

Sianlihan tuotannon lyhytaikaisia vaihteluja selittävä ennustemalli

LEO GRANBERG

Maatilahallitus, 00170 Helsinki 17

Prognosticatory model for pork production

LEO GRANBERG

Board of Agriculture 00170 Helsinki 17

Abstract. The plenary section of the National Board of Agriculture has studied the feasibility of constructing a short-term prognosticatory model for the production of pork. The prognosticatory model under investigation has been the linear regression model. The explicatory variables are the quantity of served sows, the size of the last yields (barley, oats) and the retail (deflated) price of pork one year prior to slaughtering.

The models investigated were statistically satisfactory ($R^2 = .94, .92$). However, some problems of a more general nature remained.

1) The construction of a price-variable posed problems: the difficulty lay in finding a single statistical interdependence between price development and production volume. The reason for this could be the production structure of Finnish agriculture: production mainly takes place on what are known as 'family farms'. When production takes place under low-price, non-profitable conditions there are few opportunities for 'family farms' to decrease production.

2) In order to be able to forecast, one must receive homogenous time series at least twice a year and preferably four times.

3) A prognosticatory model can never take account of all the variables which affect the matter under consideration. For example, the state may introduce new legislative measures which will profoundly influence production: no model can take such factors into account beforehand.

4) The coefficients (or their estimates) employed in a prognosticatory model age quickly which leads to the birth of cumulative forecast errors. This is particularly true when rapid changes take place in the sphere of production (centralisation, fluctuating prices). The making of accurate forecasts is therefore dependent on the stability of production and socio-economic conditions.

Johdanto

Maatalouspoliittisessa suunnittelutyössä on lukuisista syistä kyettävä ennakoidaan maataloustuotteiden tuotannon ja kulutuksen tulevaa kehitystä. Jo pelkästään valtiontalouden budjetin laadinta edellyttää tällaista ennakkointia, samoin vientiin ja varastointiin liittyvien toimenpiteiden järkevä suunnittelu sekä luonnollisesti tuotantopoliittinen suunnittelu.

Tilastotieteen menetelmien ja ATK-tekniikan kehittyessä ovat viime vuosina tulleet yhä laajempaan käyttöön ns. ekonometriset ennusteet. Yhtenä sovellutuksena esitellään sianlihan tuotannon ennustamiseksi maatilahallituksessa kehiteltyä lyhyen aikavälin regressio-tyyppistä ennustemallia ja sen käyttöön liittyviä ongelmia.

Eräitä aikaisempia malleja

Sianlihan tuotantoa on pyritty ennustamaan jo verraten varhain Suomesakin (KAARLEHTO 1959, KETTUNEN 1968). Aikaisempien mallien kehittelyä on suuressa määrin ohjannut tavoite löytää hintapoliittista päätöksentekoa palvelevaa tietoutta. Niinpä kummassakin mallissa on selittävinä muuttujina peräti kolme hintamuuttujaa: sianlihan ja perunan hinta sekä rehuviljojen hinta. Sitä vastoin näissä malleissa ei ole käytetty »biologisia» muuttujia, kuten esim. astutettujen emakoiden lukumäärää. Ilmeisesti tästä johtuu selitysasteen jääminen melko alhaiseksi, Kettusella se oli .89 ja Kaarlehdolla .72

Jos ensisijaiseksi tavoitteeksi asetetaan tuotannon mahdollisimman tarkka ennustaminen eikä esim. oikeiden hintapoliittisten ratkaisujen löytyminen, on luonnollista pyrkiä ottamaan malliin jokin biologiseen kiertoon liittyvä selittäjä. Eräs tällainen kokeilu on Yliselän (ANON. 1972) PSM:ssa estimoima malli, jossa käytettiin selittäjänä astutusindeksissä tapahtuvia muutoksia. Ilmeisesti astutusindeksin epätarkkuudesta johtui, ettei mallin soveltuvuus ennustekäyttöön ollut riittävä.

