinen esimerkiksi, jos se takaa mahdollisimman nopean tai hitaan väestönkasvun tai siinä on paljon veroja maksavia keski-ikäisiä ja vähän nuoria ja vanhuksia. Murthy (1980) käytti Australian väestön ikärakennetta vertailuikärakenteena, kun hän vertaili Aasian maiden väestöjen ikärakenteita toisiinsa. Thumerelle (1993:76-81) vertasi EY-maiden maakuntien väestöjen ikärakenteita vastaaviin stationäärisiin ikärakenteisiin.

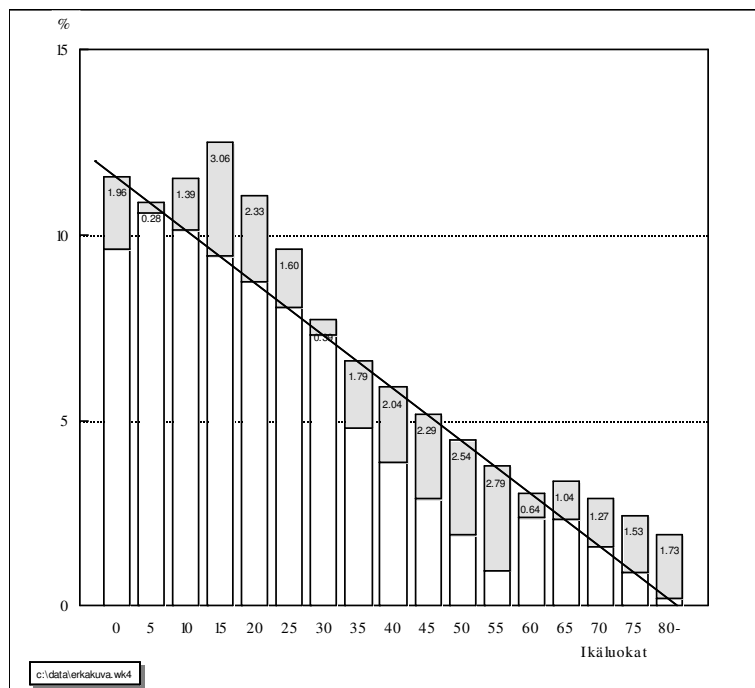
Coulson-indeksin luotettavuuden testaamiseen tässä tutkimuksessa kehitetään oma menetelmä. Havaittuja ikärakenteita verrataan havaituista ikärakenteista laskettujen regressiomallin mukaisiin ikärakenteisiin. Vertailujen lähtökohdaksi otettiin "nolla"-poikkeama ikärakennefrekvenssien ja ikäryhmien välisistä arvoista lasketusta regressiosuorasta (kuva 13). Näiden poikkeamien keskiarvo ilmaisee Coulson-indeksin luotettavuuden ja sitä kutsutaan tässä tutkimuksessa PNS-poikkeamaksi. Sanallisessa muodossa PNS-poikkeaman määritelmä on "erotusten keskiarvo, joka saadaan vähennettäessä jokaisen ikäluokan havaitusta arvosta regressioyhtälön ilmoittama arvo, summaamalla erotusten itseisarvot ja jakamalla saatu summa

ikäluokkien määrällä". Ikäluokkia on 17 kappaletta. Regressioyhtälön ilmoittama ikäjakauma on viivasuora. PNS-poikkeama ilmaisee paljonko havaittu ikärakenne poikkeaa regressioyhtälön ilmoittamasta ikärakenteesta. Ikäluokkien poikkeamat regressiosuorasta on rasteroitu kuvassa 13. Ikäjakauman keskimääräinen poikkeama pienimmän neliösumman suorasta (= PNS-poikkeama) määritellään matemaattisessa muodossa seuraavasti:

$$PNS - poikkeama = \frac{\sum_{i=1}^{17} [ikä_{i,frek} (hav) - ikä_{i,frek}(la)]^2}{17} \quad \text{Kaava 5}$$

jossa i:n arvot 1..17 vastaavat ikäluokkia 0-4 (2.5), 5-9 (7.5), 10-14 (12.5) ... yli 80 (82.5) -vuotiaat. Luokkakeskuksen arvo on suluissa.

PNS-poikkeama kertoo suoraan Coulson-indeksin luotettavuuden ilman tilastollista testausta. Lisäksi PNS-poikkeamaa voidaan käyttää ikärakennehäiriöiden yleisindikaattorina. PNS-poikkeama kehitettiin kuvaamaan Coulson-indeksin kokonaishäiriöitä. Se ei erottele tai painota mitään ikäluokkaa erikseen. Se ei reagoi havaintojen lukumäärään. Sille ei ole omaa tilastollista testiä ja PNS-poikkeamien suuruuden "merkittävyys" on jouduttu arvioimaan empiirisesti. PNS-poikkeama ei aseta mitään vaatimuksia, kuten ikäjakauman jatkuvuus tai normalisuus. Ikäluokan frekvenssi saa olla nolla, mutta mikäli nolla-luokkia on monta, niin indeksin käyttö ei ole mielekäästä.



Kuva 13. Ikärakenne ja sitä vastaava regressiosuora, PNS-poikkeama (1.68) ja Coulson-indeksi ($C_a=24000$, $C_b = -0.14$). Tarkemmat selitykset tekstissä.

Ensimmäisen asteen regressioyhtälön ilmoittama suora ikärakenne on vertailu- tai "ihanne"-ikärakenteeksi aina hieman epätodellinen. Vaikka todellisesta ikärakenteesta puuttuisivat vuosittaiset häiriöt, kuten vakiodusta ikärakenteesta, niin se ei olisi kuitenkaan muodoltaan suora, vaan "kaartuva". Tästä aiheutuu epätarkkuutta Coulson-indeksiin ja edelleen Coulson-indeksin tarkkuutta mittaavaan PNS-poikkeamaan. Epälineaariset mallit kuvaavat ainakin teoriassa ikäjakaumaa todenmukaisemmin kuin lineaarinen malli. Ne ovat kuitenkin vaikeasti yleistettäviä. Ikärakennetta kuvaavien epälineaaristen mallien, erityisesti regressiomallien, lisäksi saattaisi olla mahdollista käyttää populaation ns. vakioitua ikärakennetta vertailuarvona. Mikäli käytetään epälineaarisia malleja vertailuikäjakaumina, niin niiden ja todellisten ikäjakaumien välisten poikkeamien keskiarvot ja keskihajonnat eivät kuvaa Coulson-indeksin luotettavuutta, koska Coulson-indeksi määritellään ensimmäisen asteen regressioyhtälöllä. Tunnetuimmat vakioitujen ikärakenteiden mallit ovat YK:n julkaisemissa 158 kuolleisuustaulussa vuodelta 1955 ja 72 kuolleisuustaulussa vuodelta 1982 sekä Princetonin yliopiston vuosina 1966 ja 1983 julkaisemissa 192 kuolleisuustaulussa (Newell 1988:131-139).