Ruotsissa kehitetyistä malleista mainittakoon STOJKOVICIN (1965) malli, jossa keskeinen selittävä muuttuja on sianlihan ja ohran hintasuhde. Tuotantomalli sisältyi Stojkovicillä laajempaan moniyhtälömalliin. Tuotantomallin osalta selitysasteeksi tuli .89. SVENSSONIN (1976) Uppsalassa kehittämä malli perustui vuotta aikaisempiin teurastuksiin ja kahta vuotta aikaisempaan sianlihan hintaan. Tätä mallia pyrittiin soveltamaan alueelliseen tarkasteluun, mutta selitysasteet jäivät kovin heikoiksi (koko maassa .61).

Ennustemallin rakenne

Esiteltävä malli perustuu neljännesvuosittaisiin havaintoihin vuosilta 1965—1975. Selitettävänä muuttujana on teurastettujen lihotussikojen lukumäärä, jolloin keskiteuraspainon vaihtelut eivät aiheuta häiriöitä estimaateissa.

Vuoden päähän ulottuvan »perusmallin» (viive t-4) lisäksi on estimoitu parametrit neljännesvuotta pidempään »viivästettyyn malliin» (viive t-5). Näiden mallien muuttujat ja niiden parametreille lasketut t-arvot on esitetty oheisessa asetelmassa.

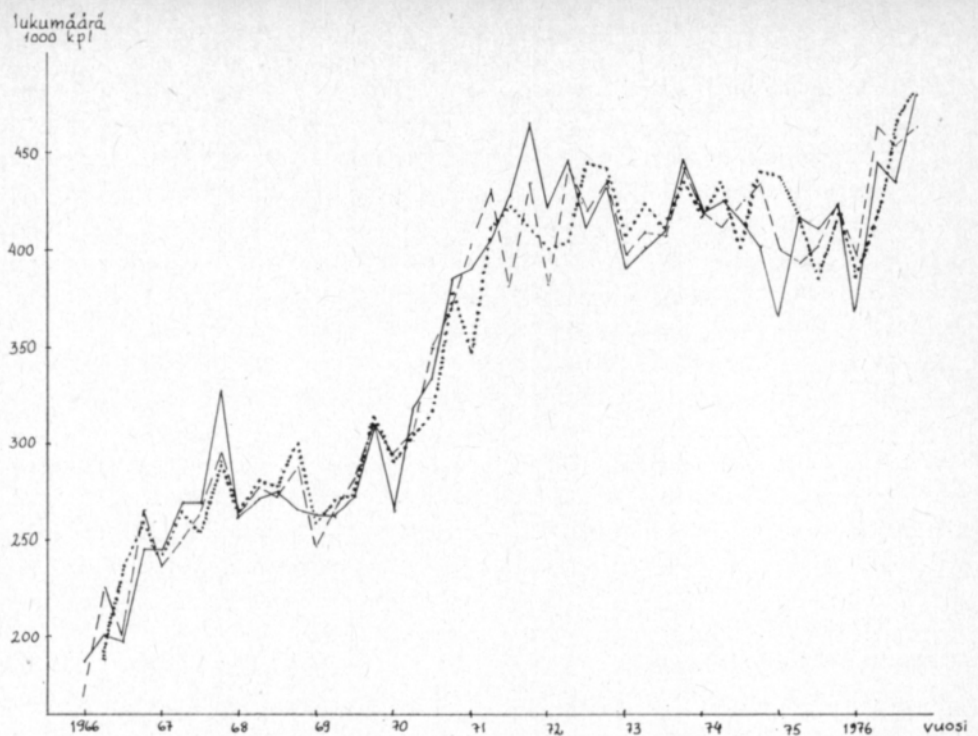
Muuttujat (lyhenne)	Perusmalli	Viivästetty malli
astutetut emakot + siat (AST)	t = 5.6	t = 3.8
rehuviljasato (SATO)	2.6	3.2
sianlihan hinta (P ^S)	0.6	1.1
Trendi	2.4	1.5
D _{II} (2. vuosineljänneksen dummy)	3.2	0.3
D _{III} (3. »)	2.1	0.4
D _{IV} (4. »)	4.8	1.5
	R ² = .94	R ² = .92
	S _e = 22'300	S _e = 24'500
	d = 1.7	d = ;1.6

Molempien mallien selitysasteet (R²) ovat varsin korkeat ja myös muut tilastolliset tunnusluvut ovat yleisesti ottaen tyydyttäviä. Perusmallissa tunnusluvut ovat järjestelmällisesti hiukan parempia kuin viiveen ollessa vuosineljänneestä pitempi, mutta perusmallin viive vastaakin paremmin astutuksen ja teurastuksen keskimääräistä aikaväliä.

Selittävien muuttujien korrelaatio on eräiltä osin (T-AST ja T-SATO) hälyttävän korkea. Tämän ei kuitenkaan käytännössä tarvitse johtaa virheisiin, koska selitysastekin on suuri. Residuaalien välistä korrelaatiota ei ole todettavissa Durbin—Watson (d-arvot) testillä, joskaan testi ei sulje pois korrelaation mahdollisuutta.

Mallista lasketut teurasmäärät poikkesivat perusmallissa keskimäärin n. 7 % todellisista määristä. Kuvasta 1 nähdään, että poikkeamat ovat lyhyellä aikavälillä paikoin varsin suuria, vaikka mallista laskettujen käyrien yleinen muoto seuraa hyvin todellista kehitystä ja vaikka malli on tilastollisilta ominaisuuksiltaan suhteellisen hyvä. Suurehkot poikkeamat osoittavat, että tästäkin mallista laskettuihin ennusteisiin tulee suhtautua varauksella. Vuositasolla suhteelliset poikkeamat ovat kuitenkin pienempiä kuin neljännesvuositain laskettuna. Alla poikkeamat on laskettu muutamalta vuodelta, olettaen että keskiteuraspaino olisi ollut 73 kg.

1972	-2.3 milj. kg
1973	+0,8 »
1974	+2.1 »
1975	-0.01 »
1976 (ennuste)	+3.5 »



Kuva 1. Teurastettujen lihotussikojen lukumäärä. Havaintoaineisto vuosilta 1965—75
 Figure 1. Fattening pigs slaughtered. Data period 1965—75.

- — — — — Malli, jossa viiveenä on 4 neljänestä
Model with 4 quarters' lag
- Malli, jossa viiveenä 5 neljänestä
Model with 5 quarters' lag
- Todelliset teurasmäärät
Realised slaughterings

Mallin muuttujien arviointi

Ongelmat kvantitatiivisen mallin kehittämissä liittyvät suurelta osin mallin ja todellisuuden suhteisiin. Puhtaasti luonnontieteellisissä malleissa tästä aiheutuvat ongelmat on helpompi ratkaista, mutta maataloustuotannon muutoksiin liittyvät yhteiskunnalliset kehityssuhteet monimutkaistavat tilannetta. Mallin muuttujia ei voida valita esim. pelkästään tilastollisten kriteerien pohjalta, vaan niille on asetettava ainakin seuraavat kolme vaatimusta.

Ensinnäkin selittävän ja selitettävän muuttujan välillä tulee todellisudessa vallita ainakin välillinen kausaalinen yhteys. Lisäksi muuttujien tulee muodostaa todellisuutta mahdollisimman hyvin vastaava kokonaisuus. Tässä mallissa perusosan muodostavaa biologista selittäjää täydentävät tuotantopäätöksiä selittävät muuttujat. Toiseksi muuttujien ja niistä muodostettavan mallin tulee täyttää tietyt tilastotieteelliset minimivaatimukset. Kolmanneksi muuttajia kuvaava aikasarja on kyettävä muodostamaan käytössä olevien tilastojen pohjalta. Tilastoihin sisältyvät virheet ja mahdolliset tilastoperusteiden muutokset voivat huonontaa ratkaisevasti päällisin puolin onnistuneenkin mallin tuloksia.

Eläinlaskentatilastosta saatavat *astutettujen emakoiden ja sikojen* lukumäärät täyttävät nämä vaatimukset varsin hyvin. Nämä määrähän kuvaavat jo toteutumassa olevia viljelijöiden tuotantopäätöksiä. Eläinlaskentatietojen käyttöä vaikeuttaa kuitenkin tilastoihin ajoittain, ennen otannan kehikon uusimista kertyvä, osin ilmeisesti systemaattinen virhe. Virhettä on voimistanut sikatalouden nopea rakennemuutos varsinkin 70-luvulla. Ekonometrinen malli asettaa käytettävälle tilastoaineistolle huomattavasti kovemmat vaatimukset kuin perinteisemmät tutkimusmenetelmät.

Toisena selittävänä muuttujana käytetty *rehusadon määrä* vaikuttaa luonnollisesti viljelijöiden tuotantopäätöksiin, koska sadon ollessa runsas ylijäämää pyritään purkamaan lihan tuotantoon. Näin ollen tämä muuttuja täyttää ensimmäisen edellä esitetystä vaatimuksista. Lisäsyö ottaa sadon *määrää* kuvaava muuttuja malliin oli se, ettei rehukustannuksia onnistuttu liittämään malliin *hintamuuttujan* muodossa. SATO-muuttujan kertoimen estimaatit ovat kummassakin mallissa myös tilastollisesti merkitseviä. Muuttuja on logaritmoitu, koska on katsottu, ettei sadon suuri kasvu lisää tuotantoa suhteellisesti yhtä paljon kuin pienempi kasvu.

Hintamuuttuja osoittautui testauksissa ongelmallisimmaksi. Kokeiltaessa Stojkavicin mallin tapaan sianlihan ja rehuohran hintasuhdetta selittävänä muuttujana oli etumerkki epälooginen. Sen mukaan tuotanto alenisi kun sianlihan hinta nousee suhteessa rehun hintaan. Edelleen yritettiin rehukustannusta saada malliin ohran tuottajahintaindeksin avulla, mutta etumerkki oli jälleen epälooginen. Muuttujien kertoimien estimaatit eivät olleet tilastollisesti merkitseviä kummassakaan tapauksessa.

Edellä esitettyyn perusmalliin otettiin hintamuuttujista mukaan vain sianlihan tuottajahinta, joka oli muutettu reaalisesti jakamalla se tuottajahintaindeksillä. Muuttuja kuvaa siis lähinnä sianlihan tuotannon »kannattavuutta» suhteessa muihin tuotantosuuntiin maatalouden sisällä. Tosin siinä ei ole otettu huomioon tuotannon rationalisoinnilla saavutettuja kannattavuuden muutoksia, jotka vaihtelevat huomattavasti eri tuotantosuunnissa ja eri aikoina. Edellä esitetystä t-arvoista näkyy, etteivät tällekkään muuttujalle estimoidut parametrit ole luotettavia. T-testissä estimaatit eivät osoitaneet edes »melkein merkitseviksi», vaikkakin etumerkit ovat molemmissa malleissa loogiset. Käytettäessä perusmallia vastaavassa mallissa vuosien 1965–72 havaintoja saatiin jonkin verran parempi t-arvo (ei kuitenkaan merkitsevä). Kun aineistona käytettiin vuosien 1970–76 havaintoja, muuttui jopa etumerkkikin epäloogiseksi. Näin ollen tilastollisia perusteita ei löydetty hintamuuttujan liittämiseksi malliin.

Tuottajahinnan kehityksellä ei siis tilastollisten kriteerien valossa näytä olevan vaikutusta tuotannon kokonaismäärään. Tämä on varsin yllättävä tulos siihen nähden, että Kettusen ja Kaarlehdon mallit nojasivat lähes yksinomaan hintamuuttujiin. Tämä tulos tukee sitä hypoteesia, ettei ainakaan koko tuotannonalaan tasaisesti kohdistuvilla hintapoliittisilla toimilla voitaisi sanottavasti vaikuttaa tuotannon kokonaismäärään. Hintakehityksellä voi tietenkin silti olla erilaisia vaikutuksia eri tyyppisiin tuotantoyksiköihin ja edelleen sillä voi olla erilaisia vaikutuksia eri ajankohtina tai samoja vaikutuksia eri pituisilla viiveillä eri aikoina ja eri tuottajaryhmiin.

Malliin on otettu myös sianlihan tuotannon varsin voimakasta kausivaihtelua kuvaavat *Dummy*-muuttujat sekä *Trendi*-muuttuja. *Dummy*-muuttujan kertoimet ovat varsin vakaat eri malleissa, joiden havaintoaineisto on eri ajanjaksoilta, mikä osoittaa kausivaihtelun vakiintuneisuutta. *Trendi*-muuttuja kuvanee lähinnä niitä tuotantoprosessissa tapahtuvia parannuksia, joita muut muuttujat eivät kykene selittämään.

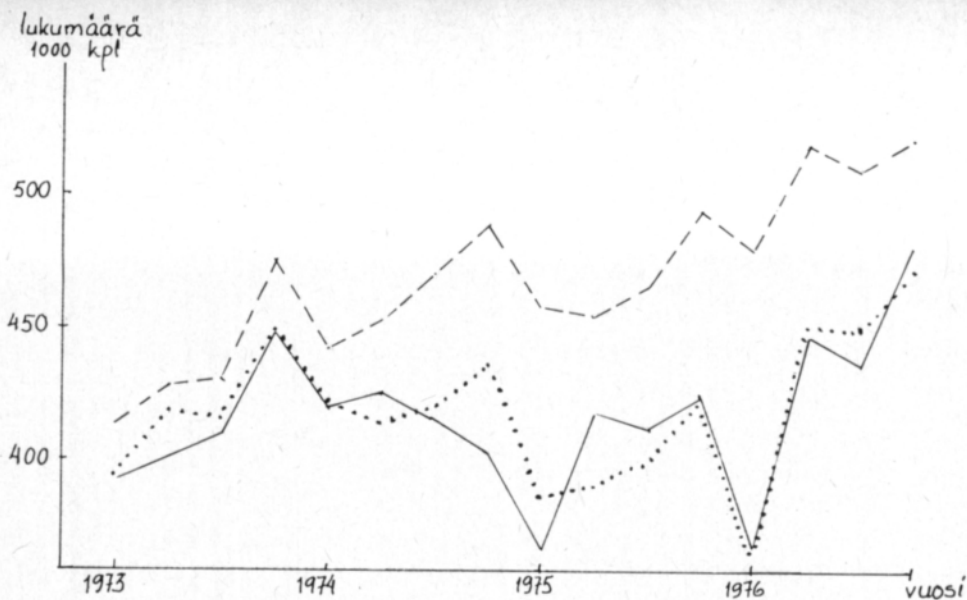
Mallin käyttökelpoisuus maatalouspoliittisessa suunnittelussa

Edellä esitelty malli sisältää jokaisen ennustemallin tapaan koko joukon ominaisuuksia, jotka rajoittavat sen käyttökelpoisuutta suunnittelussa. Ensinnäkin mallia laadittaessa joudutaan tekemään sellaisia oletuksia esim. muuttujavalinnoissa, jotka vanhenevat aikanaan. 60-luvun malleissa tuotantomuutoksia selitettiin mm. perunan hintakehityksellä, mutta tuskinpa kukaan enää nykyään niin tekisi. Muuttujien merkitys saattaa myös muuttua tai niiden vaikutusaika vaihdella havaintojaksolla ja sen jälkeen. Esim. astutuksen ja teurastuksen välinen aika vaihtelee ja taloudellinen ajattelutapa, johon tuotantopäätökset maataloilla perustuvat muuttuu ajan myötä

Toinen rajoitus johtuu siitä, että mallin kertoimet lasketaan aina tietyn havaintojakson kehityksen perusteella. Malli on näin sidoksissa menneeseen aikaan. Esim. sadon määrä saattaa olla suhteellisen vakaa havaintojaksolla, mutta jakson päätyttyä se heilahtaakin voimakkaasti. Koska havaintojaksoilla ei ole sattunut vastaavaa heilahdusta, ei muuttujan kerroin ole »varautunut» siihen ja tuloksena on todennäköisesti suhteettoman suuri muutos ennusteissa tuotantomäärissä. Lisäksi yhteen malliin ei voida teknisistäkään syistä liittää kaikkia mahdollisia asiaan vaikuttavia muuttujia, joten jokin malliin kuulumaton tekijä saattaa ratkaisevasti muuttaa tilannetta. Täysin mahdollonta olisi ollut esim. ennakolta laatia sellainen malli, joka olisi ottanut huomioon niiden tuotannon rajoittamiseen tähtäävien toimenpiteiden vaikutuksen, joita on alettu toteuttaa viime vuosina. Sellaiset muutokset on erikseen arvioitava sovellettaessa mallista laskettua ennustetta.

Ennustemallia voidaan luonnehtia konservatiiviseksi, sillä se varautuu vain sen tyyppiseen kehitykseen, jollaista havaintojaksolla on tapahtunut. Esimerkkinä esitetään kuvassa 2 kahden mallin perusteella lasketut tuotantotarvit. Molemmissa malleissa on samat, edellä esitettyyn perusmalliin sisältyvät muuttujat, mutta muuttujien kertoimet on laskettu eri havaintojaksojen tiedoista. Vaikka yhteistäkin aineistoa on käytetty 4 vuodelta, eroavat mallin 1 avulla laaditut ennusteet ja mallista 2 lasketut teurasmäärät huomattavasti toisistaan. Pääsyyinä on tässä tapauksessa se, että mallin 1 havaintojaksolla 1965—72 tuotanto on lisääntynyt voimakkaammin kuin neljää vuotta myöhemmällä havaintojaksolla mallissa 2. Näin ollen malli 1 ennustaa kasvun jatkuvan ja tuotannon muodostuvan todellista kehitystä suuremmaksi.

Pyrittäessä kehittämään eteenpäin esitetyn tyyppisiä ennustemalleja tulisi kiinnittää huomiota kolmeen osa-alueeseen, joihin edellä (jaksossa 3) on viitattu: mallin matemaattiseen kehittelyyn tilastollisten tunnuslukujen parantamiseksi, tilastotuotannon kehittämiseen ja kuvattavaa kohdetta koskevan teorian kehittämiseen. Nähdäkseni viimeksi mainitun tulisi saada nykyistä



Kuva 2. Teurastettujen lihotussikojen lukumäärä. Käyrät on laskettu samalla mallilla kuin kuvassa 1, mutta havaintojaksot poikkeavat 4 vuodella.

Figure 2. *Fattening pigs slaughtered.*

- - - - - Ennustemalli 1, aineisto vuosilta 1965-72
Prognosticatory model 1, data period 1965-72
- Ennustemalli 2, aineisto vuosilta 1969-76
Prognosticatory model 2, data period 1969-76
- Todelliset teurasmäärät
Realised slaughterings.

keskeisempi asema. Esimerkiksi hintamuutosten vaikutuksia testattaessa ilmeni, ettei ilmeisesti ole olemassa mitään yksinkertaista kausaalisuhdetta keskimääräisten hintojen ja kokonaistuotannon välillä. Tämän asian tarkempi selvittäminen ei tunnu mahdolliselta ilman perustutkimusta, jolla selvitetään tuotantopäätösten muodostumiseen liittyviä lainalaisuuksia yleensä ja erikseen eri viljelijäryhmissä.

Toisaalta olisi myös tarpeen tutkia niitä rajoja, jotka mallien ennustuskyylylle muodostuvat huolimatta kaikesta kehittäytyyöstä. Malliinhan voidaan liittää lähes loputtomasti uusia muuttujia tai matemaattisia hienouksia, mutta ennustuskyykyä ei niillä voida parantaa loputtomasti.

Erehtyvyys kuuluu ennustemallin luonteeseen. Tämä ei tietenkään merkitse sitä, etteikö ennustemalli voisi antaa pelkkää sormituntumaa parempaa pohjaa kehitysarvioita laadittaessa. Nähdäkseni lyhyen aikavälin tuotantoarvioiden luotettavuutta voidaankin merkittävästi lisätä käyttämällä apuna biologisiin muuttujiin perustuvia yksinkertaisia ekonometrisia malleja.

Yhteenveto

Maatilahallituksen suunnitteluosastolla on vuosina 1976 ja 1977 kehitelty sianlihan tuotannon lyhyen tähtäimen ennusteisiin soveltuvaa mallia. Malliin on selitettäväksi muuttujaksi valittu teurastettujen lihotussikojen määrä. Selitettävänä muuttujina ovat eläinlaskennasta saatava astutettujen ema-

koiden määrä, rehuviljasadon määrä, sianlihan reaalin hinta sekä trendi- ja kausimuuttujat. Vuosien 1965—1975 neljännesvuosiaineistosta lasketut tilastolliset tunnusluvut osoittautuivat verrattain hyviksi. Hintamuuttujan konstruoiminen muodostui kuitenkin erittäin ongelmalliseksi, mikä viittaa siihen, etteivät hintavaihteluiden tuotantovaikutukset ole lainkaan yksiselitteisiä, vaikka eräät aikaisemmat ennustemallit ovat keskeisesti rakentuneet hintamuuttujien varaan.

Pyrittäessä soveltamaan tutkitun kaltaista mallia ennustekäyttöön kohdataan useita ongelmia, joista merkittävimiksi osoittautuivat tässä yhteydessä tilastoaineiston asettamat rajoitukset, mallin ulkopuolelle pakosta jäävien tekijöiden, mm. talouspoliittisten toimenpiteiden vaikutus ja mallin »konservatiivisuus», ts. mallin kertomien verraten nopea vanheneminen, mikä johtaa kumulatiivisiin virheisiin ennusteissa.

KIRJALLISUUTTA

- ANON. 1972. Sianlihan tuotannon ennustaminen. Markkinakatsausryhmä 1972: 1—6.
- IKÄHEIMO, E. & ROUHIAINEN, J. 1974. Sianlihan tarjonnan ekonometrinen malli. Tilastotiet. prosemin. Moniste. 8 p.
- KAARLEHTO, P. 1959. Sianlihan markkinoinnista Suomessa. Maatal. Taloud. Tutk.lait. Julk. 2: 1—68.
- KETTUNEN, L. 1968. Demand and supply of pork and beef in Finland. Maatal. Taloud. Tutk.lait. julk. 11: 1—84.
- * — 1974. Maataloustuotannon ennakointi. Maatal.tiet. Aikak. 47: 82—91.
- STOJCOVIC, G. 1965. Modeller för marknadsprisbildningen. Markn. ekon. Semin. 1965: 68—88.
- SVENSSON, B.-O. 1976. Svensk svinproduktion. Rapp. Inst. Ekon. Statist. 83: 13—43.

Käsikirjoitus saapunut 2. 1. 1978